SKP Engineering College

Tiruvannamalai – 606611

A Course Material

on

Cognitive Radio

By

S.Baskaran

Professor

Electronics and Communication Engineering Department

Quality Certificate

This is to Certify that the Electronic Study Material

Subject Code:EC 6014

Subject Name:Cognitive Radio

Year/Sem:IV/VII

Being prepared by me and it meets the knowledge requirement of the University curriculum.

Signature of the Author

Name: S.Baskaran

Designation: Professor

This is to certify that the course material being prepared by Mr. S.Baskaran is of the adequate quality. He has referred more than five books and one among them is from abroad author.

Signature of HD

Name:R.Saravankumar

Seal:

Signature of the Principal

Name: Dr.V.Subramania Bharathi

Seal:

EC6014 COGNITIVE RADIO L T P C 3 0 0 3

OBJECTIVES: The student should be made to:

- ☐ Know the basics of the software defined radios.
- ☐ Learn the design of the wireless networks based on the cognitive radios
- ☐ Understand the concepts of wireless networks and next generation networks

UNIT I INTRODUCTION TO SOFTWARE DEFINED RADIO 9

Definitions and potential benefits, software radio architecture evolution, technology tradeoffs and architecture implications.

UNIT II SDR ARCHITECTURE 9

Essential functions of the software radio, basic SDR, hardware architecture, Computational processing resources, software architecture, top level component interfaces, interface topologies among plug and play modules,.

UNIT III INTRODUCTION TO COGNITIVE RADIOS 9

Marking radio self-aware, cognitive techniques – position awareness, environment awareness in cognitive radios, optimization of radio resources, Artificial Intelligence Techniques.

UNIT IV COGNITIVE RADIO ARCHITECTURE 9

Cognitive Radio - functions, components and design rules, Cognition cycle - orient, plan, decide and act phases, Inference Hierarchy, Architecture maps, Building the Cognitive Radio Architecture on Software defined Radio Architechture.

UNIT V NEXT GENERATION WIRELESS NETWORKS 9

The XG Network architecture, spectrum sensing, spectrum management, spectrum mobility, spectrum sharing, upper layer issues, cross – layer design.

TOTAL: 45

PERIODS OUTCOMES: Upon completion of the course, students will be able to

☐ Describe the basics of the software defined radios.

☐ Design the wireless networks based on the cognitive radios

☐ Explain the concepts behind the wireless networks and next generation networks

TEXT BOOKS:

1. Joseph Mitola III,"Software Radio Architecture: Object-Oriented Approaches to Wireless System Engineering", John Wiley & Sons Ltd. 2000.

2. Thomas W.Rondeau, Charles W. Bostain, "Artificial Intelligence in Wireless communication", ARTECH HOUSE .2009.

3. Bruce A. Fette, "Cognitive Radio Technology", Elsevier, 2009.

4. Ian F. Akyildiz, Won – Yeol Lee, Mehmet C. Vuran, Shantidev Mohanty, "Next generation / dynamic spectrum access / cognitive radio wireless networks: A Survey" Elsevier Computer Networks, May 2006.

REFERENCES:

1. Simon Haykin, "Cognitive Radio: Brain –Empowered Wireless Communications", IEEE Journal on selected areas in communications, Feb 2005.

2. Hasari Celebi, Huseyin Arslan, "Enabling Location and Environment Awareness in Cognitive Radios", Elsevier Computer Communications , Jan 2008.

3. Markus Dillinger, Kambiz Madani, Nancy Alonistioti, "Software Defined Radio", John Wiley, 2003.

4. Huseyin Arslan, "Cognitive Radio, SDR and Adaptive System", Springer, 2007.

5. Alexander M. Wyglinski, Maziarnekovee, Y. Thomas Hu, "Cognitive Radio Communication and Networks", Elsevier, 2010.

CONTENTS

S.No	Particulars	Page
1	Unit – I	6
2	Unit – II	41
3	Unit – III	70
4	Unit – IV	98
5	Unit – V	135

Unit - I

Introduction to Software Defined Radio

Part – A

1. **What is Software-Defined Radio? (L-1, CO-1)**

   The International Telecommunication Union (ITU) has proposed a definition of SDR as a "radio in which the operating parameters including inter alia frequency range, modulation type, and/or output power limitations can be set or altered by software".

2. **Evolution of Software-Defined Radio (L-1, CO-1)**

   Tier 0 - describes hardware based radios, and is actually not considered to fall into the realm of SDR.

   Tier 1 - The simplest SDR technology begins with Tier 1, which describes software controlled radios (SCR) with only the control functions being processed by software.

- The simplest example to this is a dual mode cell phone, which consists of two hardware radios for two different standards.
- The software simply controls which radio should be utilized.

Tier 2 - Reconfigurable software defined radios present Tier 2. SDR systems include reconfiguration by allowing control over modulation techniques, security functions (such as frequency hopping) and waveform requirements over a broad frequency range provided by software. Tier 2 SDRs include processing applications such as

- Application-specific integrated circuits (ASIC),
- Field-programmable gate arrays (FPGA) and
- Digital signal processors (DSP).

Although reconfigurable SDRs are the most commonly used systems today, especially for military applications, due to the rapid sophistication of the general SDR technology these systems become increasingly obsolete.

Tier 3 - Software defined radios, also called ideal software radios (ISR), will eventually become the mostly implemented systems within the near future. Based on the extended possibilities of programmability to the entire system, analog conversion will be completely realized only by the antenna, microphones and speakers.

Tier 4 -The SDR Forum declares that ultimate software radios (USR) as Tier 4 technologies “are defined for comparison purposes only”. In theory, these USRs are supposed to be capable of supporting a broad frequency range, air-interfaces and applications, allowing switching between air-interface formats and different applications within only milliseconds

3. **List out the potential benefits of SDR. (L-3, CO-1)**

SDR concept started in the late 1970s with the introduction of multimode radios operating in VHF band

• U.S. Air Force Avionics Laboratory initiated the Integrated Communication, Navigation, Identification and Avionics (ICNIA) program in the late 1970s

– Developed an architecture to support multifunctional, multiband airborne radios in the 30 MHz -1600 MHz band

– Successful flight test and final report delivery in 1992 – ICNIA radio was the first programmable radio

• In the late 1980s, the Air Force Research Laboratory initiated the Tactical Anti-Jam Programmable Signal processor (TAJPSP)

– Developed a processor capable of simultaneous waveform operations using modular approach

– TAJPSP later evolved into the SPEAKeasy program

• SPEAKeasy was a joint U.S. Government program to develop the architecture and technology to meet future military requirements for multimedia networking operations

– The first significant military investment to integrate various existing radio families into one family

– COTS-based architecture

– Demonstrated multiband, multimode radio capabilities in 1998

– SPEAKeasy evolved into the Joint Tactical Radio System (JTRS)

• JTRS Joint Program Office was established in 1999

4. **Define dynamic spectrum access. (L-2, CO-1)**

Dynamic spectrum access is a new spectrum sharing paradigm that allows secondary users to access the abundant spectrum holes or white spaces in the licensed spectrum bands. DSA is a promising technology to alleviate the spectrumscarcity problem and increase spectrum utilization.

5. **What is the role of spectrum policy? (L-3, CO-1)**

software radios may operate on any RF band that is within the capabilities of the underlying radio platform, and with any mode for which a software load-image is available. This raises the possibility of truly novel ap- proaches to spectrum management

6. **What are the tradeoffs required in SDR? (L-1, CO-1)**

Step 1 - Antenna Tradeoffs - The choice of antennas (step 1 in the figure)

Step 2 - **RF and IF Processing Tradeoffs -** determines the number and bandwidth of RF channels

Step 3 - **ADC Tradeoffs -** Constrains the numbers and bandwidths of ADCs

Step 4 - **Digital Architecture Tradeoffs -** Additional parallel IF processing and ADC paths may be necessary to support multiple-service bands simultaneously. The ADCs provide high-speed streams for heterogeneous multiprocessing

Step 5 - **Software Architecture Tradeoffs - -** General-purpose processors yielding a multithreaded, multitasking, multiprocessing operating environment. Software objects must be organized into real-time objects

Step 6 - **Performance Management Tradeoffs -** The effective hosting of these objects onto this complex operating environment requires a refined set of techniques unique to this text called *SDR performance management*

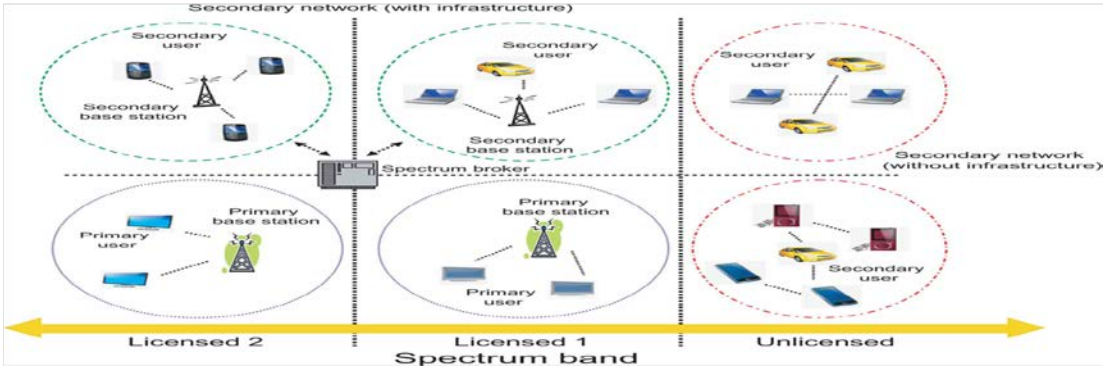
7. **What is agile radios? (L-3, CO-1)**

   In order to support new services for some users, the networks will have to become more agile in their ability to tailor content for disadvantaged users (low-data- rate users; those in severe jamming environments; those with few batteries, etc.).

8. **Define the term data explosion. (L-2, CO-1)**

   The information **explosion** is the rapid increase in the amount of published information or **data** and the effects of this abundance. As the amount of available**data** grows, the problem of managing the information becomes more difficult, which can lead to information overload.

9. **Draw the cognitive radio bands. (L-2, CO-1)**



10. **What are the security aspects of cognitive radio? (L-2, CO-1)**

    CR technology is the "intersection of personal wireless technology and computational intelligence," where CR is defined as "**a really smart radio that would be self-aware, RF-aware, user-aware, and that would include language technology and machine vision along with a lot of high-fidelity knowledge of the radio environment**

11. What are the potential benefits of SDR? (L-2, CO-1)

POTENTIAL BENEFITS

The militaristic developers hoped for the accomplishment of several benefits that a software based system could provide. These **potential benefits included**:

- **Interoperability –** Support of multiple standards through multimode, multiband radio capabilities
- **Flexibility –** Efficient shift of technology and resources
- **Cost reduction –** Less infrastructure, less maintenance, easier deployment
- **Adaptability –** Faster migration towards new standards and technologies through programmability and reconfiguration
- **Responsiveness -** To allow quick and easy incorporation of future developments

12. Draw the road map for SDR Development (L-1, CO-1)

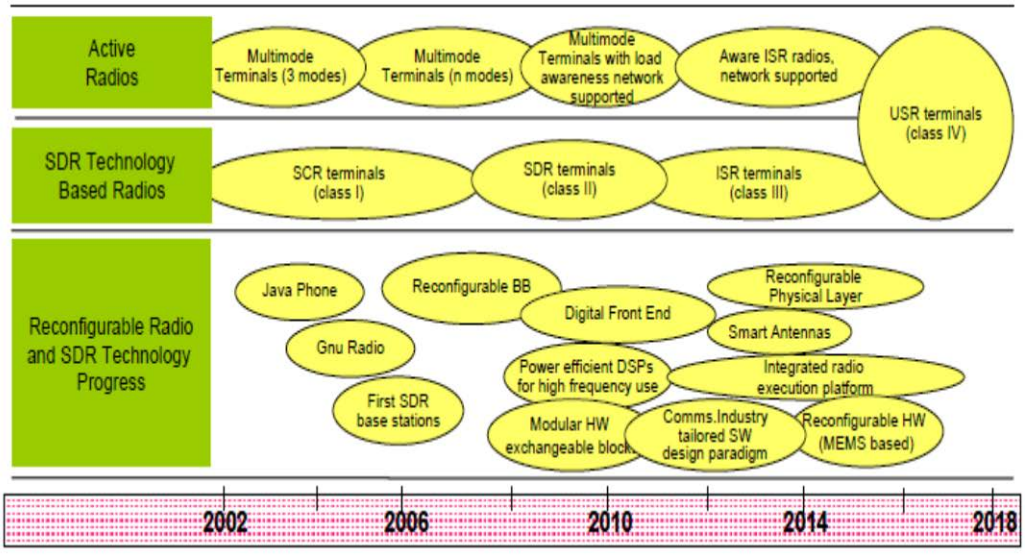


Figure 5-1: Roadmap for SDR Deployment

13. What **is the antenna design considerations in cognitive radio systems? (L-2, CO-1)**

The antenna characteristics determine not only the gain due to aperture effects, but also several critical characteristics of the SDR, including:

- The number of antenna channels required to support multiband multimode operation
- Usually, the number of parallel RF conversion chains
- Often, the number of ADCs and DACs required Parallelism is a major cost driver for software radios.
- Higher-gain antennas achieve this gain over relatively small segments of RF

14. Mention the types of antenna's used in CR(L-2, CO-1)

1. Yagi uda
2. Microchip
3. Dish Antenna
4. Reconfigurable antenna

15. Define prototyping. (L-1, CO-1)

The **Prototyping** Model is a systems development method (SDM) in which a **prototype** (an early approximation of a final system or product) is built, tested, and then reworked as necessary until an acceptable **prototype** is finally achieved from which the complete system or product can now be developed.

16. What are the goals of RF Targeoffs? (H-2, CO-1)

The goal of this tradeoff is

- to balance the noise,
- spurious components,
- inter modulation products, and
- Artifacts.

17.Draw the tradeoff for Digital architecture. (L-1, CO-1)

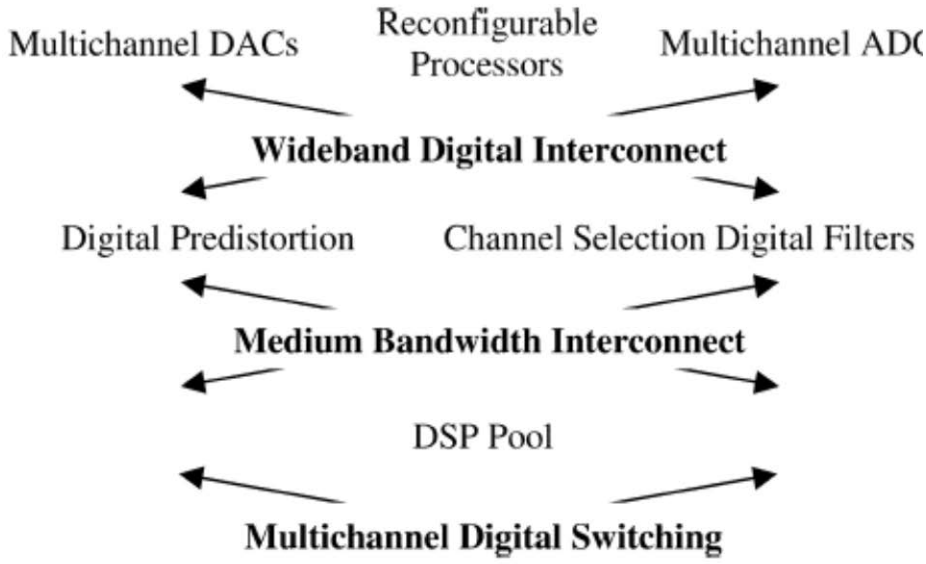


Figure 6-4 Digital architecture tradeoffs.

18..What are the implications in various architecture levels. (L-1, CO-1)

ARCHITECTURE IMPLICATIONS

An architecture implication defines various architecture levels used in SDR to perform multimode processing like DSP, FPGA and ASICs architectures.

- Programmable DSPs have the required high order language (HOL) programmability, but they are inappropriate for frontend filtering tasks.
- These DSPs would have consumed much higher power than the Harris ASICs that were eventually chosen.
- The DSPs were more appropriate for channel modem and baseband signal processing tasks.
- FPGAs were also considered for front-end filtering.
- VHDL programmability provided flexibility, but these chips were also less power efficient for finite impulse response (FIR) filters and FFTs than dedicated ASICs.
- ASICs are most power efficient,

- The mix of ASICs, FPGAs, and DSPs reproduce different types of modules. It also reduces the number of backup modes.

19. What is meant by roofing filter? (L-1, CO-1)

A "**Roofing filter**" is simply a filter in the radio's first IF through which all signals must pass before they will be "seen" by later receiver stages.

20. Explain the significance of Performance Management Tradeoffs. (L-2, CO-1)

The final major tradeoff concerns the management of processing demand offered by the software against the resources provided by the hardware platform. Accurate characterization of processing demand requires benchmarking. Accurate prediction (e.g., at proposal time), can be accomplished using queuing theory techniques that have been refined and reduced to the structured method.

PART – B

1) **What is software radio? Give the essential functions of Software Radio. (H-2, CO-1)**

Multiband technology first of all, accesses more than one RF band of Communications channel at once. The RF channel then is generalized to the channel set .This set includes RF channels, but radio nodes like PCS base stations and portable military radios also interconnected to fiber and cable; therefore these are also included in the channel set. The channel encoder of a multiband radio includes RF/channel access, IF processing, and modem. The RF/channel access includes wideband antennas, and the multi-element arrays of smart antennas. This segment also provides multiple signal paths and RF conversion that span multiple RF bands. IF processing may include filtering, further frequency translation, space/time diversity processing, beam- forming, and related functions. Multimode radios generate multiple air interface waveforms (modes) defined principally in the modem, the RF chan- nel modulator-demodulator.

These waveforms may be in different bands and may span multiple bands. A software-defined personality includes RF band, channel set (e.g., control and traffic channels), air interface waveforms, and related functions.

Although many applications do not require information security (INFOSEC), there are incentives for its use. Authentication reduces fraud. Stream encipher- ment ensures privacy. Both help ensure data integrity. Transmission security (TRANSEC) hides the fact of a communications event (e.g., by spread spec- trum techniques [4]). INFOSEC is therefore included in the functional model although the function may be null for many applications.

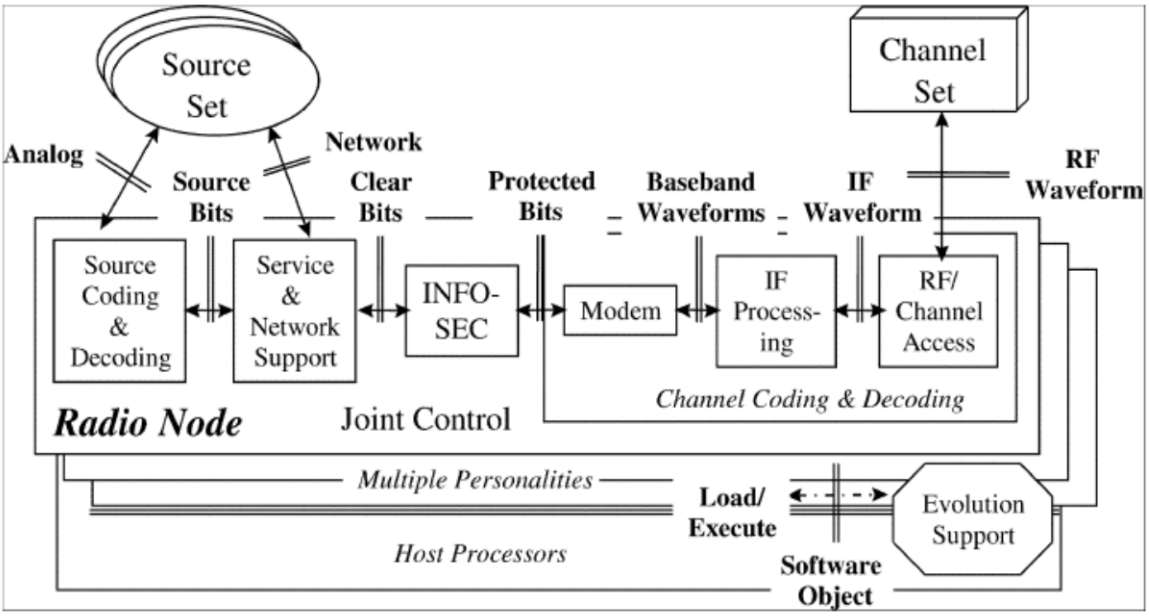
In addition, the source coder/decoder pair now includes the data, facsimile, video, and multimedia sources essential for new services. Some sources will be physically remote from the radio node, connected via the synchronous digital hierarchy (SDH), a local area network (LAN) , etc., through service and network support .

These functions may be implemented in multithreaded multiprocessor soft- ware orchestrated by a joint control function. Joint control ensures system stability, error recovery, timely data flow, and isochronous streaming of voice and video. As radios become more advanced, joint control becomes more complex, evolving toward autonomous selection of band, mode, and data for- mat. Any of the functions may be singleton (e.g., single band versus multiple bands) or null, further complicating joint control. Agile beamforming sup- ports additional users and enhances quality of service (QoS) . Beamform- ing today requires dedicated processors, but in the future, these algorithms may time-share a DSP pool along with the Rake receiver [8] and other mo-dem functions. Joint source and channel coding also yields computation- ally intensive waveforms. Dynamic selection of band, mode, and diversity as a function of QoS introduces large variations into demand, potentially causing conflicts for processing resources. Channel strapping, adaptive wave- form selection, and other forms of data rate agility further complicate the statistical structure of the computational demand. In addition, processing re- sources are lost through equipment failures. Joint control integrates fault modes, personalities, and support functions on processing resources that

in- clude ASICs, FPGAs, DSPs, and general-purpose computers to yield a reliable telecommunications object . In a software radio, the user can upload a variety of new air interface per- sonalities . These may modify any aspect of the air interface, including whether the waveform is hopped, spread, or otherwise constructed. The re- quired resources (e.g., RF access, digitized bandwidth, memory, and process- ing capacity) must not exceed those available on the radio platform. Some mechanism for evolution support is therefore necessary to define the wave- form personalities, to download them (e.g., over the air) and to ensure that each new personality is safe before being activated. The evolution-support function therefore must include a software factory. In addition, however, the evolution of the radio platform—the analog and digital hardware of the radio node—must also be supported. This may be accomplished via the design of advanced hardware modules in an integrated evolution support environment, or by the acquisition of commercial off-the-shelf (COTS) hardware modules, or both. The block diagram of the radio functional model amounts to a partitioning of the black-box functions of the ideal software radio nodes introduced above into the specific functional components shown in Table 1-1.

Functional Component	Allocated Functions	Remarks
Source Coding and Decoding	Audio, data, video, and fax	Ubiquitous algorithms (e.g., ITU, ETSI)
Service and Network Support	Multiplexing; setup and control; data services; internetworking	Wireline and Internet standards including mobility
Information Security	Transmission security, authentication, no repudiation, privacy, data integrity	May be null, but is increasingly essential in wireless applications
Channel Coding and Decoding: Modem	Baseband modem, timing recovery, equalization, channel waveforms, predistortion, black-data processing	INFOSEC, modem, and IF interfaces are not yet well standardized
IF Processing	Antenna, diversity, RF conversion	IF interfaces are not standardized
RF Access	Simultaneity, multiband propagation,	Automatically employ

	wireline interoperability	multiple channels or modes for managed QoS
Multiple Personalities	Multiband, multimode, agile services, interoperable with legacy modes	Multiple *simultaneous* personalities may cause considerable RFI
Evolution Support	Define and manage personalities	Local or network support software factory
Joint Control	Joint source/channel coding, dynamic QoS vs. load control, processing resource management	Integrates user and network interfaces; multiuser, multiband, and multimode capabilities

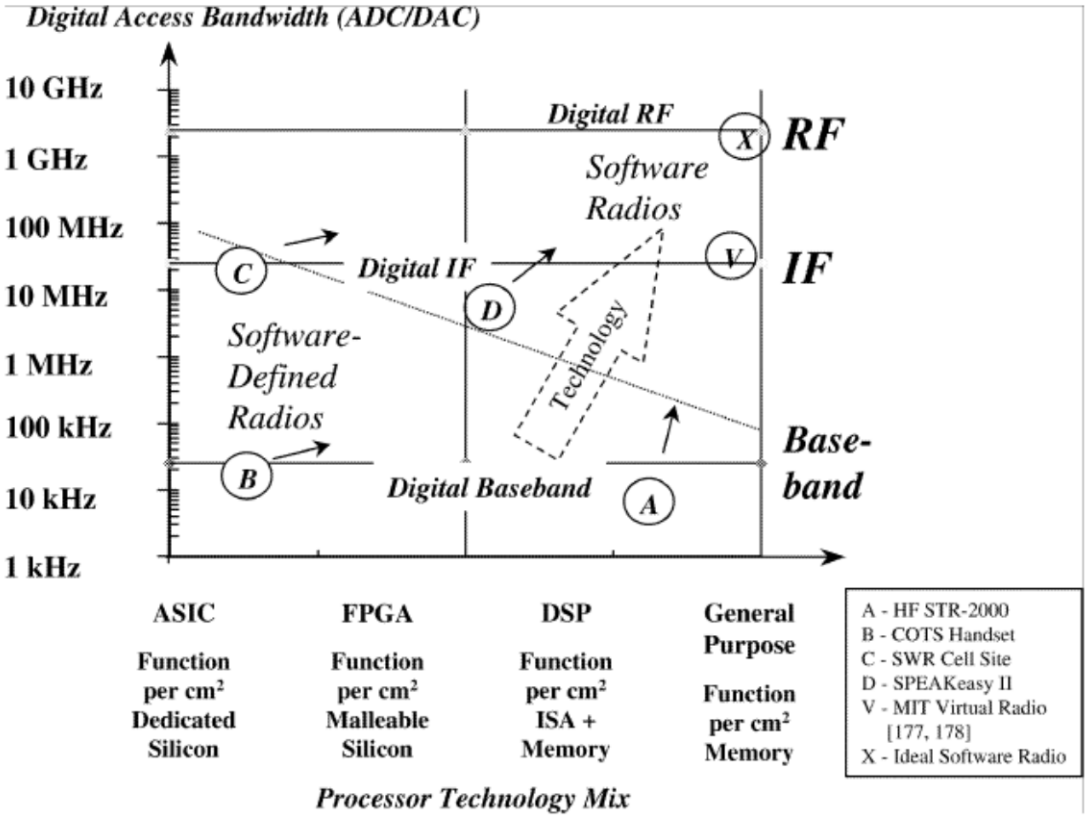


2) Explain the architecture of SDR with neat diagrams and its implications. (H-2, CO-1)

Implementation alternatives for digital radios, SDR, and software radios may be characterized in the software-radio phase space of Figure 1-7. The phase space

compares digital-access bandwidth to the flexibility of the processing platform. These are the two most critical architecture parameters of the soft- ware radio. Digital-access bandwidth is approximately half of the sampling rate of the ADC in the isochronous subscriber signal-processing path. Thus, for example, a 5 GHz conversion rate supports nominally a 2.5 GHz analog bandwidth, based on the Nyquist criterion [20]. ADCs with bandwidths of over 6 GHz exist [38], so digitizing RF is not impossible. If all the processing after the ADC were accomplished on a single general-purpose computer, one would have an ideal software radio receiver (the point marked X in the figure). Corresponding digital signal synthesis and up-conversion would yield an ideal software-radio transmitter.

Such extremely wideband ADCs consume substantial power and have a dynamic range of only about 30 dB. These limitations preclude practical im- plementations of the ideal. In addition, the digital filtering of the 5 giga-sample per second stream to access a given RF band such as 25 MHz of RF spectrum



This feat is beyond general-purpose computers. Furthermore, there is no single antenna or RF stage that can sustain the analog bandwidth from 2 MHz to 2.5 GHz required as input to the ADC (and conversely for the transmitter). Thus, the ideal software radio is not implementable with today's technology. Why even include it? The ideal software radio represents the point of maximum flexibility for a radio platform. The ideal properties of such a radio represent the best that one could ever achieve, and thus are a useful reference point for measuring progress toward generality and flexibility.

Practical implementations have limited RF coverage due to the narrow- band nature of antennas, RF conversion, and IF processing technology. They also require a mix of digital technologies including ASICs, FPGAs, DSP, and general-purpose processors. The STR-2000 (point A in the figure) was an early baseband HF DSP radio developed by Standard Marine AB. This radio digitized an HF IF signal at a 24 kHz sampling rate. It used twin Texas In- struments (TI) TMS320C30 DSPs to provide a half-dozen standard

HF signal formats digitally. This could be accomplished using a general-purpose processor today. COTS handsets (B) minimize size, weight, and power through the use of ASICs. Some handsets demodulate signals in an RF ASIC that creates a digital baseband bitstream directly from analog RF. Combining two such ASICs in a handset enclosure leads to the term "Velcro radio" for this approach .

MPLEMENTATION ALTERNATIVES

Contemporary software-radio cell-site designs (C) access the allocated up- link[8] RF using a single ADC (e.g., with 25 MHz of analog bandwidth; viz., 70 MHz conversion rate). These designs employ a bank of digital filter ASICs [40] or parallel digital filters [41] to access a hundred or more subscriber channels in parallel. Some implementations incorporate the new high-density FPGAs to provide software-driven configurability in a delivery platform that maximizes throughput for a given technology clock rate [42]. Technologically aggressive designs include SPEAKeasy [3], the military technology pathfinder. SPEAKeasy II (point D in Figure 1-7), which became the baseline for Mo- torola's WITS 6000 software radio product line [43], incorporated over a GFLOPS of processing capacity for enhanced flexibility. The Virtual Radio (point V in the figure) is the most flexible software radio research implemen- tation reported in the literature [44]. A general-purpose DEC Alpha processor running UNIX accesses a wideband IF digitally. Narrowband AM and FM broadcast receivers and an RF LAN have been implemented purely in soft- ware on this platform.

The three fundamental limitations of any SDR implementation, then, are:

1. RF access
2. Digital access bandwidth
3. Digital processing (flexibility and capacity)

The process of plotting an implementation in the software-radio phase space

is illuminating. Those that are further to the right *should* be more flexible and easier to extend. But this is sometimes not the case. Systems and software engineering disciplines described in this text are required to capitalize on the flexibility of the hardware. These techniques must be fully employed and systematically practiced throughout the system life cycle. The software design and development process chapters of this book show how to make the touted flexibility a reality. These design techniques also allow one to avoid disaster as a sequence of apparently small incremental requirements added to a simple, stable system yield an unstable "house of cards." Definition of a radio platform is one of the steps that is required to avoid such disasters.

Defining the Radio Platform

One key architecture question is the degree of programmability required for
the intended market niche. Contemporary radio designs therefore vary across the dotted line in the phase space. This represents the technology frontier, comprising a mix of ASIC, FPGAs, DSP, and general-purpose processing elements using ADCs and DACs at baseband or IF. Aggressive designs move above and to the right of this line, while conservative designs remain below and to the left. Advancing microelectronics technology moves all implementations

Software Radio Reference Platform Parameters

Critical Parameter	**Remarks**
Number of Channels	Number of parallel RF, IF, and /or baseband channels
RF Access	Continuous coverage from a minimum to a maximum RF
Digital Bandwidth	Bandwidth of the maximum ADC for each RF/IF channel
Dynamic Range	End to end, including RF, IF, ADC, AGC, and processing

inexorably upward and to the right over time. At present, handsets—even dual-mode handsets—favor the Velcro approach using RF ASICs with chip- level integration. Some implementations use the VME or PCI bus to facilitate board-level upgrades. Some applications such as law enforcement and general aviation radios are very cost sensitive and therefore typically lag the state of the art by one or two generations (2-8 years).

With such a variety of RF, ADC, and processing hardware implementations, it is extremely difficult to determine whether third-party software intended for one platform will be of any use on another. To address this question quanti-
tatively requires the following steps:

1. Definition of a radio reference platform
2. Characterization of the software processing demand in standard metrics 3. Control of critical hardware and software parameters during development and operations

A radio reference platform is a high-level characterization of the capabilities of the hardware environment of the software radio. Table 1-3 identifies the most critical radio platform parameters that determine the performance of a software radio

3) Discuss in detail about the potential benefits and technology in antenna tradeoff. (L-2, CO-1)

The antenna segment establishes the available RF bands. Although much research has been applied toward creating an "all-band" antenna, multiband radios generally require at least one antenna per decade of RF band (e.g., HF, VHF, UHF, SHF, etc.). In addition, the antenna determines the directional properties of the receiving system. Sectorized antennas, static beamforming arrays, and adaptive beamforming arrays (smart antennas) each have different spatial and temporal properties, the most significant of which is the pattern of transmit and/or receive gain. The antenna may also constrain the phase noise of the overall system. In addition, the interface between the antenna and the RF conversion stage determines VSWR, insertion loss, and other miscellaneous losses. In bands above 100 MHz, this interface can determine the overall

sys- tem noise floor. This chapter characterizes the systems-level antenna segment tradeoffs relevant to SDR architecture.

From a SDR perspective, the enabling RF-access parameters of the antenna segment are RF band and bandwidth as illustrated in Figure 7-1. Antenna- type in the figure lists the mechanical structure and the physical principle on which the antenna is based. Bandwidth is expressed either as a percent of carrier frequency or as a ratio of lowest RF to highest RF over which the antenna efficiency, VSWR, etc. are workable. Narrowband antennas have only a few percent relative bandwidth. Frequency limits are typically defined in terms of the 3 dB bandwidth of the antenna. An HF antenna, for example, that is operable between 2 MHz and 20 MHz has a relative bandwidth of 20/2 or 10 : 1. An antenna that operates effectively between 2 and 4 GHz, on the other hand, has a relative bandwidth of only 2 : 1. This ratio is one octave. Wideband antennas such as log periodic and equiangular spirals require a large number of resonant elements and therefore have a relatively high cost compared to narrowband resonant antennas. Helical antennas may be wound into whip or stub mechanical configurations for PCS applications

For the ideal software radio, one needs a single antenna element that spans all bands. Requirements of the JTRS program are illustrated in Figure 7-2a. More than forty bands and modes must be supported in that program. With conventional technology, nine or ten antenna bands would be required as shown in the figure. Anticipating the JTRS program, SPEAKeasy attempted to realize a full-band antenna. The RF range extended from 2 MHz to 2000 MHz, a ratio of 1000 : 1 or 3 decades. Figure 7-1 shows that this requires a tech- nology breakthrough, since the maximum relative bandwidth of the well-established designs is 10 : 1, or one decade. Through in-depth antenna studies conducted by Rockwell, Hazeltine, and others, it was determined that at least 3 bands are needed for this range. In fact, SPEAKeasy employed three bands as follows: (a) 2-30 MHz; (b) 30-400 MHz; and (c) 400-2000 MHz. To be precise, only band b was fully implemented in SPEAKeasy I and only bands a and b were implemented in SPEAKeasy II. For the foreseeable future, affordable RF access will probably be limited to octave

coverage in the bands above 100 MHz. One configuration of antenna coverage that employs four conservatively designed bands is illustrated.

PARAMETER CONTROL

From a systems-engineering perspective, one must allocate end-to-end per- formance to parameters of the appropriate segment. The use of wideband antennas that enable SDR levels of performance complicates the control of SNR, timing, and phase parameters as follows.

Linearity and Phase Noise

Wide bandwidth is sufficient for detection, but high SNR is necessary for good SDR algorithm performance. As the antenna bandwidth is increased, the thermal noise power increases linearly. Thus, the antenna channels must be filtered to select only those subsets of the band required to service subscriber signals. This is accomplished in the RF conversion and digital IF processing segments.

Low phase noise is also critical for phase-sensitive channel modulations such as high-order QAM (> 16 states). Phased array antennas that form beams through the switching of delay elements can have high phase noise induced by switching transients, making high-order QAM impractical.

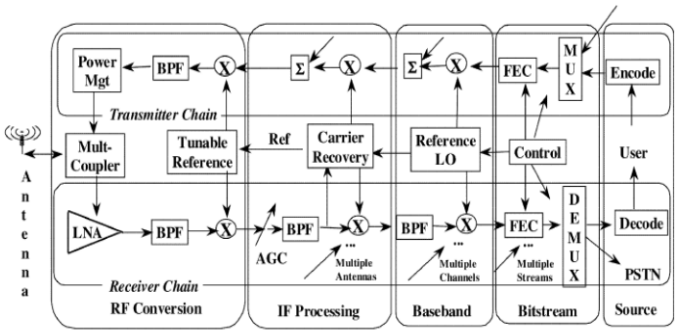
Parameters for Emitter Locations

In addition, precision timing or RF phase control may be necessary. For ex- ample, the commercial sector now has requirements for the location of mo- bile stations from which emergency calls are placed. The US E911 service requires location to within 125 meters. Network-based emitter location tech- niques include time-difference of arrival (TDOA) and angle of arrival (AOA) estimation using phase interferometry. TDOA [218] requires timing precision on the order of 100 ns, systemwide, to meet E911 requirements. Similarly, AOA [219, 220] requires phase measurements equivalent to a

few degrees of angle uncertainty, which is equivalent to a few electrical degrees of phase error.

Smart antennas generally derive some estimate of the direction-manifold of the received signals. This information can be translated into AOA. In ad- dition, TDOA techniques may be used alone or in conjunction with smart antennas to estimate the location of mobile subscribers. TDOA is particularly

4) Discuss in detail about the potential benefits and technology in IF/RF tradeoffs. (H-2, CO-1)



. The antenna segment may provide a single element for both trans- mission and reception. In this case, a multicoupler, circulator, or diplexer pro- tects the receiver from the high-power transmission path. In other cases, the transmit and receive antennas may be physically separate and may be sepa- rated in frequency. First-generation cellular radio and GSM systems separate downlink and uplink bands by typically 45 MHz to limit interference.

The transmission subsystem intersects the RF conversion segment as shown in Figure 8-1. This includes a final stage of up-conversion from an IF, band- pass filtering to suppress adjacent channel interference, and final power am- plification. First-generation cellular systems did not employ power control to any significant degree. CDMA systems, including third-generation (3G) W-CDMA, require power control on each frame (50 to 100 times per sec- ond). SDRs may be implemented with a DAC as the interface between IF up-conversion and the RF segment. Alternatively, a high-speed DAC may di-rectly feed the final power amplifier.

Power amplifiers have less-than-ideal performance, including amplitude ripple and phase distortion. Although these effects may be relatively small, failure to address them may have serious consequences on SDR performance. Amplitude ripple, for example, degrades the transmitted power across the band, particularly near the band edges. IF processing may compensate by preemphasizing the IF signal with the inverse of the power amplifier's band- edge ripple. Feher [238] describes techniques for compensating a sequence of channel symbols, shaping the transmitted waveform in the time domain to yield better spectral purity in the frequency domain. The concept behind Feher's patented design is straightforward. Sequential symbols may have the same relative phase, yet the channel-symbol window in which the sinusoids are generated modulates the amplitude at the symbol boundaries. When adja- cent symbols have different phase, this symbol weighting reduces frequency domain sidelobes and hence adjacent-channel interference. Feher suppresses the modulation further with an extended symbol that includes the sequential symbols of the same phase generated with constant amplitude, thus without the weighting-induced amplitude modulation. The result is that energy that normally is redirected into the adjacent channels by the phase discontinu- ities remains within the channel because the discontinuities have been sup- pressed.

The receiver subsystem intersection with the RF conversion segment is shown in Figure 8-1 also. This includes the low noise amplifier (LNA), one or more stages of bandpass filtering (BPF), and the translation of the RF to an IF. In conventional radios, a tunable-reference local oscillator (LO) may be shared between the transmitter and receiver subsystems. FH radios of- ten share a fast-tuning LO between the transmitter and receiver. In military applications, the LO executes a frequency-hopping plan defined by a trans- mission security (TRANSEC) module. In commercial systems (e.g., GSM), a fixed frequency-hopping plan that suppresses fades may be used instead of a complex TRANSEC plan. The radio then either transmits or receives on the frequency to which the LO is tuned. Any radio which employs a physi- cally distinct programmable LO may be a programmable digital radio (PDR), a type of SDR, but it is not a software radio. Software radios use lookup tables to define the instantaneous hop frequencies, not

physical LOs. This ap- proach, of course, requires a wideband DAC. One advantage of using such a DAC is that the hop frequency settles in the time between DAC samples, typically Wa=2:5—hundreds of nanoseconds. The hop frequency is pure and stable instantly, subject to minor distortions introduced by the final power amplifier.

Since the receiver must overcome channel impairments, it may be more complex and technically demanding than the transmitter. Thus, this chapter focuses on receiver design.

Again referring to Figure 8-1, IF processing may be null, as may baseband processing. The direct conversion receiver, for example, modulates a reference signal against the received RF (or IF) signal to yield a baseband binary analog waveform in the in-phase and quadrature (I&Q) channels. Although this kind of RF conversion has nonlinear characteristics, it is particularly effective for single-user applications such as handsets. It may not work well for multiuser applications, however.

This chapter examines the SDR implications of the RF conversion segment. The following section describes receiver architectures. Programmable compo- nent technology including MEMS and EPACs is described. RF subsystem specifications are then analyzed. The chapter concludes with an assessment of RF/IF conversion architecture tradeoffs.

RECEIVER ARCHITECTURES

This section describes the superheterodyne architecture used in base station applications, the direct conversion receiver used in handsets, and related re- search.

The Superheterodyne Receiver

The Watkins-Johnson company [239] publishes the frequency plans of its re-ceivers, an example of which is shown in Figure 8-2. This superheterodyne receiver [240] consists of a preselector and two conversion stages. The prese- lector consists of a matrix of bandpass filters and amplifiers that are switched as defined by the frequency plan for the specific frequency to which the re- ceiver is tuned. The preselector filters

cascade with a low-pass filter and step attenuator that keep the total power of the signal into the first conversion stage within its linear range.

Each conversion stage includes one LO and additional filtering and amplification. The first local oscillator is tuned in relatively coarse steps (e.g., 2.5 MHz in Figure 8-2). The first conversion stage converts the RF to 3733.75 MHz. Higher IF frequencies minimize the physical size of the inductors and capacitors used in the filters. The modulator that converts the RF into the initial IF generates sum and difference frequencies in addition to the desired frequency. The bandpass filter then suppresses these intermodulation products. The low-pass filter further suppresses out-of-band energy. An amplifier and pads with variable gain determine the power into the second conversion stage. The operation of the second stage is similar to the first except that it down- converts the 3733.75 MHz to a standard wideband IF, in this case, 21.4 MHz. In addition, this stage has fine-tuning steps of 1 kHz.

Artifacts must be controlled in the conversion process [241, 242]. In addi- tion to the desired sideband, the conversion process introduces thermal noise, undesired sidebands, and LO leakage into the IF signal as shown in Figure 8-3. Thermal noise is shaped by the cascade of bandpass and low-pass filters. Depending on the RF background environment, thermal noise in the receiver may dominate or thermal-like noise or interference from the environment may dominate the noise power. Superconducting IF filters suppress noiselike interference generated in one cellular half-band from a second, immediately adjacent half-band (e.g., .5 MHz of active signals). See [243] for superconducting filter test results that show a 30 dB suppression of such noise. Undesired sidebands are al- ways present at some very low level because filtering operations suppress sideband energy but do not completely eliminate it. LO leakage occurs be- cause a modulator acts in some ways as a transmission line with imperfect matching. Consequently, part of the power of the LO is transmitted through the modulator to the output.

When the IF is processed digitally, these artifacts can be characterized. Long-term averaging using an FFT, for example, will reveal shape of the noise and the degree of

suppression of the LO leakage and of the undesired sidebands. When designing a PDR, one is concerned that these artifacts not distort the baseband enough to degrade the output SNR or BER unacceptably. When designing an SDR, none of these artifacts should degrade any of the subscriber channels by more than the degradation of the least significant bit (LSB) of the ADC. To accomplish this, the in-band artifacts need to be as uni- form as possible and the maximum level anywhere in the operating band (e.g., in the cell channels) cannot exceed half of the LSB of the ADC. As shown below, this constraint implies that the ADC, postprocessing algorithms, and RF plan must be designed to mutually support each other. Algorithm design- ers who employ floating-point precision at design time may not be familiar with the noise, spurs, and other analog artifacts of the analog RF circuits that limit useful dynamic range constraints. These effects limit the digital dynamic range, and thus reduce the requirements for arithmetic precision in the digi- tal hardware and software. Thus, the effects of each of the disparate analog, digital, and software-signal processing stages have an effect on the sampled signal. When these effects are properly balanced, the wideband superheterodyne receiver yields hundreds of analog subscriber channels that have been struc- tured for the ADC. As a result, the ideal software radio base station replaces hundreds of parallel narrowband analog channels with one wideband chan- nel digitized by a wideband ADC, followed by hundreds of parallel digital channels. Since the digital channels inherently cost less than analog channels, the software radio base station may be more cost-effective than the baseband digital design. Yet most base stations deployed up to 1999 had a baseband digital architecture, not an SDR architecture. The inadequacy of the prior generation of ADC technology explains this situation as discussed in the se- quel. Wideband ADCs were within about 6 to 10 dB of the performance required to effectively compete with baseband architectures in the base sta- tion. By June 2000, digital IF base stations began shipping, but manufacturers did not publically disclose this fact in order to protect this competitive advan- tage. Tsurumi's discussion of zero-IF filtering with up-conversion in a handset architecture provides an innovative approach to multiple-conversion receivers for

handsets [244]. By heterodyning multiple bands to zero-IF, Tsurumi pre- filters any of the commercial standards using a simple programmable low-pass

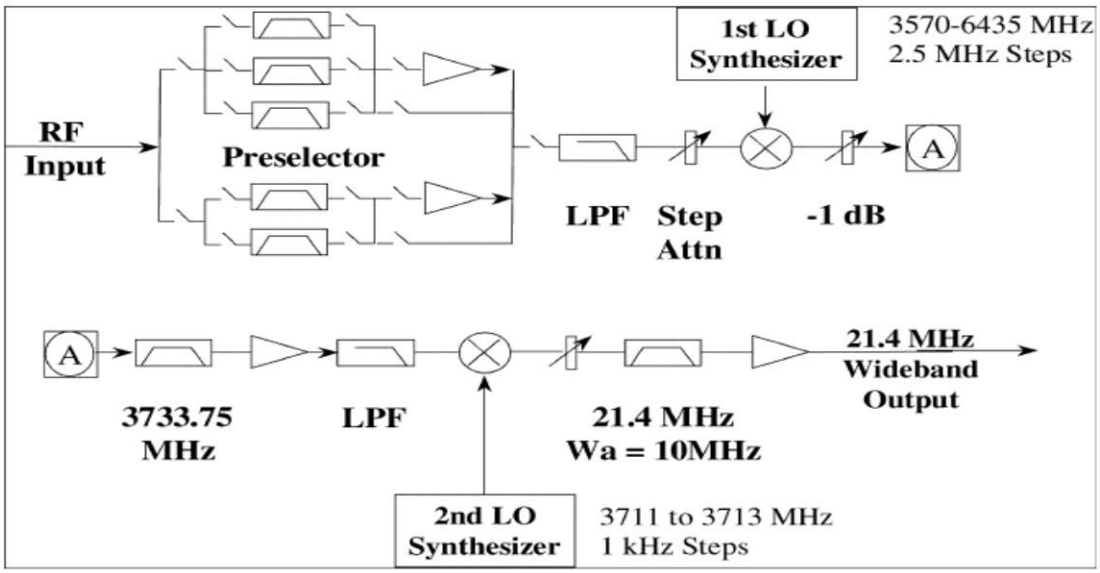


Figure 8-2 Superheterodyne receiver architecture.

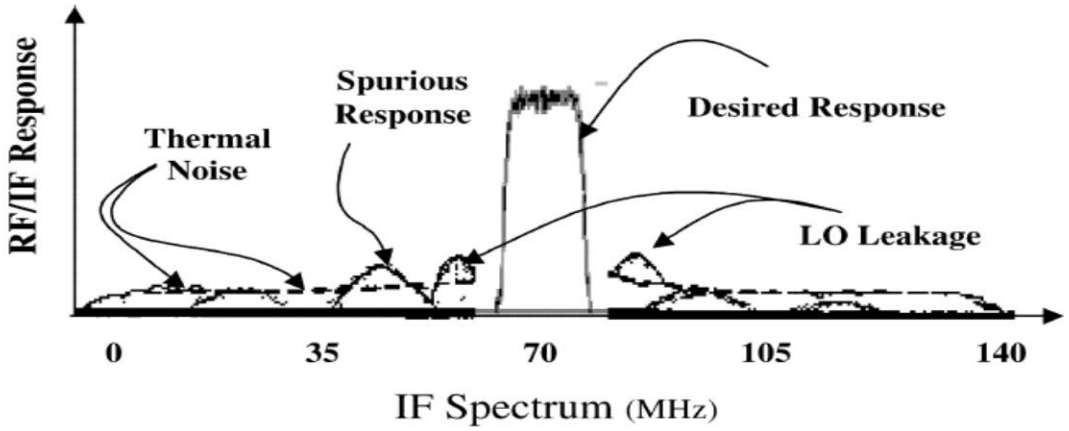
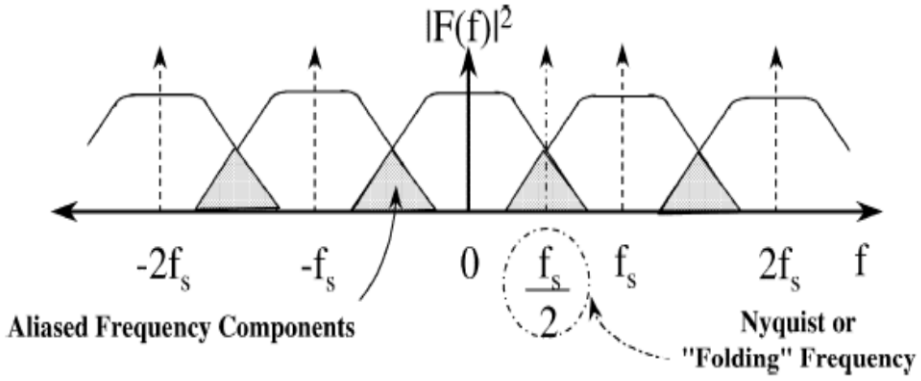


Figure 8-3 Frequency plan suppresses spectral artifacts.

5) Discuss in detail about the potential benefits and technology in ADC/DAC tradeoffs. (L-2, CO-1)

REVIEW OF ADC FUNDAMENTALS

Since the wideband ADC is one of the fundamental components of the soft- ware radio, this chapter begins with a review of relevant results from sam- pling theory. The analog signal to be converted must be compatible with the capabilities of the ADC or DAC. In particular, the bandwidths and linear dy- namic range of the two must be compatible. Figure 9-1 shows a mismatch between an analog signal and the ADC. For uniform sampling rate f_S, the maximum frequency for which the analog signal can be unambiguously re- constructed is the Nyquist rate, f_S=2. The wideband analog signal extends beyond the Nyquist frequency in the figure. Because of the periodicity of the sampled spectrum, those components that extend beyond the Nyquist fre- quency fold back into the sampled spectrum as shown in the shaded parts of the figure (thus the term *folding frequency*). This is well known as alias- ing [274, 275]. Although some aliasing is unavoidable, an ADC designed for software-radios must keep the total power in the aliased components below the minimum level that will not unacceptably distort the weakest subscriber signal.



Clipping Distortion

In most applications, one cannot control the energy level of the maximum signal to be exactly equal to the most significant bit. One must therefore allow for some AGC

or for some peak power mismatch. Clipping of the peak energy level introduces frequency domain sidelobes of the high power signal. These sidelobes have the general structure of the convolution of the signal's sinusoidal components with the Fourier transform of a square wave, which has the form of a sin(x)=x function. Frequency domain sidelobes have a power level of □□ 11 dB, other signals in a wideband passband. In practice, avoiding clipping may occupy the entire most significant bit (MSB). Usable dynamic range may therefore be one or two bits less than the ADCs resolution.

Aperture Jitter

Sample-and-hold circuits also limit ADC performance as illustrated in Fig- ure 9-5. Consider a sinusoidal input signal, V(t) = A cos(!t), where ! is the maximum frequency. The rate of change of voltage is as shown, yielding a maximum rate of change of $2A=(2^B)$ or $A=(2^{(B+1)})$. The time duration of this differential interval is inversely proportional to the frequency and the exponential of the number of bits in the ADC. This period is the aperture uncertainty, the shortest time taken for a maximal-frequency sine wave to traverse the LSB. The timing jitter must be a small fraction of the aperture uncertainty to keep the total error to less than $\frac{1}{2}$ LSB. Therefore, the timing jitter should be 10% or less of the uncertainty shown in the figure. An 8-bit ADC sampling at 50 MHz requires aperture jitter that is less than a picosecond (ps).

This stability must be maintained for a period of time that is inversely pro- portional to the frequency stability that one requires. If, for example, the min- imum resolvable frequency component for the signal processing algorithms should be 1 kHz, then the timing accuracy over a 1 ms interval should be less than the aperture uncertainty. Short-term jitter can be controlled to less than 1 ps for 1 ms with current technology. If the spectral components should be accurate to 1 Hz, then the stability must be maintained for 1 second. Due to drift of timing circuits, such performance may be maintained for 10^9 to 10^{11} aperture periods, or on the order of 1 to 100 ms. Stability beyond these rela- tively short intervals is problematic due to drift induced by thermal changes, among other things. A sampling rate of 1 GHz with 12 bits of resolution requires

about 2 fs of aperture jitter or less. This stability is beyond the cur- rent state of the art, which corresponds to 6.5 to 8 bits of resolution at these sampling rates.

Quantization and Dynamic Range

Quantization step size is related to power according to [279]:

$P_q = q^2=12R$

where q is the quantization step size, and R is the input resistance. The SNR

at the output of the ADC is

$SNR = 6:02\ B + 1:76 + 10\ \log(f_S=2f_{max})$

where B is the number of bits in the ADC, f_S is the sampling frequency, and f_{max} is the maximum frequency component of the signal.

For Nyquist sampling, $f_S = 2f_{max}$, so the ratio of these quantities is unity.

Since the log of unity is zero, the third term of the equation for SNR above is

eliminated. The approximation for Nyquist sampling, then, is that the dynamic range with respect to noise equals 6 times the number of bits. This equation suggests that the SNR may be increased by increasing the sampling rate be- yond the Nyquist rate. This is the principle behind the *sigma-delta/delta-sigma* ADC.

Technology Limits

The relationship between ADC performance and technology parameters has been studied in depth by Walden [280, 281]. His analysis addresses the electronic parameters, aperture jitter, thermal effects, and conversion-ambiguity. These are related to specific devices in Figure 9-6. The physical limits of ADCs are bounded by Heisenberg's uncertainty principle. This core physical limit suggests that one could implement a 1 GHz ADC with 20 bits (120 dB) of dynamic range. To accomplish this, one must overcome thermal, aperture jitter, and conversion ambiguity limits. Thermal limits may yield to research in Josephson Junction or high-temperature superconductivity (HTSC) research. For example, Hypress has demonstrated a 500 Msa/sec (200 MHz) ADC with dynamic range of 80 dB operating at 4K [435]. Walden notes that advances in ADC technology have been limited. During the last eight years, SNR has improved only 1.5 bits. Substantial investments are required for continued progress. DARPA's Ultracomm

program, for example, funded research to realize a 16-bit 100 M
[282]. Commercial research continues as well, with Analog Devices' announcement of the

ADC AND DAC TRADEOFFS

The previous section characterized the Nyquist ADC. This section provides an overview of important alternatives to the Nyquist ADC, emphasizing the tradeoffs for SDRs. It also includes a brief introduction to the use of DACs.

Sigma-Delta (Delta-Sigma) ADCs

The sigma-delta ADC is also referred to in the literature as the delta-sigma
ADC. The principle is understood by considering an analogous situation in visual signal (e.g., image) processing. The spatial frequency of a signal is in- versely proportional to its spatial dimension. A large object in a picture has low spatial frequency while a small object has high spatial frequency. Spa- tial dynamic range is the number of levels of grayscale. A black-and-white image has one bit of dynamic range, 6 dB. But consider a picture in a typi- cal newspaper. From reading distance, the eye perceives levels of grayscale, from which shapes of objects, faces, etc. are evident. But under a magnifying glass, typical black-and-white newsprint has no grayscale. Instead, the picture is composed of black dots on a white background. These dots are one-bit digitized versions of the original picture. The choice between white and black is also called zero-crossing. The dots are placed so close together that they oversample the image. The eye integrates across this 1-bit oversampled im- age. It thus perceives the low-frequency objects with much higher dynamic range than 6 dB. The gain in dynamic range is the log of the number of zero- crossings over which the eye integrates. Zakhor and Oppenheim [283] explore this phenomenon in detail, with applications to signal and image processing. Thao and Vetterli [284] derive the projection filter to optimally extract max- imum dynamic range from oversampled signals. Candy and Temes offer a definitive text [285].

1. Principles The fundamentals of an oversampling ADC for SDR appli- cations are illustrated in Figure 9-7. A low-resolution ADC such as a zero- crossing detector oversamples the signal, which is then integrated linearly. The integrated result has greater dynamic range and smaller bandwidth than the oversampled signal. The amount of oversampling is the ratio of the sampling frequency of the analog input to the Nyquist frequency, shown as k in the figure.

Since $f_{Nyquist} = 2f_{max}$, the oversampling rate must be at least $2kf_{max}$. With continuous 1 : k integration of the zero-crossing values, the output register contains a Nyquist approximation of the input signal. Since the integrated output has an information bandwidth that is not more than the Nyquist bandwidth, the integrated values may be decimated without loss of information. Decimation is the process of selecting only a subset of available digital samples. Uniform decimation is the selection of only one sam- ple from the output register for every k samples of the undecimated stream. If the signal bandwidth is 0.5 MHz, its Nyquist sampling rate is 1 MHz. A zero- crossing detector with a sampling frequency of 100 MHz has an oversampling gain of ten times the log of the oversampling ratio (100 MHz/1 MHz), 20 dB. The single-bit digitized values may be integrated in a counter that counts up to at least 100. Although this is the absolute minimum requirement, real signals may exhibit DC bias. A counter with only a capacity of 100 could tolerate no DC bias. A counter with range that is a power of two, e.g., 128, tolerates up to log(28) bits or 4.7 of DC bias. For a range of 128, a signed binary counter requires $\log_2(128)$ bits or 7 bits plus a sign bit. The counter treats each zero- crossing as a sign bit, +1 or [illegible] of this 8-bit counter, with [illegible] counter, [illegible] every 10 an output-sampling rate to 1 MHz as required for

Nyquist sampling. Zero-crossing detectors do not work properly, however, if there are insuffi- cient crossings to represent the signal. For example, if DC bias drifts beyond the full-scale range of the detector, then there will be no zero-crossings and no signal. A signal may be up-converted, amplified, and clipped to force the re- quired zero-crossings. A similar effect can be realized in linear oversampling ADCs through the addition of dither. A dither signal is a pseudorandomly generated train of positive and

negative analog step-functions. The dither is added to the input of the ADC before conversion (but after anti-alias filter- ing). The corresponding binary stream is subtracted from the oversampled stream. Alternatively, an integrated digitized replica of the dither signal may be subtracted from the integrated output stream. This forces zero-crossings, enhancing the SNR. One may view dithering as a way of forcing spurs gen- erated by sample-and-hold nonlinearities to average across multiple spectral components, enhancing SNR.

In addition, high power out-of-band components will be sampled directly by the zero-crossing detector. These components will then be integrated, sub- ject to the bandwidth limitations imposed by the integrator-decimator. The anti-aliasing filter therefore must control total oversampled power so that it conforms to the criteria for Nyquist ADCs.

2. Tradeoffs There are several advantages to oversampling ADCs. First, sam- ple-and-hold requirements are minimized. There is no sample-and-holdircuit in a zero-crossing detector. Simple threshold logic, possibly in con- junction with a clamping amplifier, yields the single-bit ADC.

Aperture jitter remains an issue, but the jitter is a function of the number of bits, which is 1 at the oversampling rate. This minimizes aperture jitter requirements for a given sample rate. As the oversampled values are integrated, the jitter averages out. In order to support large dynamic range for narrowband signals, the timing drift (the integration of aperture jitter) should contribute negligibly to the frequency components of the narrowband signal. This means that integrated jitter should be less than 10% of the inverse of the narrowband signal's bandwidth, for the corresponding integration time.

In addition, the anti-aliasing filter requirements of a sigma-delta ADC are not as severe as for a Nyquist ADC. The transfer-function of the anti-aliasing filter is convolved with the picket-fence transfer-function of the decimator. Thus, the anti-aliasing filter's shape factor may be 1=k that of a linear ADC for equivalent performance. Many commercial products use oversampling and decimation within an ADC chip to achieve the best combination of bandwidth and dynamic range.

Oversampled ADCs work well if the power of the out-of-band spectral components is low. In cell site applications, Q must be very high in the filter that rejects adjacent band interference. Superconducting filters [286] may be appropriate for such applications.

6.) Discuss in detail about the potential benefits and technology in ADC/DAC tradeoffs. (L-2, CO-1)

This chapter addresses digital hardware architectures for SDRs. A digital hard- ware design is a configuration of digital building blocks. These include ASICs, FPGAs, ADCs, DACs, digital interconnect, digital filters, DSPs, memory, bulk storage, I/O channels, and/or general-purpose processors. A digital hardware architecture may be characterized via a reference platform, the minimum set of characteristics necessary to define a consistent family of designs of SDR hardware. This chapter develops the core technical aspects of digital hardware architecture by considering the digital building blocks. These insights permit one to characterize the architecture tradeoffs. From those tradeoffs, one may derive a digital reference platform capable of embracing the necessary range of digital hardware designs. The chapter begins with an overview of digital processing metrics and then describes each of the digital building blocks from the perspective of its SDR architecture implications.

METRICS

Processors deliver processing capacity to the radio software. The measure- ment of processing capacity is problematic. Candidate metrics for processing capacity are shown in Table 10-1. Each metric has strengths and limitations. One goal of architecture analysis is to define the relationship between these metrics and achievable performance of the SDR. The point of view employed is that one must predict the performance of an unimplemented software suite on an unimplemented hardware platform. One must then manage the compu- tational demands of the software against the benchmarked capacities of the hardware as the product is implemented. Finally, one must determine whether an existing software personality is compatible with an existing hardware suite.

1. Differentiating the Metrics MIPS, MOPS, and MFLOPS are differentiated by logical scope. An operation (OP) is a logical transformation of the data in a designated element of hardware in one clock cycle. Processor architectures typically include hardware elements such as arithmetic and logic units (ALUs), multipliers, address generators, data caches, instruction caches, all operating in parallel at a synchronous clock rate. MOPS are obtained by multiplying the number of parallel hardware elements times the clock speed. If multiple operations are required to complete a machine instruction (e.g., a floating-

point multiply), then

MIPS = ®MOPS, ®<1

If, on the other hand, the processor has a *very long instruction word* (VLIW), ® may be greater than 1. Suppose, for example, that a processor includes a "smart" cache, an ALU, and two parallel multiplier units with a 250 MHz system clock. One could characterize this processor in terms of the operations of the ALU and multipliers. If ® = 1, then it can deliver 250

point multiply on every clock cycle, then the processor provides 500 MFLOPS. Thus, one may characterize such a device as capable of a peak of 750 MIPS/500 MFLOPS. This notation means "750 MIPS of which up to 500 may be MFLOPS." Digital filtering takes more floating-point operations than, say, protocol processing, or FEC algorithms. If the SDR application uses a mix of 50% ALU and 50% floating point operations, then the processor delivers a maximum of

0:5

Clearly, processing capacity realized is a function of instruction mix.

Alternatively, one could consider just the memory cache operations, at- tributing 250 MOPS of memory operations (MEOPS). If the memory cache operates fast enough so that the ALU and multipliers are never waiting for data or instructions, then the memory cache is not a bottleneck. If, however, there are states in which it must wait, then the potential 750 MIPS will not be realized. In this case, since MEOPS < MIPS, then the peak of 750 MIPS cannot be sustained beyond the capacity of the cache. For extremely

com- putationally intensive operations like digital filtering, one may in fact realize the maximum capacity because all the data is resident in cache. Cache-misses then degrade performance.

2. Processor-Memory Interplay The execution of an instruction requires ac- cessing memory for instructions and data or accessing local registers. Pro- cessors that are more complex may fill a pipeline with instructions to be executed concurrently. Pipelines produce no results until the pipeline is full. Thereafter, pipelines produce a result per clock cycle. Newer architectures may employ set-associative cache coherency and other schemes to yield a higher number of instruction executions for a given clock speed. In addi- tion, there is statistical structure to the application, which will determine whether the data and instruction necessary at the next step will be in the cache (cache hit) or not (cache miss). Statistical structure is also present in the mix of input/output, data movement in memory, logical (e.g., masking and finding patterns), and arithmetic needed by an application. Some appli- cations like FFTs are very computationally intensive, requiring a high pro- portion of arithmetic instructions. Others such as supporting display windows require more copying of data from one part of memory to another. And sup- port of virtual memory requires the copying of pages of physical memory to hard disk or other large-capacity primary storage. This gives the programmer the illusion that physical memory is relatively unlimited (e.g., 32 gigabytes) within a physically confined space of, say, 128 Mbytes of physical mem- ory.

3. Standard Benchmarks Consequently, MIPS are hard to define. Often, the popular literature attributes MIPS based on a nonstatistical transformation of MOPS into instructions that *could be executed in an ideal instruction mix*. This approach makes the chip look as fast as it possibly could be. Since most manufacturers do this, the SDR engineer learns that achievable per- formance on the given application will be significantly less than the nomi- nal MIPS rating. The manufacturer's MIPS estimate is useful because it de- fines an upper bound to realizable performance. Most chips deliver 30 to 60% of such nominal MIPS as usable processing capacity in a realistic SDR mix. In

the 1970s, scientists and engineers concerned with quantifying the ef- fectiveness of supercomputers developed the Whetstone, Dhrystone, and other benchmarks consisting of standard problem sets against which each new gen- eration of supercomputer could be assessed. These benchmarks focused on the central processor unit (CPU) and on the match between the CPU and the memory architecture in keeping data available for the CPU. But they did not address many of the aspects of computing that became important to prospective buyers of workstations and PCs. The speed with which the dis- play is updated is a key parameter of graphics applications, for example. The SPECmarks evolved during the 1990s to better address the concerns of the early-adopter buying public. Consequently, SPECmarks are informative but these also are not the ideal SDR metric in that they do not generally reflect the mix of instructions employed by SDR applications. Turletti [293], how- ever, has benchmarked a complete GSM base station using SPECmarks, as discussed further below.

4. SDR Benchmarks At this point, the reader may be expecting some new "SDR benchmark" to be presented as the ultimate weapon in choosing among new DSP chips. Unfortunately, one cannot define such a benchmark. First of all, the radio performance depends on the interaction among the ASICs, DSP, digital interconnect, memory, mass storage, and the data-use structure of the radio application. These interactions are more fully addressed in Chapter 13 on performance management. It is indeed possible to reliably estimate the performance that will be achieved on the never-before-implemented SDR application. But the way to do this is not to blindly rely on a benchmark. Instead, one must analyze the hardware and software architecture (using the tools described later). One may then accurately capture the functional and statistical structure of the interactions among hardware and software. This systems analysis proceeds in the following steps:

Identify the processing resources.

- Characterize the processing capacity of each class of digital hardware. Characterize the processing demands of the software objects

- Determine how the capacity of the hardware supports the processing demands of the software by mapping the software objects onto the sig- nificant hardware partitions. here is a trap in identifying the hardware processor classes. ASICs and DSPs are easily identified as processing modules. But one must traverse each signal processing path through the system to identify buses, shared memory, disks, general-purpose CPUs, and any other component that is on the path from source to destination (outside the system). Each such path is a process- ing thread. Each such processor has its own processing demand and priority structure against which the needs of the thread will be met. One then abstracts the block diagram into a set of critical resources, as illustrated in Figure 10-1. This chapter begins the process of characterizing the capacity of SDR hard- ware. It summarizes the tradeoffs among classes of processor, functional ar- chitecture, and special instruction sets. Other source material describes how to program them for typical DSP applications [294]. The extensive literature available on the web pursues detailed aspects of processors further [295-298]. The popular press provides product highlights (e.g., [299-303]). This text, on the other hand, focuses on characterizing the processors with respect to the support of SDR applications. This is accomplished by the derivation of a dig- ital processing platform model that complements the RF platform developed previously.

Unit-II

Sdr Architecture

Part-A

1)Define DDI? (L-1,CO-2)

The Defense Information Infrastructure (DII) of the United States consists of the fixed plant of telecommunications and information processing systems plus the mobile infrastructure that military forces must take with them on deployments around the world.

2)What are the architecture goals of SDR? (L-2,CO-2)

In long term,software defined radios are expected by proponents like SDRForum to become the dominant technology in radio communications.

3)what are the military requirements of radio services? (L-2,CO-2)

Mobility of both subscribers and infrastructure

INFOSEC (TRANSEC and COMSEC)

Ruggedness and reliability in austere operating environments

Growth from voice and low-speed data to high-speed *tactical internets*

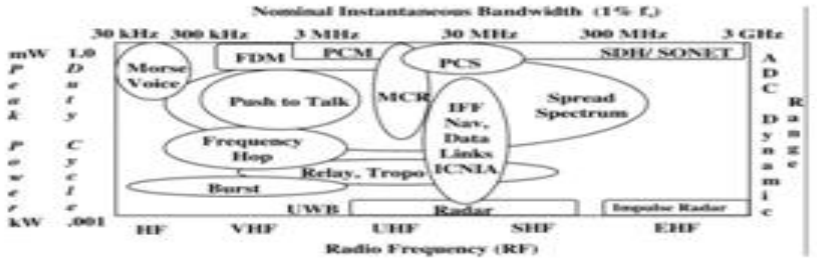
Interoperability with legacy radios and coalition partners

Affordability

4) what is architecture revolution road map? (L-2,CO-2)

Software-radio architecture has been evolving from its roots in military com-munications. In addition to the traditional emphasis on radar and radar jamming, some developed wideband digital techniques for radio.

5) Draw communication clusters in RF signal space? (L-2,CO-2)



6) Define software flexibility and affordability. (L-2,CO-2)

No air-interface-defining software (0), Single-supplier software (1), Multiple-supplier but single-host platform (2), Multiple-supplier multiplatform software

7) Mention the applications of SDR(L-1,CO-2)

JTRS-used in military Amateur & home use USRP-uses a USB 2.0 interface

8)Define low band noise? (L-2,CO-2)

the lower radio bands—HF, VHF, and lower UHF—include significant sources of radio noise and interference. The incidental and unavoidable interference includes automobile ignitions, microwave ovens, power distribution systems, gaps in electric motors, and the like.

9) Define noise interference? (L-2,CO-2)

The noise/interference levels are defined with respect to thermal noise:

$$P_n = kTB$$

where k is Boltzmann's constant, T is the system temperature (T_0 is the refer-ence temperature of 273 Kelvin), and B is the bandwidth (e.g., per Hz).

10) How VLHF multi channel air interference? (L-2,CO-2)

FM frequency division multiplexing (FM/FDM) for military LVHF applications includes modes with four channels per RF carrier.

These meet the con- nectivity needs of radiotelephony operations of relatively low-echelon military forces. Due to the relatively narrow coherence bandwidths of LVHF, conven- tional FM/FDM is limited to about 60 channels.

11) What is satellite communication mode? (L-2,CO-2)

Communications satellites operate in the three orbital regimes .Geosynchronous satellites have an orbital period which is nearly identical to the earth's rotational period, resulting in an apparent stationary position above the equator at an altitude of 22,500 miles, approx- imately.

12) Define disaster relief case study? (L-2,CO-2)

This case study considers a mobile communications capability for disaster re- lief. The capability includes mobile infrastructure, mobile nodes, and handsets. The design emphasis is on defining an open architecture for the infrastructure.

13) Expand and explain JTRS? (L-2,CO-2)

The Joint Tactical Radio System (JTRS) Programmable Modular Communications System (PMCS) integrated process team (IPT) recommended the consolidation of the more than 200 nomencla- tured U.S. radio families into a single program, JTRS, under the joint manage- ment of the three U.S. military services. The Joint Tactical Radio (JTR) mis- sion needs statement (MNS) and Operational Requirements Document (ORD) express the vision for the functional capability of the JTRS.

14) Define Transparent bridging? (L-3,CO-2)

Software radios not only support standard services, but they also can provide background routing services, which the military calls *transparent bridging*.

15) Define software objects? (L-2,CO-2)

One may apply the principles of object-oriented design to the design of an SDR node in a top-down way, as outlined in this section.

16) Define network layer? (L-2,CO-2)

The network layer contributes additional constraints on SDR design. First,

the well-established network analysis tools may be employed to analyzing routing, queuing, and related aspects of the network. This level of analysis establishes resource bounds, such as maximum buffer sizes and processing latency in a node.

17) Define homeomorphism? (L-2,CO-2)

one may compare the range of one primitive to the domain of another using a topological map called a *homeomorphism*, a topology- preserving mapping.

18) Define Programmable Digital Radio (PDR)? (L-2,CO-2)

To develop the relationship between architecture and implementations, atten-tion turns to a series of case studies of the progenitors and research implemen- tations of the software radio.

19) What is industry standard zone archietecture? (L-2,CO-2)

The approach to defining open-architecture wireless taken by the SDR Forum is considered first. These are both open- architecture standards. There are many PDRs and touted SDRs in existence, but there is as yet no single manufacturer that so dominates the industry that one could say a de facto standard exists in the year 2000.

PART B

1)Define and explain each essential functions of the software radio. (L-2,CO-2)

Technology advances have ushered in new radio capabilities that require an expansion of the essential communications functions of source coding and channel coding. The new aspects are captured in the software radio functional model.

A. The Software Radio Functional Model

Multiband technology [1], first of all, accesses more than one RF band of

Communications Services	*Applications and related services* (e.g., over-the-air downloads)
Radio Applications	***Air interfaces ("waveforms")*** State machines, modulators, interleaving, multiplexing, FEC, control and information flows
Radio Infrastructure	Data movement: drivers, interrupt service routines, memory management, shared resources, semaphores
Hardware Platform	Antenna(s), analog RF hardware, ASICS, FPGAs, DSPs, microprocessors, instruction set architecture, operating systems

communications channel at once. The RF channel then is generalized to the channel set of Figure 1-2. This set includes RF channels, but radio nodes like PCS base stations and portable military radios also interconnected to fiber and cable; therefore these are also included in the channel set. This segment also provides multiple signal paths and RF conversion that span multiple RF bands. IF processing may include filtering, further frequency translation, space/time diversity processing, beam- forming, and related functions. Multimode radios [3] generate multiple air interface waveforms (modes) defined principally in the modem, the RF chan- nel modulator-demodulator

B. Functional Interfaces

After identifying the functions to be accomplished in a software radio, one must define the interface points among the functional components. Figure 1-3 identifies these interfaces. The notation "RF waveform" is shorthand for air interface. The IF waveform includes most aspects of the air interface, but the signals have been filtered and converted to an IF that facilitates processing.

C. Architecture

Since industrywide agreement on anything can be challenging, one should begin with a definition of architecture. The *Random House Unabridged Dictionary* defines *architecture* as "a fundamental underlying design of computer hard- ware, software, or both [25]". While this is an agreeable definition, it provides no prescription of what "underlying design" entails.

1. ***Functions, Components, and Design Rules*** None of the many possible definitions of architecture suit the purposes of defining architecture for the software radio. One that best relates services, systems, technology, and economics is best suited to the software radio.

2. ***2. Plug-and-Play*** If an architecture supports plug-and-play, then the design rules have been crafted so that hardware and software modules from different suppliers will work together when plugged into an existing system. Hardware modules will plug-and-play if the physical interfaces and logical structure of the functions supplied by that module are compatible with the physical interfaces, allocation of functions, and other design rules of the host hardware platform. Software modules will plug-and-play if there is a comprehensive but simple interface to the host environment, and if the module offers to the environment the information that it needs in order to employ it as a resource.

D. Levels of Abstraction

Clearly, software radio functions do not all share the same logical level of abstraction. A modem, for example, supports data movement from baseband to IF, data transformation from bits to channel symbols, timing recovery, FEC and the related functions. It is therefore not accurate to think of software radio architecture as merely a collection of functions with associated interfaces.

One must then define interfaces among these levels. One approach is the definition of an applications programming interface (API) from one horizontal layer to the next. The API calls may be thought of as the vertical interfaces among horizontal layers. This approach has been used with reported success on technology pathfinders [30], and will be dealt with in some detail in this text. Not all APIs that have been described conform to the four layers identified above. These four layers, however, are conceptual anchors that help organize the process of evolving the software radio architecture.

2)What are the architectural goals of SDR? Explain with neat diagrams. (L-2,CO-2)

OPEN ARCHITECTURE AND STANDARDS EVOLUTION

Industry organizations such as the SDR Forum are in the process of develop- ing open architecture for SDR [90]. In addition, work of the Object Manage- ment Group (OMG), the Telecommunications Industries Association (TIA), the Internet Engineering Task Force (IETF), the Wireless Applications Proto- col (WAP) Forum, and the IEEE includes standards relevant to software radio architecture.

A. The Software-Defined Radio (SDR) Forum

In March 1996, the U.S. government invited industry to participate in what it named the Modular Multifunction Information Transfer Systems (MMITS) forum. It hoped this group would become an industry body to establish open- architecture standards for SPEAKeasy. The initial DoD thinking was that MMITS might be a study group of the VME International Trade Associa- tion (VITA) because of the success of VME as an open-architecture standard for the technology pathfinder SPEAKeasy I program. The author was elected to chair the nascent organization, which decided not to align with VITA or any other standards organization per se. Instead, it attempted to function like the ATM Forum, a quick-response consortium that would publish recommenda- tions based on current engineering practice. It planned to delegate the formal standards-setting process to others (OMG, TIA, IEEE, ANSI, ITU, ETSI, etc.). That choice proved to be a wise one. For example, SPEAKeasy II chose the PCI bus, ITT chose PC-104, GEC chose a narrowband control bus, and, in general, there was no consensus on VME or any other backplane as a paradigm for industry cooperation.

B. Product Standards Organizations

The plethora of hardware standards potentially relevant to SDR can be out- lined but not exhaustively enumerated. In part, this is so because product standards have been emerging at breakneck speed. In 1995 and early 1996, for example, the VME standard backplane/bus had a majority of rack-mount open-architecture designs. By late 1996, the PCI bus had taken the lead with a large variety of DSP cards, ADCs, host processors, and other modules nec- essary for open-architecture software radio. By 1999 compact PCI (cPCI) had increasing popularity, while PC-104 retained a strong market niche. By April, 2000, systems-on-chip using DSP cores had a strong following. The families of product standards that have a continuing relevance to SDR are shown in Table 2-3.

Analog hardware standards may be useful for defining interfaces with an- tennas. Interconnect and backplanes will probably not be standardized per se, but emerging

open-architecture middleware will hide the details of this inter- connect technology. Internetworking standards, similarly, have to be accom- modated in any viable software radio architecture. Object-oriented standards support the design process (e.g., the Unified Modeling Language, UML [92]). CORBA provides the middleware essential for open-architecture in software radio.

C. Air Interface Standards

Finally, air interface standards organizations define channel modulations, frequency llocations, access protocols, and other characteristics of the radio interface needed for interoperability over the air. The ITU has organs that ad- dress the radio aspect (ITU-R) and the telecommunications aspect (ITU-T). The European Telecommunications Standards Institute (ETSI) sets radio stan- dards in Europe, while the Telecommunications Industries Association (TIA), the Electronics Industries Association (EIA), and the Institute of Electrical and Electronics Engineers (IEEE) set regional standards in the United States. ARIB sets standards in Asia. And there are numerous other regional, national, and local standards organizations.

The variability of standards creates the need for flexibility of band and mode for a "world phone." A handset that accommodates first-generation, second- generation GSM, and third-generation (3G) CDMA would be such a world phone. Industrial goals for 3G software radio handsets contemplate a mix of ASICs, DSP cores, and general-purpose microcontrollers in 3G handsets. Infrastructure costs are also such a driver that continuing proliferation requires future-proof infrastructure to ensure affordability of future mobile wireless.

D. The Global Deliberative Process

The United States, Europe, and Asia each have central perspectives on soft- ware-radio architecture. These perspectives are reflected in the actions of those who

most aggressively push the technology. Participation in the SDR Forum provides a useful gage of interests. U.S. interests currently center on the needs of the cellular service providers and of the military.

the software radio concept offered the GSM proponents a low-cost means of migrating toward 3G. Given a 20 MHz W-CDMA de- spreader ASIC, one could easily digitize the 200 kHz GSM subscriber signals using minimum chip area on the despreader ASIC. The DSP power necessary for 3G could then be employed to filter the subscriber signals in software, yielding a GSM SDR mode. In addition, W-CDMA generation would employ high-speed circuits that could be adapted to generating the GSM waveform. The GSM MoU committee therefore recommended that software radio technology be employed to provide graceful migration to 3G.

3)Explain the RF front-end architecture of SDR.

In a radio receiver circuit, the **RF front end** is a generic term for all the circuitry between the antenna and including the mixer stage.[1] It consists of all components in the receiver that process the incoming signal at the original radio frequency (RF), before it is converted to a lower intermediate frequency (IF). In microwave and satellite receivers it is often called the *low-noise block* (LNB) *or low-noise downconverter* (LND) and is often located at the antenna, so that the signal from the antenna can be transferred to the rest of the receiver at the more easily handled intermediate frequency.

For most superheterodyne architectures, the RF front end consists of:[2]

A 'gentle' band-pass filter (BPF) to reduce strong out-of-band signals and image frequency response;

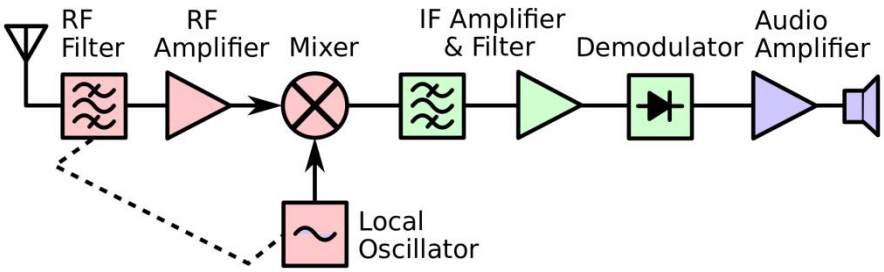
An RF amplifier, often called the low-noise amplifier (LNA). Its primary responsibility is to increase the sensitivity of the receiver by amplifying weak signals without contaminating them with noise, so that they can stay above the noise level in succeeding stages. It must have a very low noise figure (NF). The RF amplifier may not be needed and is often omitted (or switched off) for frequencies below 30 MHz, where the signal-to-noise ratio is defined by atmospheric and man-made noise.

A local oscillator (LO) which generates a radio frequency signal at an offset from the incoming signal, which is mixed with the incoming signal.

The mixer, which mixes the incoming signal with the signal from the local oscillator to convert the signal to the intermediate frequency (IF).

In many modern integrated receivers, particularly those in wireless devices such as cell phones and Wifi receivers, the intermediate frequency is digitized; sampled and converted to

a binary digital form, and the rest of the processing - IF filtering and demodulation - is done by digital filters (digital signal processing, DSP), as these are smaller,



use less power and can have more selectivity.[3] In this type of receiver the RF front end is defined as everything from the antenna to the analog to digital converter (ADC) which digitizes the signal.[3] The general trend is to do as much of the signal processing in digital form as possible, and some receivers digitize the RF signal directly, without down-conversion to an IF, so here the front end is merely an RF filter.

4)Discuss about the use of MEMS in RF for SDR. (L-2,CO-2)

RF MEMS switches, vacastors and inductors : DC-120 GHz Micromachined hyperfrequency components: transmission lines, high-Q resonators, filter, antenna (12-200 GHz). No mobile parts, no operation in mechanical domain. Not a truly MEMS devices, but using technologies similar with MEMS devices. FBAR (thin Film Bulk Acoustic Resonators), filters : integrable very high Q filters/resonators for

Capacitive switches

The most efficient and promising, since no mechanical contact: a large lifetime Drawbacks : efficient only at high frequencies, limited insulation

Capacitive switches

•Up-capacitance : the dielectric layer can be neglected

•Tens of fF

•Holes in the upper membrane : needed for the releasing of the mobile part

•Holes : 4-6 mm diameter, spaced by 5-6 mm period

. •Typical gap : 3-4 mm

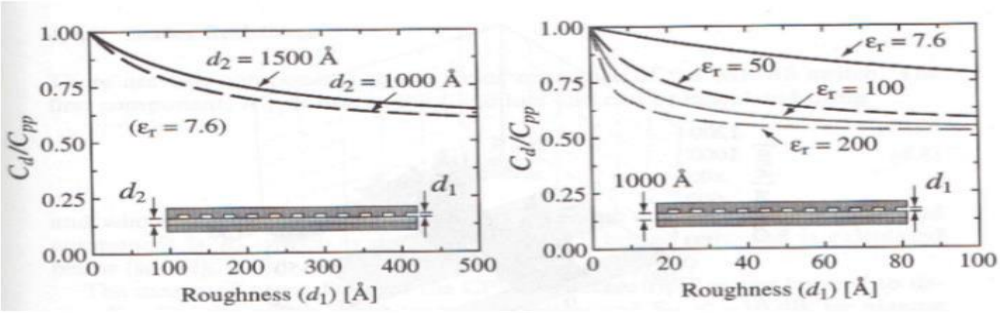
•The holes dont affect the up-state capacitance: fringe field

Capacitive switches

• Down capacitance : defined by the dielectric (thickness, dielectric constant).

• Should be as high as possible, however, limited by the Minimal thickness of dielectric

(~1000-1500 Å) which should support the actuation voltage (20-50 V) Roughness of the surface : a degradation of the down-capactiance



DC-contact series switches Relevant parameters: Up-state capacitance Contact series resistance Inductance Gold-to-gold contact : 0.1 Ohms for applied force of 100-500 µN, contact area of 20 µm2,

MEMS switches and pull-in phenomenon

• Two parameters: pull-in voltage and hold-down voltage

• Hysteretis characteristic x(V)

• Exercice: calculate the pull-in voltage and the pulldown voltage if W=100µm, w=80 µm, k=30 N/m, td=0.1 µm, eps=7. k is supposed to be constant

• Calculate the contact force in down position

5)With neat diagrams, explain the functional components and properties of SDR

architecture. (L-2,CO-2)

Software Architecture and Components:-

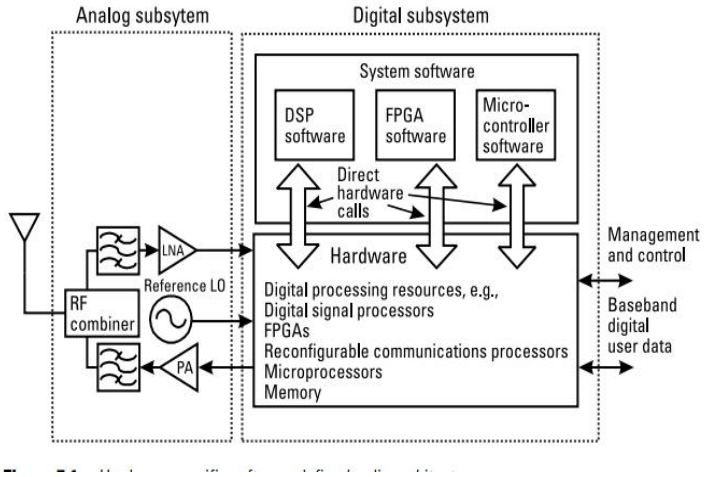
In previous chapters we have focused on system-level design, hardware selection, and functional partitioning. These topics currently dominate the early design stages of a software defined radio project. This is particularly true for 3G cellular mobile radio, because state-of-the-art hardware is required to meet demanding performance specifications. Ultimately it is the software that provides the functionality, and the software architecture must include characteristics and mechanisms that allow for an efficient utilization of the underlying hardware platform.

Major Software Architectural Choices

Hardware-Specific Software Architecture In Chapter 1 we introduced the concept of the ideal software defined radio (see Figure 1.2) and proposed a layered abstracted software architecture. This concept allows the application software to be independent of an underlying standardized hardware platform. The aim of this approach is that ultimately any

investment in the development of application software is maintained when ported to new (and presumably better) hardware platforms that comply with the standard.

Commercially designed cellular mobile radio equipment (terminals and base stations) has traditionally been developed as a black box. A software-level interface is not provided; only high-level functional and physical interfaces are exposed (e.g., mobile phone serial interface or BTS Abis interface [1]). In most cases radio equipment (1G and 2G) from different vendors will have incompatible software architectures, where the driving requirements have been to support legacy hardware or software or both. The degree to which these traditional developments have utilized common interfaces and object-oriented (OO) design is difficult to gauge, because details of the developments are most often kept in-house.



Abstracted Open Software Architecture

The term “bloatware” is part of the PC software lexicon as a result of inefficient object-oriented implementations. These poorly designed applications and operating systems are characterized by a tendency, following upgrade, to consume far more hardware resources (RAM, disk space, and CPU cycles) than the previous version for questionable improvements in functionality. As an

example, the Microsoft NT operating system reportedly expanded from 16 million lines of code in version 4 to approximately 40 million lines of code in version 5.

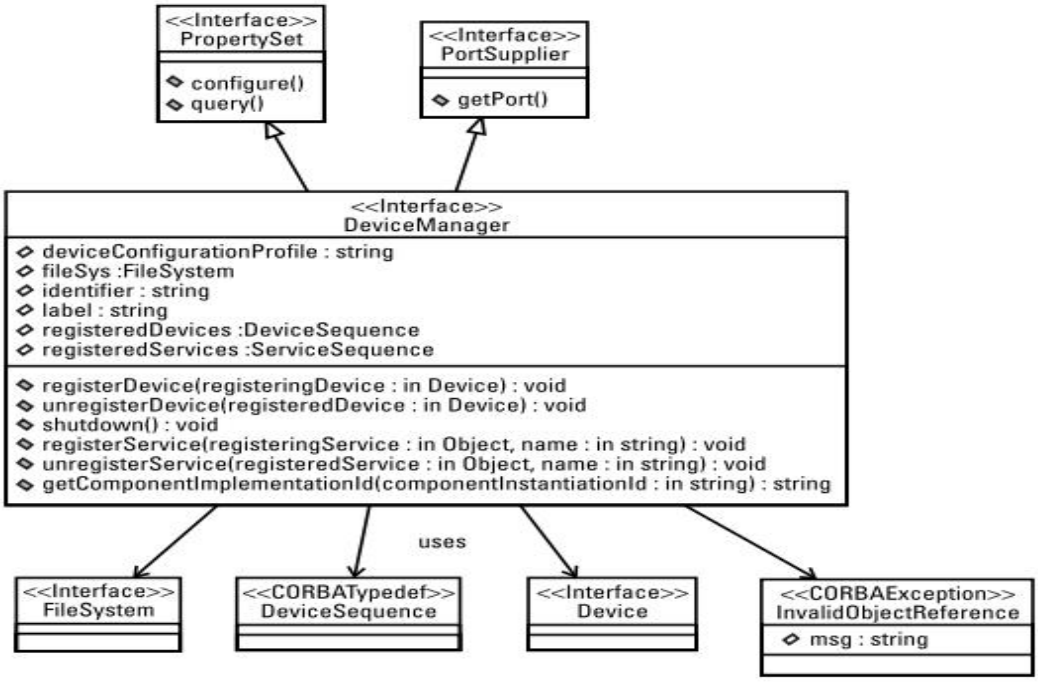
Software Standards for Software Radio

There are two key organizations pushing forward the adoption of software standards for the development of software defined radios. The Software Communications Architecture Specification (SCAS) [3] is published by the United States Joint Tactical Radio System (JTRS) Joint Program Office (JPO). This defense program has set goals for future communications systems (i.e., to increase flexibility and interoperability; ease up grade ability; and reduce acquisition, operation, and support costs). The JTRS states that the SCAS is not a system specification but a set of rules that constrain the design of systems to achieve these objectives. The U.S. government expects the basic SCAS to become an approved commercial standard through the Object Management Group (OMG) and has designed the specification to meet commercial as well as military application requirements. The OMG is becoming more involved in software radio specification and has created a special interest group.

6).Discuss briefly about architecture partitions of SDR with diagrams. (L-2,CO-2)

The SCAS provides guidance on partitioning the SDR hardware using an object-oriented (OO) approach. The OO method describes a hierarchy of hardware class and subclass objects that represent the architecture. Class structure is a hierarchy that depicts how object-oriented classes and subclasses are related. The class structure in the SCAS identifies functional elements that are used in the creation of physical system elements or hardware devices. As per the OO approach, devices inherit from their parents and share common physical and interface

attributes; theoretically, this should make it easier to identify and compare device interchangeability.

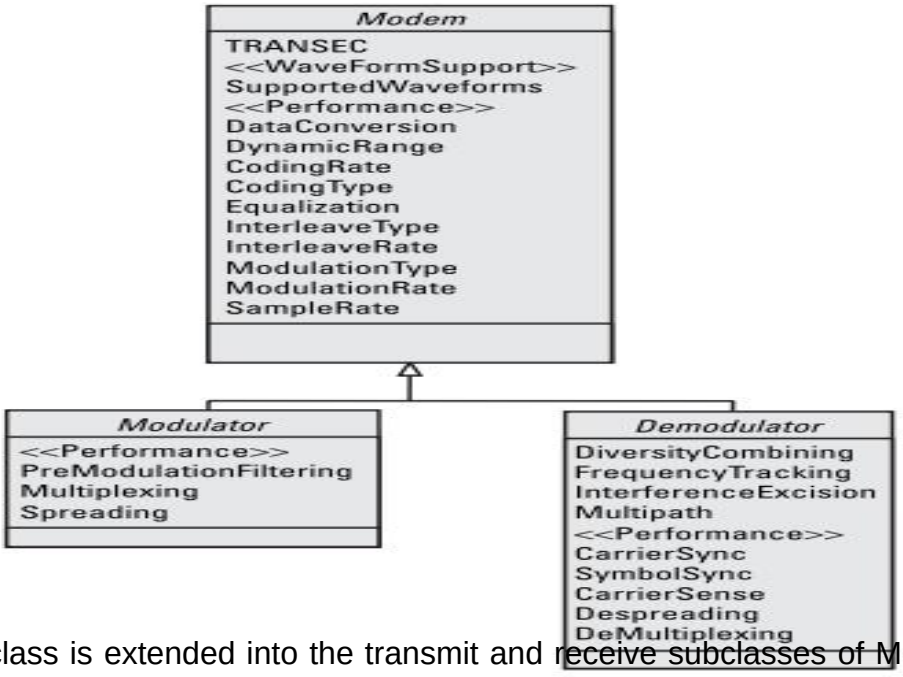


The overall hardware parent is the SCA-Compliant Hardware class; it defines attributes such as maintainability, availability, physical, environmental, and device registration parameters. SCA-Compliant Hardware has two child classes: Chassis and HW Modules. The Chassis subclass includes the attributes of module slots, form factor, back plane type, platform environmental, power, and cooling requirements. HW Modules is the parent to all module subclasses (e.g., RF

Power Supply, Modem, GPS, Processor, Reference Standard, and I/O). Each of the hardware child classes can be further extended, and examples of the granularity of the extensions are depicted in Figure 7.12 for the RF class and Figure 7.13 for the Modem

class. The RF class is extended by the addition of Antenna, Receiver, and Power Amplifier child classes. The Receiver class includes many of the parameters discussed in previous chapters.

Cosite Mitigation implies careful monitoring and control of interference-related parameters; this issue is expected to be a significant one, especially during the deployment of cosited 2G TDMA and 3G TDMA/CDMA systems. The SCAS notes that antennas have been historically passive elements attached to the structure that houses the communications system. In anticipation of technological advances and "smart" antenna (see Chapter 9) functionality, Antenna is included as an RF subclass with smart antenna parameters, such as BeamSteering and Nulling.



The modem class is extended into the transmit and receive subclasses of Modulator and Demodulator, as illustrated in Figure 7.13. For a multiple air interface mobile cellular radio the SupportedWaveforms attribute may have the valid values of GSM, IS 136, IS-95B, CDMA2000-1xRTT, or UMTS-FDD.

7)Explain RF conversation architecture? (L-2,CO-2)

RF CONVERSION ARCHITECTURES

The RF conversion segment of the canonical software radio is illustrated in Figure 8-1. The antenna segment may provide a single element for both transmission and reception. In this case, a multicoupler, circulator, or diplexer protects the receiver from the high-power transmission path. In other cases, the transmit and receive antennas may be physically separate and may be separated in frequency. First-generation cellular radio and GSM systems separate downlink and uplink bands by typically 45 MHz to limit interference.

Power amplifiers have less-than-ideal performance, including amplitude ripple and phase distortion. Although these effects may be relatively small, failure to address them may have serious consequences on SDR performance. Amplitude ripple, for example, degrades the transmitted power across the band, particularly near the band edges.

IF processing may compensate by preemphasizing the IF signal with the inverse of the power amplifier's band-edge ripple. Feher [238] describes techniques for compensating a sequence of channel symbols, shaping the transmitted waveform in the time domain to yield better spectral purity in the frequency domain. The concept behind Feher's patented design is straightforward. Sequential symbols may have the same relative phase, yet the channel-symbol window in which the sinusoids are generated modulates the amplitude at the symbol boundaries. When adja-cent symbols have different phase, this symbol weighting reduces frequency domain sidelobes and hence adjacent-channel interference.

Feher suppresses the modulation further with an extended symbol that includes the sequential symbols of the same phase generated with constant amplitude, thus without the weighting-induced amplitude modulation. The result is that energy that normally is redirected into the adjacent channels by the phase discontinu-ities remains within the channel because the discontinuities have been sup-pressed.

The receiver subsystem intersection with the RF conversion segment is shown in Figure 8-1 also. This includes the low noise amplifier (LNA), one or more stages of bandpass filtering (BPF), and the translation of the RF to an IF. In conventional radios, a tunable-reference local oscillator (LO) may be shared between the transmitter and receiver subsystems. FH radios of-ten share a fast-tuning LO between the transmitter and receiver.

- In military applications, the LO executes a frequency-hopping plan defined by a trans-mission security (TRANSEC) module. In commercial systems (e.g., GSM), a fixed frequency-hopping plan that suppresses fades may be used instead of a complex TRANSEC plan. The radio then either transmits or receives on the frequency to which the LO is tuned. Any radio which employs a physi-cally distinct programmable LO may be a programmable digital radio (PDR), a type of SDR, but it is not a software radio.

- Software radios use lookup tables to define the instantaneous hop frequencies, not physical LOs. This ap-proach, of course, requires a wideband DAC. One advantage of using such a DAC is that the hop frequency settles in the time between DAC samples, typically Wa=2:5—hundreds of nanoseconds. The hop frequency is pure and stable

instantly, subject to minor distortions introduced by the final power amplifier.

Since the receiver must overcome channel impairments, it may be more complex and technically demanding than the transmitter. Thus, this chapter focuses on receiver design.

Again referring to Figure 8-1, IF processing may be null, as may baseband processing. The direct conversion receiver, for example, modulates a reference signal against the received RF (or IF) signal to yield a baseband binary analog waveform in the in-phase and quadrature (I&Q) channels.

Although this kind of RF conversion has nonlinear characteristics, it is particularly effective for single-user applications such as handsets. It may not work well for multiuser applications, however.

This chapter examines the SDR implications of the RF conversion segment. The following section describes receiver architectures. Programmable component technology including MEMS and EPACs is described. RF subsystem specifications are then analyzed. The chapter concludes with an assessment of RF/IF conversion architecture tradeoffs.

8)Explain the Reviews of ADC fundamental? (L-3,CO-2)

REVIEW OF ADC FUNDAMENTALS

Since the wideband ADC is one of the fundamental components of the soft-ware radio, this chapter begins with a review of relevant results from sam-pling theory. The analog signal to be converted must be compatible with the capabilities of the ADC or DAC. In particular, the bandwidths and linear dy-namic range of the two must be compatible. Figure 9-1 shows a mismatch between an analog signal and the ADC. For uniform sampling rate fs, the maximum frequency for which the analog signal can be unambiguously re-constructed is the Nyquist rate, fs=2. The wideband analog signal extends beyond the Nyquist frequency in the figure. Because of the periodicity of the sampled spectrum, those components that extend beyond the Nyquist fre-quency fold back into the sampled spectrum as shown in the shaded parts of the figure (thus the term *folding frequency*). This is well known as alias-ing [274, 275]. Although some aliasing is unavoidable, an ADC designed for software-radios must keep the total power in the aliased components below the minimum level that will not unacceptably distort the weakest subscriber signal.

A. Dynamic Range (DNR) Budget

If acceptable distortion is defined in terms of the BER, then dynamic range (DNR) may be set by the following procedure:

1. Set BERTHRESHOLD from QoS considerations
2. BER = f(MODULATION, CIR, FEC)
3. BER < BERTHRESHOLD " CIR > CIRTHRESHOLD, from f( )
4. $DNR = DNR_{ADC} + DNR_{RF\#IF} + DNR_{OVERSAMPLING} + DNR_{ALGORITHMS}$
5. $P_{ALIASING+RFIF+NOISE} < \frac{1}{2}^{(DNR_{ADC} + CIR_{THRESHOLD})}$

Consider the situation where the channel symbol modulation, MODULA-TION, is fixed (e.g., BPSK). BER is a function of the CIR. The first step in es-

tablishing the acceptable aliasing power is to set the BERTHRESHOLD by considering the QoS requirements of the waveform (e.g., voice). The BERTHRESHOLD for PCM voice is about $10^{\#3}$. The next step is to characterize the relationship between BER and CIR. In the simplest case, this relationship is defined in the BER-SNR (CIR or Eb/No) curve for MODULATION (e.g., from [275]). In other cases, FEC reduces the net BER for a given raw BER from the mo-dem. In such cases, net BER has to be translated into modem BER using the

properties of the FEC code(s) [276, 277]. BERTHRESHOLD is then translated to CIRTHRESHOLD using f (e.g., 11 dB). Finally, one must incorporate the instantaneous dynamic range requirements of the ADC. Total dynamic range must be partitioned into dynamic range that the AGC, ADC, and algorithms must supply.

B. Anti-Aliasing Filters

When the aliased components are below the minimum acceptable power level (e.g., $\frac{1}{2}$ LSB) the sampled signal is a faithful representation of the analog signal, as illustrated in Figure 9-2. The wideband ADC, therefore, is preceded by anti-aliasing filter(s) that shape the analog spectrum to avoid aliasing. This requires anti-aliasing filters with sufficient stop-band attenuation. Figure 9-3 shows the stop-band attenuation required for a given number of bits of dy-namic range. Since the instantaneous dynamic range cannot exceed the reso-lution of the ADC, the number

of bits of resolution is a limiting measure of the dynamic range. High dynamic range requires high stop-band attenuation. To reduce the power of out-of-band energy to less than $\frac{1}{2}$ LSB, the stop-band attenuation of the anti-aliasing filter of a 16-bit ADC must be #102 dB. This includes the contributions of all cascaded filters including the final anti-aliasing filter.

C. Clipping Distortion

In most applications, one cannot control the energy level of the maximum signal to be exactly equal to the most significant bit. One must therefore allow for some AGC or for some peak power mismatch. Clipping of the peak energy level introduces frequency domain sidelobes of the high power signal. These sidelobes have the general structure of the convolution of the signal's sinusoidal components with the Fourier transform of a square wave, which has the form of a sin(x)=x function. Frequency domain sidelobes have a power level of #11 dB, which is clearly unacceptable interference with other signals in a wideband passband. In practice, avoiding clipping may occupy the entire most significant bit (MSB). Usable dynamic range may therefore be one or two bits less than the ADCs resolution.

D. Aperture Jitter

Sample-and-hold circuits also limit ADC performance as illustrated in Fig-ure 9-5. Consider a sinusoidal input signal, V(t) = Acos(!t), where ! is the maximum frequency. The rate of change of voltage is as shown, yielding a maximum rate of change of $2A=(2^{B})$ or $A=(2^{(B+1)})$. The time duration of this differential interval is inversely proportional to the frequency and the exponential of the number of bits in the ADC. This period is the aperture uncertainty, the shortest time taken for a maximal-frequency sine wave to traverse the LSB. The timing jitter must be a small fraction of the aperture uncertainty to keep the total error to less than $\frac{1}{2}$ LSB. Therefore, the

timing jitter should be 10% or less of the uncertainty shown in the figure. An 8-bit ADC sampling at 50 MHz requires aperture jitter that is less than a picosecond (ps).

E. Quantization and Dynamic Range

Quantization step size is related to power according to [279]:

$$Pq = q^2=12R$$

where q is the quantization step size, and R is the input resistance. The SNR at the output of the ADC is

$$SNR = 6:02\ B + 1:76 + 10\ \log(fs=2fmax)$$

where B is the number of bits in the ADC, fs is the sampling frequency, and fmax is the maximum frequency component of the signal.

For Nyquist sampling, fs = 2fmax, so the ratio of these quantities is unity. Since the log of unity is zero, the third term of the equation for SNR above is eliminated. The approximation for Nyquist sampling, then, is that the dynamic range with respect to noise equals 6 times the number of bits. This equation suggests that the SNR may be increased by increasing the sampling rate be-yond the Nyquist rate. This is the principle behind the *sigma-delta/delta-sigma* ADC.

F. Technology Limits

The relationship between ADC performance and technology parameters has been studied in depth by Walden [280, 281]. His analysis addresses the elec-tronic parameters, aperture jitter, thermal effects, and conversion-ambiguity. These are related to specific devices in Figure 9-6. The physical limits of ADCs are bounded by Heisenberg's uncertainty principle. This core phys-ical limit suggests that one could

implement a 1 GHz ADC with 20 bits (120 dB) of dynamic range. To accomplish this, one must overcome thermal, aperture jitter, and conversion ambiguity limits.

9)Briefly explain the applications of SDR? (L-2,CO-2)

SDR APPLICATIONS

ADC and DAC applications are constrained by sampling rate and dynamic range. The pace of product insertion into wireless devices is also determined by power dissipation. Infrastructure applications that are not power-constrained may evolve toward digital RF. This section highlights these aspects of ADC and DAC applications.

A. Conversion Rate, Dynamic Range, and Applications

ADC sampling rates and dynamic range requirements depend on the appli-cation. Figure 9-10 shows how increasingly wideband applications require increasingly large instantaneous dynamic range. Analog filtering and AGC achieve 90 to 100 dB or more of total dynamic range. As one increases the instantaneous bandwidth, one must also increase the instantaneous DNR as shown in Figure 9-10. It differentiates baseband (BB), IF, and RF ADC re-quirements. Baseband refers to the bandwidth of modulation of a single RF carrier

B. ADC Product Evolution

the relationship of commercially available ADC perfor-mance to research devices, emerging technology, and maximum requirements from Figure 9-10. Many viable

SDR applications are workable with currently available technology. Fielded applications include baseband digital signal pro-cessing in programmable digital radios. Emerging applications include SDRs that use IF conversion and parallel ADC channels to obtain high dynamic range. SPEAKeasy I and II, for example, both employed IF conversion with moderate (1 MHz) and wideband (70 MHz) ADC channels. The dynamic range of these implementations did not fully address the maximum require-ments for radio applications.

C. Low-Power Wireless Applications

The recent evolution of ADC product has been driven significantly by the wireless marketplace. Handheld commercial audio devices motivate invest-ment in devices with less than 1 MHz sampling rates but more than 100 dB SNR. Wireless handset applications provide much of the impetus behind low-power wideband ADC chips. Figure 9-12 shows the difference in sampling rate and dynamic range between low-power ADCs and ADCs for board-level products (e.g., for research and laboratory instrumentation markets). The 10-and 12-bit 70 MHz ADCs are rapidly evolving to 14-bit products.

D. Digital RF

As ADCs continue to evolve, they will enable the digital RF architecture illus-trated in Figure 9-13. Traditional RF subsystems include preamplifiers, LNAs, filters, RF distribution, and frequency translation and filtering stages that trans-late RF to usable IF signals. Such RF subsystems may comprise upward of 60% of the manufacturing cost of a radio node. These subsystems require large amounts of expensive touch-labor to assemble waveguide, coaxial cable, and other discrete components. The

digital RF alternative, also shown in Fig-ure 9-13, uses a preamplified ADC and multiplexer at the antenna to create a Gbps fiber optic signal [292]. Digital RF distribution via gigabit fiber optics weighs less and costs less per meter than RF distribution via coax or waveg-uide. In addition, fiber optics costs less to install and maintain than coax and waveguide. Lack of dynamic range, digital-RF's major shortfall at this time, can be enhanced using digital filtering techniques.

Unit III

Introduction To Cognitive Radios

Part-A

(1) How to make a radio cognitive? (L-2, CO-3)

We designed a cognitive radio software package, called a Cognitive Engine (CE), overlaid on radio hardware platform. CE manages radio resource to accomplish cognitive functionalities and adapts radio operation for performance optimization.

(2) What is the cognition model? (L-1, CO-3)

A specific machine learning model (embedding the two-loop cognition cycle as the learning core) is designed to enable cognition capability for wireless applications. Case based reasoning and evolutionary search are combined in the learning process.

(3) Which host radio architecture? (L-2, CO-3)

Any software defined radio platform and any radio with a certain level of reconfigurability can be supported by a cognitive engine with the platform independent radio interface.

(4) Which communication layers have cognition? (H-2, CO-3)

Currently the cognitive radio functionality is focusing on the PHY and MAC layers for cross-layer optimization. The designed cognition algorithms can easily be extended to the network and application layers due to its general learning core.

(5) How to establish a cognitive wireless network? (L-2, CO-3)

CWT2 provides a cognitive radio node which can be deployed for both centralized and distributed network intelligence. As network nodes, cognitive radios can work individually or jointly on resource management and performance optimization.

(6)**Define orient in Congnitive cycle. (L-1, CO-3)**

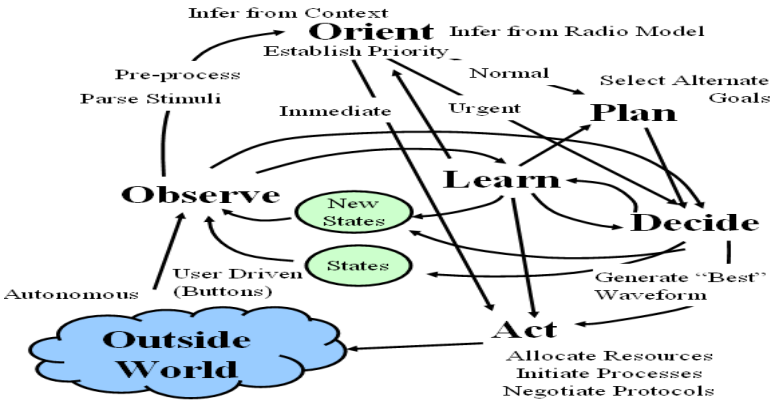
The orient phase determines the significance of an observation by binding the observation to a previously known set of stimuli of a “scene.” The orient phase

contains the internal data structures that constitute the equivalent of the short-term memory (STM) that people use to engage in a dialog without necessarily remembering everything with the same degree of long-term memory (LTM).

(7) **What is cognition cycle? (L-2, CO-3)**

This cycle implements the capabilities required of iCR in a reactive sequence. Stimuli enter the CR as sensory interrupts, dispatched to the cognition cycle for a response. Such an iCR continually observes (senses and perceives) the environment, orients itself, creates plans, decides, and then acts. In a single-processor inference system, the CR's flow of control may also move in the cycle from observation to action. In a multiprocessor system, temporal structures of sensing, preprocessing, reasoning, and acting may be parallel and complex. Special features synchronize the inferences of each phase.

8) **Draw the cognition cycle. (L-2, CO-3)**



9) **What do you meant by Optimization? (L-2, CO-3)**

Optimization is defined as a capability achieved through refining the solution and adapting practice accordingly. Optimal performance is a goal of such operations by directing the practice toward better results. List some characteristics of Radio Cognition Task.

10) **What is meant by Waveform? (L-1, CO-3)**

Waveform is defined as a super set of PHY parameters describing the format of a communication signal (PHY) and its related processing protocols (MAC, LLC, Net, etc.). This parameter set completely defines the wireless method of transceiving information between two communicating nodes. Such definition conforms to the waveform definition of the Software Communication Architecture (SCA) [41]. Such a waveform definition supports the standardization and portability of software defined communication applications.

11) **What is position awareness in cognitive radio? (H-1, CO-3)**

For cognitive radio (CR) to reach its full potential as an efficient member of a network or as an aid in users' daily tasks, and even to conserve the precious spectrum resource, a radio must primarily know its position and what time it is. From position and time, a radio can: (1) calculate the antenna pointing angle that best connects to another member of the network; (2) place a transmit packet on the air so that it arrives at the receiver of another network member at precisely the proper time slot to minimize interference with other users; or (3) guide its user in his or her daily tasks to help achieve the user's objectives, whether it be to get travel directions, accomplish tasks on schedule, or any of a myriad of other purposes. Position and time are essential elements to a smart radio. Furthermore, from position and time, velocity and acceleration can be inferred, giving the radio some idea about its environment

12) **Define spectrum pooling. (L-2, CO-3)**

Spectrum pooling is a spectrum management strategy in which multiple radiospectrum users can coexist within a single allocation of radio spectrum space.

13) **What is a LORAN? (L-1, CO-3)**

LORAN systems transmit a known burst signal from multiple transmitters with a known and published periodicity. Furthermore, the exact location of each transmitter is known. Three such transmitters cooperate to enable TDoA measurements. Ships at sea receive these transmissions and measure the time difference between each received signal.

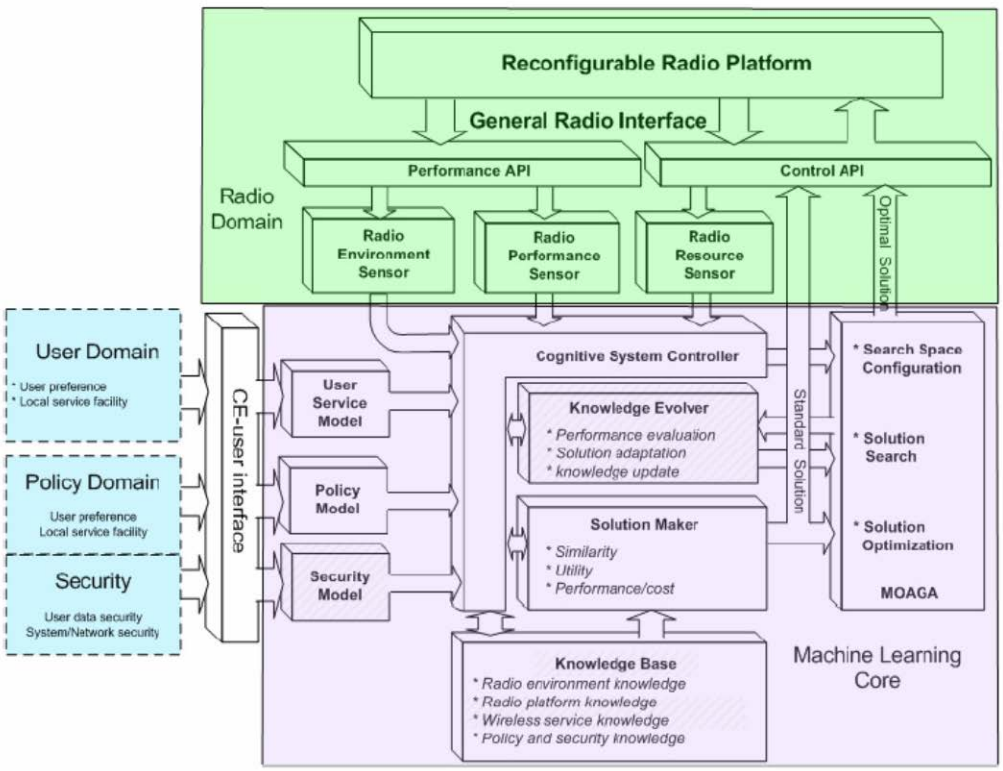
From these time differences, ships are able to calculate the TDoA hyperbolas. To simplify the process, the TDoA hyperbolas have been converted into published charts so that a navigator can directly look up the time differences for each transmitter pair and find an intersection of two time difference pairs to perform a location at sea

14) **What is menat by OCON-ANN? (L-1, CO-3)**

A One-Class-One-Network ANN (OCON-ANN) structure is constructed, in which one sub-network is created for each modulation type and these networks each output a value, a probability of a match according to the input signal. There is a judgment network, called MAXNET [101] collects the outputs from all these subnets, and the one with the highest output value wins as the modulation type

PART-B

1) Draw the cognitive radio framework and explain each block (H-2, CO-3)



The design objective of machine-learning capability leads to the development of a cognition cycle where artificial intelligence keeps evolving through the loop of reasoning, decision making, adaptation and knowledge accumulation. As the creator of the cognitive radio term, Mitola also proposed a cognition cycle which is well known and widely cited [8]. However, there are several reasons that the CWT2 defined a new cognitive cycle instead of using Mitola's. First, Mitola's cognitive cycle is defined at the application layer, with a simplified model of low-level communication layer (PHY and MAC) behaviors and optimistic capabilities assumptions, especially for 31 wireless links and networks. This makes it relatively difficult to realizing its envisioned cognition directly in the radio domain. Second, Mitola's cognitive cycle is defined to realize human-like cognitive capabilities which largely rely on today's super computation. These are well beyond the capability and scope of current radio engineering and device technologies. Third, the functional diagram is a bit complicated for embedded system design as the widely-spread relations are not suitable for efficient finite-state-machine design, and the definitions of functional nodes are too general or vague to be implemented directly with computational techniques. In 2004, I felt that a straightforward machine learning cycle is needed for radio's cognition, which should be straightforward with clear functional definitions but complete enough to carry general machine learning capabilities for various wireless applications. The skeleton of its autonomous running mechanism is a finite state machine that seamlessly connects the radio platform and artificial intelligence together. Therefore a new cognition cycle was defined. It has two feedback loops that realize two levels of intelligence; it also separates the general machine-learning core from radio-specific operations, which reflects the platform independence of the cognitive engine system, as stated in previous sections. As shown in Figure 2.6, the cognition cycle has two layers of loops. The outer loop consists of information recognition and behaviour adaptation, which are directly coupled with radio domain knowledge. The inner loop is a machine-learning loop where artificial intelligence methods are tailored and combined for a general solution making and self learning. This two-loop cognition cycle matches the two-layer hierarchical machine learning structure. Recalling the Egg Model in Figure 2.2, the outer layer consists of radio domain-specific

operations; while the inner layer is a general machine-learning core. 32 Figure 2.6: Cognition cycle block diagram The full cognition cycle collects environment information from the recognition modules, synthesizes this information into parametric models for scenario representation, compares this scenario with remember knowledge and decides whether to directly apply a previous solution, adapt a previous solution, or develop a new solution, estimates the solutions' anticipated performance with formulated objectives, applies the solution to the radio platform and monitors the actual performance, compares the real performance with the anticipated performance, records the differences as lessons, and updates the knowledge base associated with this scenario-solution pair for future use. Steps and are carried by the outer loop that provides environment recognition and practice adaptation, together with the radio platform; and steps

and are carried by the inner loop that forms the machine learning cycle. Such a tiered structure has several advantages. First, it enables platform independence. General learning should 33 be able to cooperate with different radio platforms and its intelligence should be applicable to various scenarios. A lower layer of radio domain-specific intelligence serves as middleware that interprets and abstracts the environment into standard representation and procedure for the learning core. This reduces the complexity of developing machine learning algorithms for the higher level intelligence by avoiding domain-specific considerations. Second, it simplifies system interface design. API efficiency and transparency are important to improve system efficiency, functional verification, operation management, and code reuse. Since the cognitive engine interacts with both software and hardware for a real-time performance response, the API between the cognitive algorithm and radio platform should be designed to be as straightforward as possible. A radio-domain-specific layer takes care of the physical meaning of all domain parameters, thus allowing the API to be simply script parsing. It also supports parallel software development and maintenance. System upgrade and porting become flexible by allowing both layers of cognition to be under parallel development. This feature is especially important for applying software defined approach for radio platform design. In the cognition cycle, the outer loop serves the inner loop. It observes the environment and

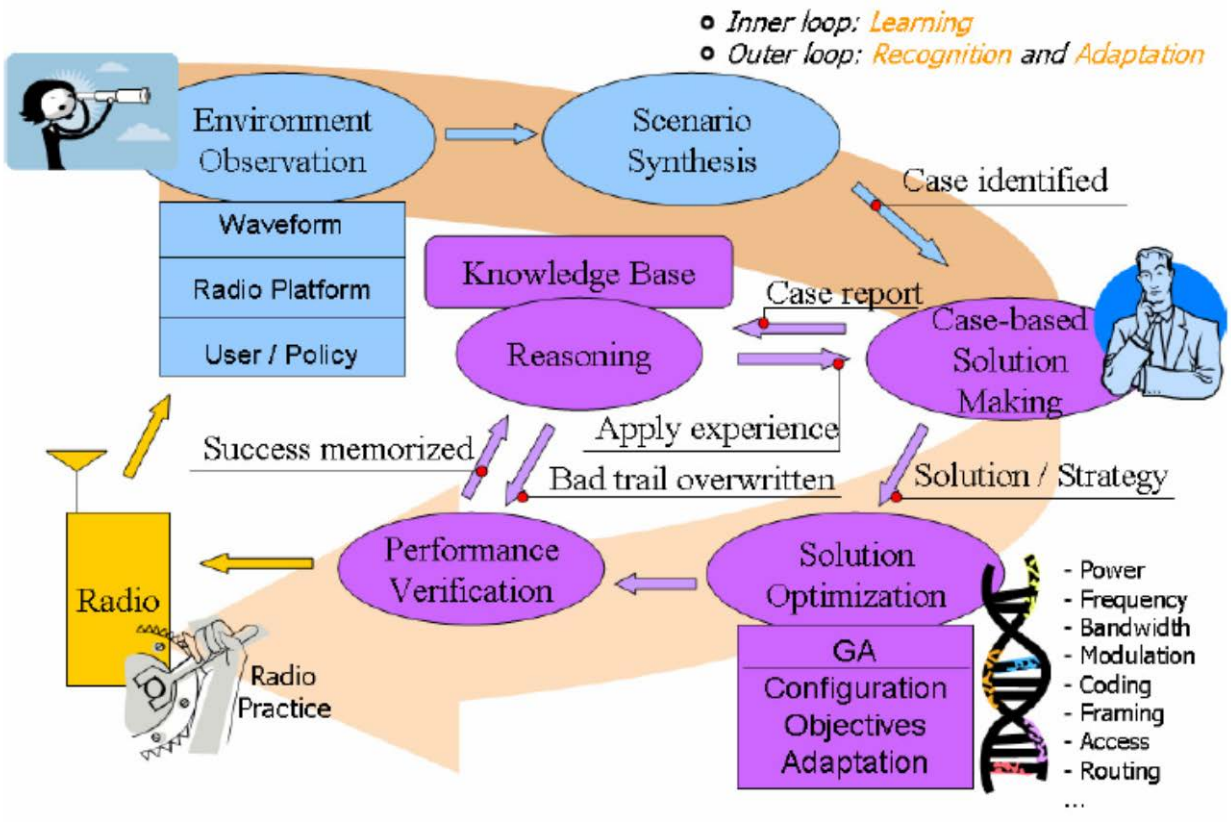
reports derived information like waveform features, interference and propagation channel characteristics, user service preferences, service policy and spectrum regulations, to the machine reasoning core. It also refines the solution from the reasoning core to improve performance, and formulates adaptation instruction, according to the solution, to feed the radio platform for action. The outer loop can be viewed as a lower-layer of cognition directly on top of radio application domain. Its intelligence level only guarantees that the environment observation is sufficient and accurate for the machine reasoning and that the adaptation exactly follows the instructions from the decision maker. The design of the radio environment recognition portion is detailed in Chapter 3. Generally speaking, learning is a cyclic process which consists of making decisions based upon the awareness of the environment and the available knowledge, applying them as instructions for current (and future) behaviours, then observing the results of practice and educating itself by updating the knowledge base for future practice. This machine 34 learning approach is commonly referred to as knowledge based learning, and the knowledge is evolved through the iterative reinforcement learning processing. The inner loop, the learning loop, follows this learning approach. It synthesizes the problem scenario interpreted from the lower layer recognition and the performance objectives interpreted from the user service demands, and then search its database to associate relevant knowledge for solutions suitable for this interpreted situation. The reasoning process may use pre-loaded decision principles or accumulated experience from past practice, or both. When facing novel or unknown situations where a suitable knowledge instance is difficult to associate, a generally viable (but not optimal) solution can be further adapted with an evolutionary algorithm to fit such novelty; or a completely creative solution can be constructed through evolutionary search for a totally unknown situation. With rationality and creativity combined in solution making, the solution made and then the action taken have a higher probability of success facing an unknown situation, and meanwhile the action to take can be guaranteed correct facing a known situation. The performance of practice is evaluated and then reported to the learning core as the lesson associated to the deployed solution. In such a way, the knowledge is updated through experience.

Specifically, AI techniques are applied in both the outer and inner loops. Statistical pattern recognition and feature clustering methods are designed for radio environment recognition; the multi-objective genetic search method is designed for novel solution search to guide the radio's adaptation. In the inner loop, case based reasoning (CBR) is the primary solution making method, and is aided by genetic search to enhance solution adaptability., and AI for inner loop is detailed in Chapter 4. Knowledge is the key in machine learning. It can be divided into two categories based on different sources. One type of knowledge is experience that includes situations, actions taken, and their consequences; the other consists of pre-set principles and non-adaptive examples. For machine learning system, the first type knowledge is achieved from selfpractice, just like in humans, while the second type can be simply pre-loaded rather than 35 through the long-term education that humans require. For the CWT2 cognitive radio system, the knowledge is implemented as a relational database, called the Cognitive Radio Knowledge Base (CRKB). CRKB is a metadata base that consists of several sub-databases, such as the radio environment map (REM) [40] for environment awareness, user service knowledge for performance objectives, case base knowledge for scenario-solution association, radio resource knowledge for solution boundary, and regulation knowledge for legality and security verification. As stated in Section 2.1.3, the CRKB concept matches the functional view of cognition. It can be flexibly implemented in a distributed way, unifying the cognitive behaviour of both CR nodes and cognitive networks.

2) With a neat diagram, explain the simplified cognition cycle. (L-2, CO-3)

The design objective of machine-learning capability leads to the development of a cognition cycle where artificial intelligence keeps evolving through the loop of reasoning, decision making, adaptation and knowledge accumulation. As the creator of the cognitive radio term, Mitola also proposed a cognition cycle which is well known and widely cited [8]. However, there are several reasons that the CWT2 defined a new cognitive cycle instead of using Mitola's. First, Mitola's cognitive cycle is defined at the application layer,

with a simplified model of low-level communication layer (PHY and MAC) behaviors and optimistic capabilities assumptions, especially for 31 wireless links and networks. This makes it relatively difficult to realizing its envisioned cognition directly in the radio domain. Second, Mitola's cognitive cycle is defined to realize human-like cognitive capabilities which largely rely on today's super computation. These are well beyond the capability and scope of current radio engineering and device technologies. Third, the functional diagram is a bit complicated for embedded system design as the widely-spread relations are not suitable for efficient finite-state-machine design, and the definitions of functional nodes are too general or vague to be implemented directly with computational techniques. In 2004, I felt that a straightforward machine learning cycle is needed for radio's cognition, which should be straightforward with clear functional definitions but complete enough to carry general machine learning capabilities for various wireless applications. The skeleton of its autonomous running mechanism is a finite state machine that seamlessly connects the radio platform and artificial intelligence together. Therefore a new cognition cycle was defined [20]. It has two feedback loops that realize two levels of intelligence; it also separates the general machine-learning core from radio-specific operations, which reflects the platform independence of the cognitive engine system, as stated in previous sections. As shown in Figure 2.6, the cognition cycle has two layers of loops. The outer loop consists of information recognition and behavior adaptation, which are directly coupled with radio domain knowledge. The inner loop is a machine-learning loop where artificial intelligence methods are tailored and combined for a general solution making and selflearning. This two-loop cognition cycle matches the two-layer hierarchical machine learning structure. Recalling the Egg Model in Figure 2.2, the outer layer consists of radio domain-specific operations; while the inner layer is a general machine-learning core.



The full cognition cycle (1) collects environment information from the recognition modules, (2) synthesizes this information into parametric models for scenario representation, (3) compares this scenario with remember knowledge and decides whether to directly apply a previous solution, adapt a previous solution, or develop a new solution, (4) estimates the solutions' anticipated performance with formulated objectives, (5) applies the solution to the radio platform and monitors the actual performance, (6) compares the real performance with the anticipated performance, records the differences as lessons, and updates the knowledge base associated with this scenario-solution pair for future use. Steps (1), (2) and (5) are carried by the outer loop that provides environment recognition and practice adaptation, together with the radio platform; and steps (3), (4) and (6) are carried by the inner loop that forms the machine learning cycle.

Such a tiered structure has several advantages. First, it enables platform independence. General learning should 33 be able to cooperate with different radio platforms and its intelligence should be applicable to various scenarios. A lower layer of radio domain-specific intelligence serves as middleware that interprets and abstracts the environment into standard representation and procedure for the learning core. This reduces the complexity of developing machine learning algorithms for the higher level intelligence by avoiding domain-specific considerations. Second, it simplifies system interface design. API efficiency and transparency are important to improve system efficiency, functional verification, operation management, and code reuse. Since the cognitive engine interacts with both software and hardware for a real-time performance response, the API between the cognitive algorithm and radio platform should be designed to be as straightforward as possible. A radio-domain-specific layer takes care of the physical meaning of all domain parameters, thus allowing the API to be simply script parsing. It also supports parallel software development and maintenance. System upgrade and porting become flexible by allowing both layers of cognition to be under parallel development. This feature is especially important for applying software defined approach for radio platform design. In the cognition cycle, the outer loop serves the inner loop. It observes the environment and reports derived information like waveform features, interference and propagation channel characteristics, user service preferences, service policy and spectrum regulations, to the machine reasoning core. It also refines the solution from the reasoning core to improve performance, and formulates adaptation instruction, according to the solution, to feed the radio platform for action. The outer loop can be viewed as a lower-layer of cognition directly on top of radio application domain. Its intelligence level only guarantees (1) that the environment observation is sufficient and accurate for the machine reasoning and (2) that the adaptation exactly follows the instructions from the decision maker. The design of the radio environment recognition portion is detailed in Chapter 3. Generally speaking, learning is a cyclic process which consists of making decisions based upon the awareness of the environment and the available knowledge, applying them as instructions for current (and future) behaviors, then observing the results of practice and educating

itself by updating the knowledge base for future practice. This machine 34 learning approach is commonly referred to as knowledge based learning, and the knowledge is evolved through the iterative reinforcement learning processing [39]. The inner loop, the learning loop, follows this learning approach. It synthesizes the problem scenario interpreted from the lower layer recognition and the performance objectives interpreted from the user service demands, and then search its database to associate relevant knowledge for solutions suitable for this interpreted situation. The reasoning process may use pre-loaded decision principles or accumulated experience from past practice, or both. When facing novel or unknown situations where a suitable knowledge instance is difficult to associate, a generally viable (but not optimal) solution can be further adapted with an evolutionary algorithm to fit such novelty; or a completely creative solution can be constructed through evolutionary search for a totally unknown situation. With rationality and creativity combined in solution making, the solution made and then the action taken have a higher probability of success facing an unknown situation, and meanwhile the action to take can be guaranteed correct facing a known situation. The performance of practice is evaluated and then reported to the learning core as the lesson associated to the deployed solution. In such a way, the knowledge is updated through experience. Specifically, AI techniques are applied in both the outer and inner loops. Statistical pattern recognition and feature clustering methods are designed for radio environment recognition; the multi-objective genetic search method is designed for novel solution search to guide the radio's adaptation. In the inner loop, case based reasoning (CBR) is the primary solution making method, and is aided by genetic search to enhance solution adaptability. AI for outer loop is explained in Chapter 3, and AI for inner loop is detailed in Chapter 4. Knowledge is the key in machine learning. It can be divided into two categories based on different sources. One type of knowledge is experience that includes situations, actions taken, and their consequences; the other consists of pre-set principles and non-adaptive examples. For machine learning system, the first type knowledge is achieved from selfpractice, just like in humans, while the second type can be simply pre-loaded rather than 35 through the long-term education that humans require. For the

CWT2 cognitive radio system, the knowledge is implemented as a relational database, called the Cognitive Radio Knowledge Base (CRKB) [26]. CRKB is a metadatabase that consists of several sub-databases, such as the radio environment map (REM) [40] for environment awareness, user service knowledge for performance objectives, case base knowledge for scenario-solution association, radio resource knowledge for solution boundary, and regulation knowledge for legality and security verification. the CRKB concept matches the functional view of cognition. It can be flexibly implemented in a distributed way, unifying the cognitive behavior of both CR nodes and cognitive networks.

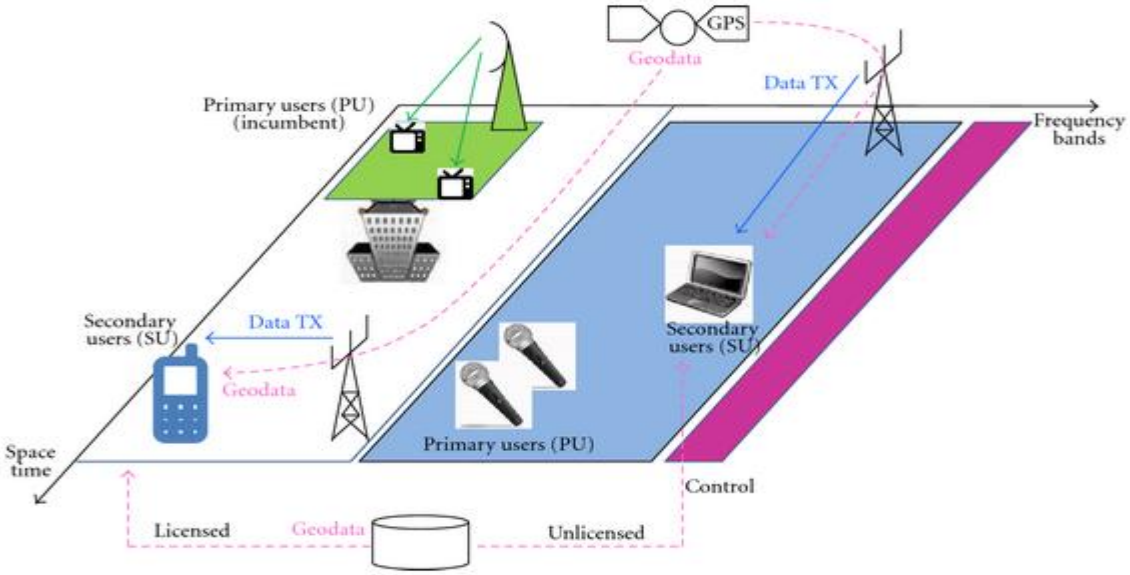
3) List and explain the characteristics of radio cognition task. (L-2, CO-3)

The information and communication technology industry is today faced with a global challenge: develop new services with improved quality of service (QoS) and at the same time reduce its environmental impact. Clearly, there is a deep need of global efficiency not only in the energy domain, but also in the spectral domain.

In order to address the problem of spectrum usage efficiency, the cognitive radio (CR) concept was proposed. The detailed definition of cognitive radio systems (CRSs) will be given in Section 2. Cognitive radio technology has the potential of being a disruptive force within spectrum management.

A very popular example is opportunistic radio or opportunistic spectrum access whose principle is temporal, spatial, and geographic “reuse” of licensed spectrum as shown in Figure 1 where an “unlicensed” secondary user (SU) can be permitted to use licensed spectrum, provided that it does not interfere with any primary users (PUs). In that way, the efficiency of spectrum usage is significantly improved. Various measurements of spectrum utilization have shown that spectrum is underutilized, in the sense that the typical duty cycle of spectrum usage at a fixed frequency and at a random geographical location is low. This means that there are many “holes” in the radio spectrum that could be exploited [3]. Opportunistic radio system should be able to exploit these spectrum holes by detecting them and using them in an opportunistic manner. Because of the outstanding propagation characteristic in the television (TV) bands with strong wall and

floor penetration capability, long range, and flexible bandwidth, it could be used to allow a brand-new class of services and increase the limited capacity of existing systems.



But CRS is not only limited to opportunistic spectrum access. Also, it includes heterogeneous networks where a heterogeneous radio framework management is performed. Current spectrum allocation approaches, fixed by nature, do not allow for the allocation of frequency bands to different radio access technologies (RATs) dynamically. However, the coexistence and cooperation of diverse technologies, which form part of a heterogeneous infrastructure, have brought about the possibility of flexibly managing the spectrum in a dynamic manner. No longer are fixed frequency bands guaranteed to apply to specific RATs, but conversely, through intelligent management mechanisms, bands can be allocated to RATs dynamically in a way such that the capacity of each RAT is maximized and interference is minimized. on long terms, will be considered also the flexibility in spectrum management where network operator may employ different RATs dynamically over time/frequency/location and acquire or exchange the spectrum usage rights. The devices may autonomously and dynamically adapt to the diverse heterogeneous radio networks.

This new technology opens up many potential applications and exciting opportunities. For example, the rural connectivity, content distribution networks, city and campus wide coverage, and giant wireless hotspots could benefit from the new spectrum access and management.

This paper is organized as follows: after this introduction, the definition and high level concept of CRS is presented in Section 2. In Section 3, the status on regulation and standardization activities are described. Section 4 will present the research challenged related to CRS. Section 5 will describe the implementation challenges of CR devices. The conclusion of the paper is drawn in Section 6.

Definition and High Level Concept

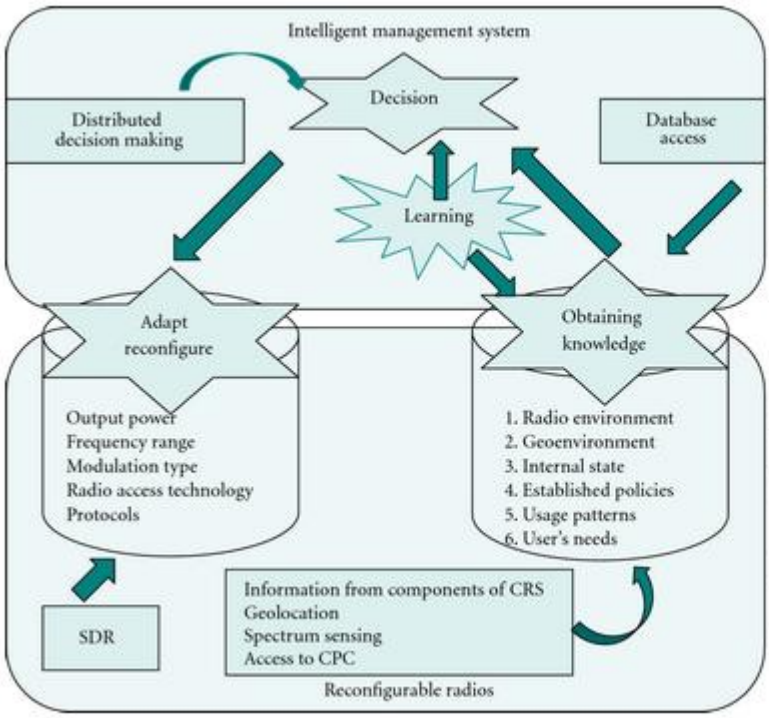
There are different definitions of CRS, from many authors and organizations. The definition giving the common understanding about CRS and now adopted for most is from International Telecommunication Union (ITU). CRS is a radio system employing technology that allows the system:

(i) To obtain knowledge of its operational and geographical environment, established policies, and its internal state (cognitive capability);

(ii) To dynamically and autonomously adjust its operational parameters and protocols according to its obtained knowledge in order to achieve predefined objectives (reconfigurable capability);

(iii)To learn from the results obtained (learning capability).

At high level concept presented in Figure 2, the main components of the CRS are the intelligent management system and reconfigurable radios . CRS is also able to take action including obtaining knowledge, making decision, reconfiguration, and learning. The knowledge used by the CRS includes knowledge about operational radio and geographical environment, internal state, established policies, usage patterns, and users' needs.



Status in Regulations and Standardizations

Regulations

The major regulatory agencies are developing rules for the unlicensed use of TVWS such as the FCC in the United States, Ofcom in the UK, and the Electronic Communications Committee (ECC) of CEPT in Europe.

The FCC provided the final rules for TVWS in 2010. There is ongoing proceeding for secondary use of the 2.36 GHz to 2.4 GHz band for medical area networks. Other opportunistic spectrum access beyond the already completed TVWS proceedings and cognitive techniques to better utilize the radio spectrum are currently under investigation. Ofcom has also made significant progress in developing regulations for the TVWS with a first public consultation in 2009. The statement on white spaces devices and

implementation of geolocation databases was released on September 1st, 2011. The detailed rules will be released in the future.

The ECC studied the technical and operational requirements for the operation of CRS in the WS of the UHF broadcasting band (470–790 MHz). This work is used as the starting point for regulatory activities within the ECC.

Standardizations

Currently international standardization of CRS is performed at all levels (ITU, IEEE, ETSI, and ECMA). They are considering multiple deployment scenarios and business directions.

In ITU, Working Party (WP) 1B has worked on the definition of SDR and CRS and their relationship and summarized the technical and operational studies, and relevant recommendations. It has considered the SDR and CRS usage scenarios in different radio services and regulation implications. The WP 5A is currently addressing the definition, description, and application of CRS in the land mobile service.

IEEE is very active in CRS. In 802 WGs (LAN/MAN), the activity to define CRSs is currently performed in the 802.11 and 802.22, while the activity to specify components of a CRS is currently performed in 802.19, 802.21, and 802.22. 802.11y is an amendment for 3650–3700 MHz operation in USA defining new regulatory classes, transmit power control, and dynamic frequency selection for 802.11 to share frequency bands with other users. Draft standard P802.11af is an amendment for TVWS operation defining standardized modifications to both the 802.11 physical (PHY) layers and medium access control (MAC) sublayer to meet the legal requirements for channel access and coexistence in the TVWS. Draft standard P802.19.1 concerns TVWS coexistence methods. IEEE 802.21 focuses on media independent handover services enabling the optimization of handover between heterogeneous IEEE 802 networks, and facilitating handover between IEEE 802 networks and cellular networks. The draft standard P802.22 is on policies and procedures for operation in the TV bands. It specifies the air interface, including the cognitive MAC and PHY, of point-to-multipoint wireless regional area

networks, operating in the unlicensed TV bands between 54 MHz and 862 MHz. Draft standard P802.22.1 is to enhance harmful interference protection for PUs operating in TV bands.

The ETSI Reconfigurable Radio Systems (RRS) Technical Committee (TC) is also active in standardizing SDR and CRS. TC RRS main responsibility is to carry out standardization activities related to reconfigurable radio systems (RRS) encompassing both SDR and CR with a focus on specific needs of the European Regulatory Framework, and CR/SDR TV white space standards adapted to the digital TV signal characteristics in Europe. Two out of the four working groups within ETSI RRS have activities resulting in standardization of potential regulatory aspects of CRS and SDR. Working group 3 has proposed and investigated the feasibility of standardizing a functional architecture for management and control of reconfigurable radio systems and cognitive pilot channel. SDR-related standardization is considered for both base station and mobile device. Working group 2 relies mainly on mobile device SDR related interface standardization. ETSI RRS is also working on operation in WS frequency bands and coexistence architecture for cognitive radio and investigating security and threats issues.

4)Explain the structuring knowledge for cognition tasks(H-2, CO-3)

Cognition tasks that might be performed range in difficulty from the goal-driven choice of RF band, air interface, or protocol to higher-level tasks of planning, learning, and evolving new protocols. characterizes these tasks in terms of nine levels of capability. Table 4-1 Characteristics of Radio Cognition Task Level Capability Task Characteristics 0 Pre-programmed The radio has no model-based reasoning capability

1 Goal-driven Goal-driven choice of RF band, air interface, and protocol

2 Context Awareness Infers external communications context (minimum user involvement)

3 Radio Aware Flexible reasoning about internal and network architectures

4 Capable of Planning Reasons over goals as a function of time, space, and context

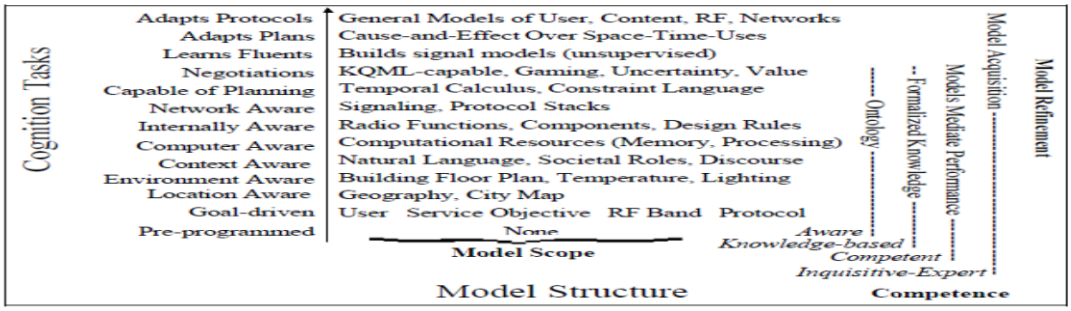
5 Conducts Negotiations Expresses arguments for plans/ alternatives to user, peers, networks

6 Learns Fluents Autonomously determines the structure of the environment

7 Adapts Plans Autonomously modifies plans as learned fluents change

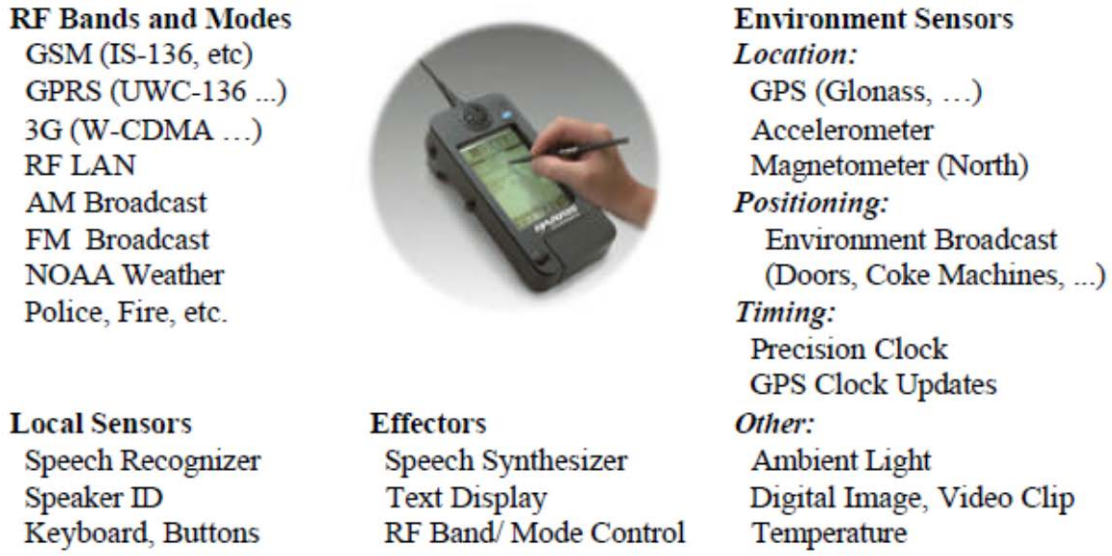
8 Adapts Protocols Autonomously proposes and negotiates new protocols

The Foundation for Cognition (Capability Levels 0-3) Cognition capability level 0 represents a conventional PDR, SDR, or software radio. In some cases, the tradeoff between network intelligence and mobile unit intelligence may favor the minimization of computational intelligence in the mobile unit. In the past, the power consumption of hand-held devices was a major design driver limiting processing capacity. During the past five years, however, semiconductor device density has increased 3.5 times while power consumption has been reduced by a factor of 40. By the year 2002, 0.12 micron production lines may again double device density while the 1.0 Volt power supplies probably will reduce power by another factor of At that point, even handheld devices will support the GFLOP processing capacities and hundreds of Mbytes of random access memory (RAM) needed for the high levels of processing capacity implicit in cognitive radio.



The pre-programmed level of cognition may therefore become limited to extremely small wearable devices such as badges supported by cognitive infrastructure. Goal-driven reasoning, cognition capability level 1 in the table, may be achieved with rulebased expert systems technology, and thus is not of research interest by itself. Environment- 51 awareness extends goal-driven reasoning to goals defined in a spatial context. This level of capability requires a suite of environment sensors and multiband RF like that

summarized in Figure 4-4. Approximate location may be known from a global positioning receiver like GPSor Glonass. GPS may not be available inside buildings, and it may not be reliable in urban areas. Therefore, additional positioning sensors are needed. Environments mapped via sensors with wireless read-out or broadcast can provide very detailed data about objects in one's immediate vicinity. Agent-based control of services using the resulting positioning and status information is being explored in a companion research project at KTH.



Context awareness (level 2) augments environment awareness in a way that is unique to cognitive radio. This entails the processing of incoming and out-going media in order to infer user communications context. This includes the detection of significant events that may shape the nature of computing and communications services as discussed in the use cases of the next chapter. The processing does not necessarily require in-depth machine understanding of every stimulus. On the contrary, in order to be practicable, all stimuli are scanned for surface features indicating that a context-defining event may be present. If such features warrant, in-depth analysis of the stimulus (e.g. using natural language techniques) may be performed to extract the parameters of the event. Alternatively, the user may be asked what to do in the presence of the surface

stimuli. It may then store the situation and user action as a problem-solution case for case-based reasoning. This is the approach taken in CR1, but that does not exclude in-depth 52 analysis from the resulting cognitive radio architecture. Level 2 also requires proficiency in air interfaces and protocols (e.g. GSM, DECT, RFLAN, 3G). One specific objective is to relieve the user from the need to manually select and configure air interface modes as 3G services enter the marketplace. In cognitive radio, this proficiency is obtained by modeling the air interfaces shown in the figure. For every air interface available to the radio, there is a corresponding set of internal models expressed in RKRL. These models describe the mode's behavior, parameters, and bindings by which the cognition cycle can control that mode. The level of detail of control is a function of the richness of the internal knowledge and the level at which the interfaces to those modules have been defined. For example, if the modem is a parameterized software module the cognitive radio could control the parameters of physical layer air interface. If the protocol uses a conventional stack like TCP, then it is likely to be highly encapsulated, limiting the cognitive radio's ability to tailor services. If, however, the protocol is dynamically defined then the intra-module interfaces may be modeled in RKRL and controlled by cognitive radio. If the protocol is defined in a customizable framework, then it may be tailored dynamically to the application. General awareness of the wireline network and its contrasting properties (e.g. short latency, large bandwidths) is also required for radio awareness. Level 3, radio awareness, in some sense turns the notion of sensing the environment inward. This includes the definition of interfaces between the operational software modules of the SDR and the cognitive control component. In addition, the radio's internal model must reflect substantial knowledge of the behavior of networks. Propagation modeling, QoS models, queuing models, and such system-level models would be characteristic of a cognitive network. A cognitive PDA would include a subset of that knowledge useful in support of local decisions. This could include which of several available RF modes (on different networks) to use given the user's constraints and the PDA's rate of movement and destination, which is a part of communications context for a mobile user.

Core Cognition Capabilities: Natural Language, Planning and Learning Achieving level 4 capability requires computer-based planning. There is a well-established technology base for computer-based planning [159, 160]. Applications to intelligent control 53 generally employ some form of hierarchy to modularize the knowledge [161]. A typical decision making architecture from the planning literature analyzes a situation to determine goals. The goals imply candidate plans, each of which is supported or refuted by arguments. Those arguments imply specific plans, which lead to actions, which influence the situation, closing the loop. Domino exemplifies the planning languages that have been published in the literature [162]. Since Domino has classical, temporal, and modal axioms, it represents axioms of beliefs as well as models of time. These approaches tend to ascribe logical structure to situations in which there may be none. For example, a network may change air interfaces or upgrade protocols arbitrarily. A user may change his mind arbitrarily. Thus, cognitive radio incorporates elements of planning technology, but in a way that relaxes the requirement for consistency and completeness in its (prototype) planning processes. Closely related to planning is level 5, negotiation capability. This is the subject of much current research stimulated by Internet commerce [163]. This includes managing conflicting plans [164]. Conducting negotiations with other radio entities requires an ability to execute negotiation protocols. These may be artificially constrained to finite-state grammars to insure prompt convergence. In addition, some user interactions may be organized as negotiation dialogs. For example, the user may ask for communications services that violate a-priori constraints. The PDA then should express the constraint violations (e.g. “Cannot send this file with a time delay of less than two minutes and cost below five dollars”). It should understand the user’s side of the negotiation (e.g. “Why not?”). It should generate an explanation (e.g. “The file size of two megabytes requires a data rate of over 384 kbps which increases your cost to seven dollars”). Finally, it should understand when the user has made a decision (e.g. “I do not care about cost on this email, so send it”). Thus, effective use of plan-generation capability in interactions with the user may employ negotiation dialogs. This requires substantial naturallanguage analysis and generation capability along with causality

analysis. Levels 6-8 require machine learning. Statistical learning of patterns is well in hand, but the reliable, incremental acquisition of (valid) new models remains on the frontiers of research. Learning applications in cognitive radio could include autonomously determining the structure of the radio environment as it changes. This would require the learning of fluents [149], events that have duration greater than zero as evidenced, for example, by temporal consistency in a time- 54 varying stream. An example of a fluent that that a cognitive radio might learn is the fact that the impulse response of a particular set of RF channels is usually not equalized "near the football stadium." A cognitive radio should modify its plans as the fluents change. For example, when the stadium is under renovation, its multipath reflection properties change, so the PDA no longer should us a low data rate when operating "near the stadium." If it always uses the low data rate from prior learning, then it will never discover the availability of the high data rate. This can be solved by advice from the network. But the network would only discover the feasibility of the higher data rate if some small percentage of subscribers attempt that higher data rate. Ultimately, cognitive radios with these higher capability-levels could propose and negotiate new protocols among themselves. In order to operate in these higher capability levels, radio domain knowledge must be organized for the performance of these cognition tasks.

5) What are the primary concepts of location aware cognitive radio? Explain with neat architecture. (L-2, CO-3)

In the emerging spectrum sharing paradigm, cognitive radio (CR) nodes seek spectrum opportunities that can be characterized both in temporal and spatial domains. Temporal spectrum holes provide us with prospective opportunities for transmission when a primary user (PU) is in idle state. Whereas spatial spectrum holes enable CRs nodes to coexist on occupied primary frequency bands without causing harmful interference to the PUs. A system capable of using both spatial and temporal spectrum holes is bound to increase system performance significantly. However spatial spectrum holes require either location sensing or location estimation i.e., localization of primary and secondary users (SUs), and this cost has to be balanced with performance gains offered. To understand

the tradeoffs, this work aims to evaluate performance of location aware cognitive radio networks in scenarios where CR users and PUs operate on overlapping frequency bands. This chapter focuses on the paradigm of location assisted dynamic spectrum management for scenarios where primary and secondary users operate on overlapping frequency bands. Using the spectrum utilization efficiency as performance metric, we analyze the performance gain offered by location awareness with respect to key system parameters of legacy networks and wireless channel parameters. The goal is to evaluate the feasibility of coexistence of CR users and PUs in spatial domain. The concept of probability of successful concurrent transmission (PSCT) is introduced in, but its evaluation hinges on some restrictive assumptions on the system setup, such as single channel availability, constant and same transmission power for both the CR users and PUs. In this paper we build up on the notion of (PSCT) and extend it to a much more realistic case. In lieu of the single-channel case in, we consider the multi-channel, multi-user case, in which PUs operate on orthogonal channels. In addition, PUs are equipped with adaptive transmit power control (ATPC), which enables them to adjust 25 Resource Management in CRNs - Single CR User 26 their transmission power in accordance to their distance from the intended receiver. We investigate the relationship between performance of overlaid CR ad hoc networks with respect to wireless channel parameters (path loss and shadowing) and system parameters (system capacity, detection thresholds and transmit power control). In addition, we provide analytical expressions for PSCT and average probability of successful concurrent transmission (Pavg,SCT). Our analysis also sheds light on the location sensing and estimation requirements for such overlaid cognitive radio networks, which provides useful guideline for system designers to trade off the benefits and costs of location aware networking.

System Setup and Problem Statement Consider a primary network (PN)

Consisting of a single base station (BS) and M active mobile stations (MSs a.k.a PUs) uniformly distributed within its coverage RBS. Further, a CR ad hoc network comprising of N CR nodes uniformly distributed within a radius R, with BS at its center, coexists with the given PN. Secondary/CR nodes present in the area can lie either

outside or inside the coverage of the primary BS. Each of the M PUs/MSs present is allocated a distinct frequency band for communication with its BS (in uplink) and vice versa, and are assumed to be stationary. In the current setting CR nodes concurrently seek to form single hop links with their intended receiver, by reusing one of the M PU bands subject to constraints on interference caused to PUs. Such a single hop link (CRT x,CRRx), consists of a CR transmitter (CRT x) located at a distance r 0 from the BS such that ($0 < r\ 0 \leq R$) and, a CR receiver (CRRx). We assume that CR nodes are aware of the location of their neighboring CR nodes, which lie within one hop distance from them. Each of these CR peer-to-peer connections uses a distinct PU band, so that interference caused within the CR ad hoc network can be ignored. Let SIRI and SIRA denote the received signal to interference ratios (SIR) for the CR ad hoc and primary network respectively. The essential condition for both a pair of CR nodes and, a primary receiver and transmitter pair to operate successfully on a common frequency band can be expressed as: $(SIRI > SIRI,T) \cap (SIRA > SIRA,T)$ where SIRI and SIRA represent the received SIR at CR receivers and primary receivers respectively. The quantities SIRI,T and SIRA,T denote the minimum SIR requirements for the PN and CR respectively. The minimum SIR requirement for the PN (SIRI,T) dictates the amount of interference that can be tolerated by PUs, whereas SIRI,A represents the minimum rate requirement of CR users. However, in order to estimate Resource Management in CRNs - Single CR User 27 SIRI at the PUs the CR nodes need to be aware of the locations of the PU transmitters and receivers. But embedding location awareness in overlaid CR ad hoc networks increases system complexity, and is justified if the resultant gains (increase in system spectrum utilization) are significant. To observe the impact of incorporating location awareness in CR ad hoc networks, we determine the performance gap between two scenarios. • Location Aware Scenario (LAS) - CR nodes are aware of the location of PU transmitters and receivers. • Location Unaware Scenario (LUS) - CR nodes are aware of the location of the PU transmitters only. Location information concerning PN can be collected via methods discussed in chapter 2. In this paper, the adopted figure of merit is

probability of successful concurrent transmission (PSCT), which represents the probability of coexistence of a PU and a CR link on a common frequency band :

PSCT = P((SIRI > SIRI,T)∩(SIRA > SIRA,T))

Notice that a CR decides to transmit only when it confirms that 4.1 is satisfied. Assuming perfect utilization of temporal spectrum opportunities, the probability of opportunistic transmission (POT) can be expressed as: POT = PSCT .Pr(PUon) +1.Pr(PUo f f) where Pr(PUon) and Pr(PUo f f) represent the probability of the PU being in on/off state, which can be learned statistically over long term via temporal spectrum sensing. To analyze POT , the only term that remains to be assessed is PSCT , which is the focus of this paper. The probability of successful concurrent transmission can therefore be used as a measure of the gain achieved in spectrum utilization, by reusing PU bands. Further we consider two scenarios: Downlink and Uplink for our analysis since both differ in their location estimation or sensing requirements, as discussed later. We shall assume that all transmissions are omni directional and the signal propagation is governed by a log normal shadowing model. The received power at node i (Pi) due to node j can be expressed as:

Pi = sj −g(di, j ,ni, j) +W [dBm]

Resource Management in CRNs - Single CR User 28 where sj represents the transmission power of node j and di, j denotes the distance between nodes i and j.

Further plg(d,n) = 10nlog10(d/do)[dB]

denote it by g(d),

where do is the reference distance.

We assume that shadowing noise W ∼ (0,σ 2 w) where σ 2 w denotes the shadowing variance in dB. The goal of this work is to evaluate 4.2 for both LAS and LUS in order to shed light on advantages of embedding location awareness in CR nodes. Before we proceed, we present an analytical result that will be useful for our ensuing analysis. Result: If X and Y are two normally distributed random such that

X ∼ (μx,σ 2 x) and Y ∼ (μy,σ 2 y). Then, the probability that ratio of X and Y exceeds a constant T (T 0) i.e.,

Pr(X/Y) ≥ T is given by:

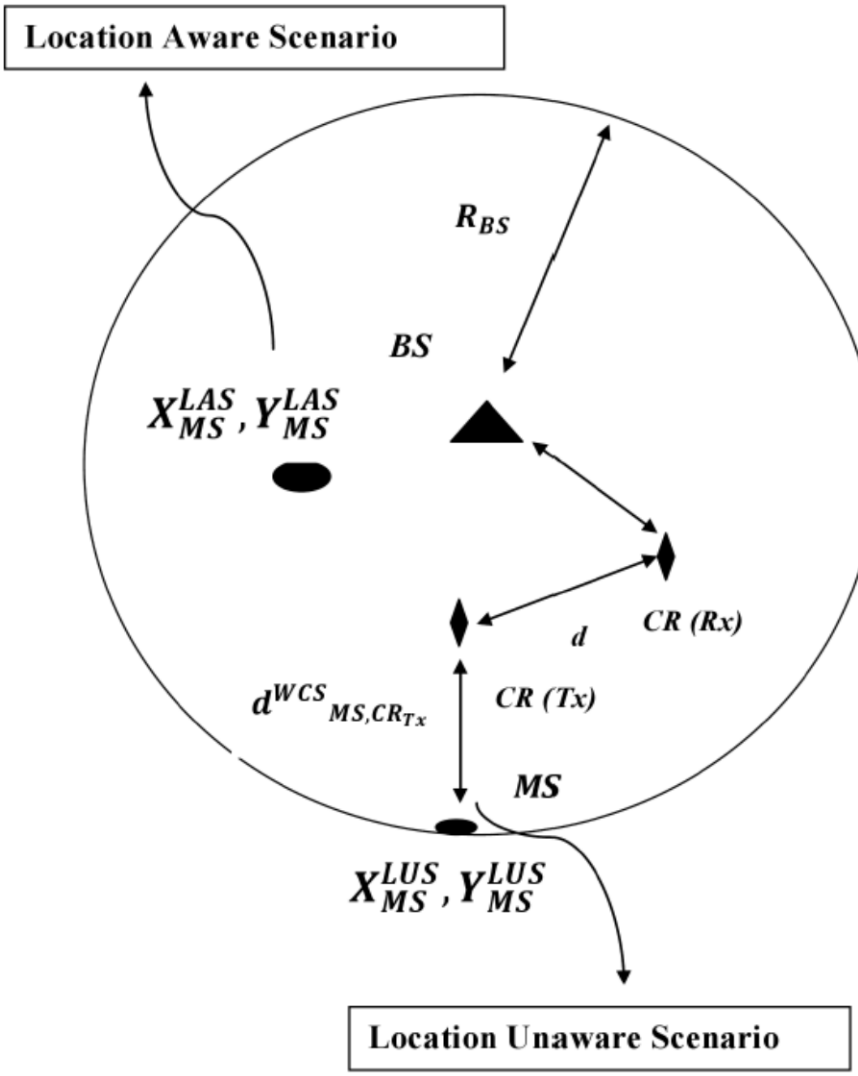
Pr(X/Y) ≥ T = Q((−μz)/σz)

μz = μx +T μy

σz = q (σ2 x +T 2σ2 y) and Q function can be defined as: Q(x) = 1/2(1−er f(x/ √ 2))

(4.5) Downlink Analysis In the downlink, .

A BS tries to establish connections with active MSs within its coverage area, using wireless channels. In the current setup we assume M PUs to be active at any time, each of which is connected to the BS via an orthogonal frequency band. The BS is equipped with ATPC, which enables it to adjust its transmission power in accordance to its distance from the intended MS. Concurrently CR nodes seek to form single hop links with their intended receivers, by reusing one of the M PU bands. However, a (CRT x,CRRx) pair overlays with PU transmission on a common frequency band, if only when this CR link confirms that 4.1 is satisfied. If 4.1 is not satisfied for a particular band, the CR pair continues to scan the remaining M − 1 PU bands until is satisfied for any one of them. If 4.1 is not satisfied for neither of the M PU bands, the CR pair fails to establish a link. Location Aware Scenario (LAS) Under the LAS we assume that the CR nodes are aware of the locations of PU transmitter (BS) and PU receivers (MSs) along with their characteristic frequency bands.


Location Aware Scenario
R_{BS}
BS
$X_{MS}^{LAS}, Y_{MS}^{LAS}$
d
CR (Rx)
$d^{WCS}{}_{MS,CR_{Tx}}$
CR (Tx)
MS
$X_{MS}^{LUS}, Y_{MS}^{LUS}$
Location Unaware Scenario

Unit IV

Cognitive Radio Architecture

Part-A

1) **What are the primary functions of Cognition? (L-1, CO-4)**

Cognition function should:

— maintain a model of space-time

— reliably infer user's communication context & inform SDR

— model propagation of own radio signal (estimate interference)

— infer/adjust the parameters to support running applications

— administer the computational resources

— recognize preemptive actions by user and give control to him/her

2) **What is the objective of cognitive radio architecture? (L-1, CO-4)**

- make flexible use of Radio Spectrum (e.g. SDR)
- see what user sees (e.g. image/video pattern recognition)
- hear what user hears (e.g. voice/speaker recognition)
- protect user's data (e.g. soft/hard biometrics & encryption)
- judge Quality of Information (QoI) needed by user

3) **What is waking behavior? (L-1, CO-4)**

Waking behavior is optimized for real-time interaction with the user, isochronous control of SWR assets, and real-time sensing of the environment. The conduct of the waking behavior is informally referred to as the awake-state, although it is not a specific system state, but a set of behaviors

4) **Define sleeping behavior. (L-2, CO-4)**

Cognitive PDAs (CPDAs) detect conditions that permit or require sleep and dreaming. For example, if the PDA predicts or becomes aware of a long epoch of

low utilization (such as overnight hours), then the CPDA may autonomously initiate sleeping behavior. Sleep occurs during planned inactivity, for example, to recharge batteries. Dreaming behavior employs energy to retrospectively examine experience since the last period of sleep.

5) What is a conflict? (L-2, CO-4)

Conflict is a context in which the user overrode a CPDA decision about which the CPDA had little or no uncertainty. Map _ may resolve the conflict. If not, it will place the conflict on a list of unresolved conflicts (map).

6) Define prayer behavior. (L-1, CO-4)

Attempts to resolve unresolved conflicts via the mediation of the PDA's home network may be called prayer behavior, referring the issue to a completely trusted source with substantially superior capabilities. The unresolved-conflicts list is mapped to RXML queries to the PDA's home CN expressed in XML, OWL, KQML, RKRL, RXML, or a mix of declared knowledge types. Successful resolution maps network responses to integrated knowledge.

7) Define world model. (L-1, CO-4)

The World Model, W, consists primarily of bindings between a priori data structures and the current scene. These associative structures are also associated with the observe phase. Dialog states, action requests, plans, and actions are additional data structures needed for the observe, orient, plan, and act phases, respectively.

8) What is meant by Biniding? (L-1, CO-4)

Binding occurs when there is a nearly exact match between a current stimulus and a prior experience and very general criteria for applying the prior experience to the current situation are met. One such criterion is the number of unmatched features

of the current scene. If only one feature is unmatched and the scene occurs at a high level such as the phrase or dialog level of the inference hierarchy, then binding is the first step in generating a plan for behaving in the given state similar to the last occurrence of the stimuli

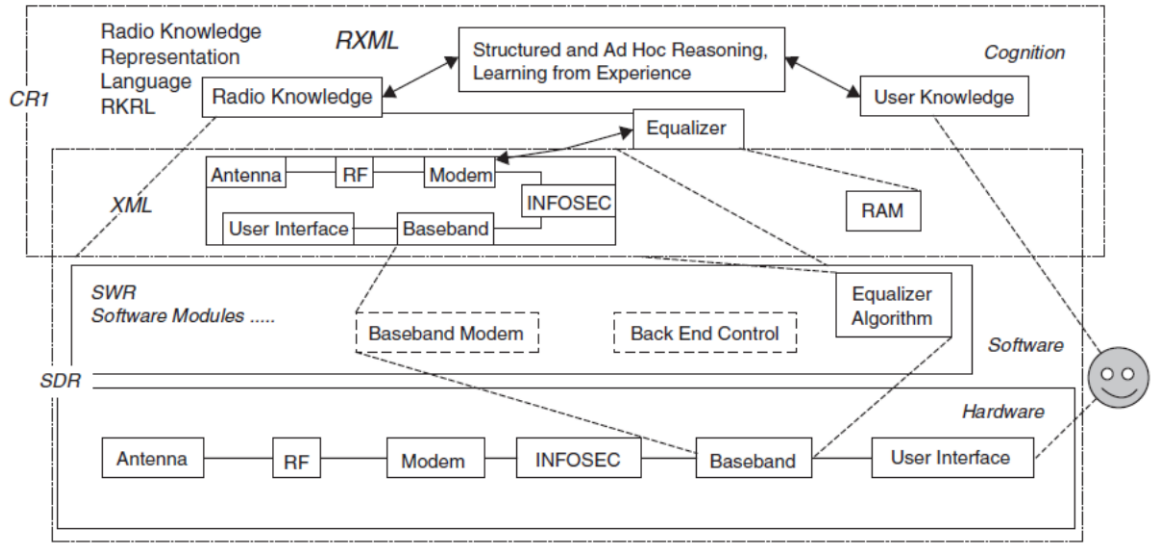
9) Define Scene. (L-1, CO-4)

A scene is a context cluster, a multidimensional space–time–frequency association, such as a discussion of a baseball game in the living room on a Sunday afternoon. Such clusters may be inferred from unsupervised ML (e.g., using statistical methods or nonlinear approaches such as SVMs).

PART-B

1) Discuss about the primary functions of cognitive radio with diagram . (L-1, CO-4)

The SDR components and the related cognitive components of iCR appear in Figure. The cognition components describe the SDR in RXML so that the <Self/> can know that it is a radio and that its goal is to achieve high QoI tailored to its own users. RXML intelligence includes a priori radio background and user stereotypes as well as knowledge of RF and space–time <Scenes/> perceived and experienced. This includes both structured reasoning with iCR peers and cognitive wireless networks (CWNs), and ad hoc reasoning with users, all the while learning from experience



The detailed allocation of functions to components with interfaces among the components requires closer consideration of the SDR component as the foundation of CRA.

SDR Components

SDRs include a hardware platform with RF access and computational resources, plus at least one software-defined personality. The SDR Forum has defined its SCA [3] and the Object Management Group (OMG) has defined its SRA [4]. These are similar fine-grained architecture constructs enabling reduced-cost wireless connectivity with next-generation plug-and-play. These SDR architectures are defined in Unified Modeling Language (UML) object models [5], Common Object Request Broker Architecture (CORBA) Interface Design Language (IDL) [7], and XML descriptions of the UML models. The SDR Forum and OMG

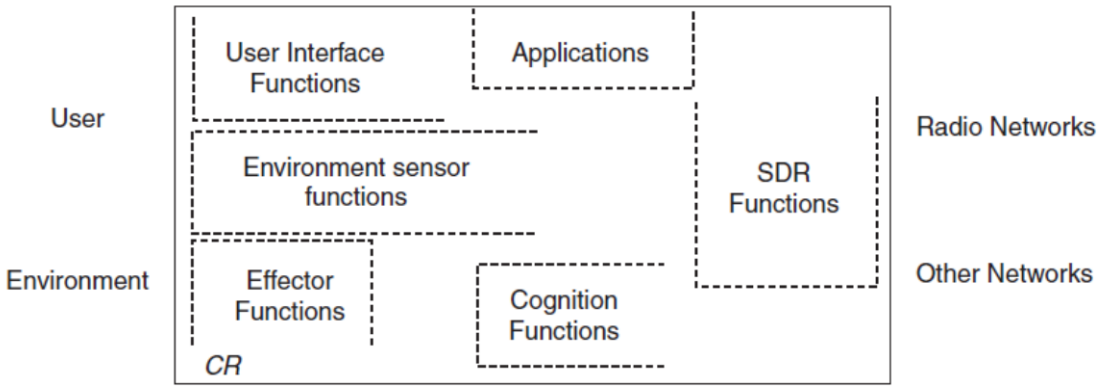
standards describe the technical details of SDR both for radio engineering and for an initial level of wireless air interface (“waveform”) plug-and-play. The SCA/SRA was sketched in 1996 at the first US Department of Defense (DoD) inspired modular multifunctional information transfer system (MMITS) Forum, was developed by the DoD in the 1990s and the architecture is now in use by the US military [7]. This architecture emphasizes plug-and-play wireless personalities on computationally capable mobile nodes where network connectivity is often intermittent at best.

The commercial wireless community [8], in contrast, led by cell phone giants Motorola, Ericsson, and Nokia, envisions a much simpler architecture for mobile wireless devices, consisting of two application programming interfaces (APIs)—one for the service provider and another for the network operator. Those users define a knowledge plane in the future intelligent wireless networks that is not dissimilar from a distributed CWN. That community promotes the business model of the user→service provider→network operator→large manufacturer→device, in which the user buys mobile devices consistent with services from a service provider, and the technical emphasis is on *intelligence in the network*. This perspective no doubt will yield computationally intelligent networks in the near- to mid-term.

The CRA developed in this text, however, envisions the computational intelligence to create ad hoc and flexible networks with the *intelligence in the mobile device*. This technical perspective enables the business model of user→ device→ heterogeneous networks, typical of the Internet model in which the user buys a device (e.g., a wireless laptop) that can connect to the Internet via any available Internet service provider (ISP). The CRA builds on both the SCA/SRA and the commercial API model, but integrates Semantic Web intelligence in RXML for more of an Internet business model. This chapter describes how SDR, AACR, and iCR form a continuum facilitated by RXML.

AACR Node Functional Components

A simple CRA includes the functional components shown in Figure 14.2. A functional component is a black box to which functions have been allocated, but for which implementation is not specified. Thus, while the applications component is likely to be primarily software, the nature of those software components is yet to be determined. User interface functions, however, may include optimized hardware (e.g., for computing video flow vectors in real time to assist scene perception). At the level of abstraction of this figure, the components are functional, not physical.



These functional components are as follows:

1.The *user sensory perception* (SP), which includes haptic, acoustic, and video sensing and perception functions.

2. The local *environment* sensors (location, temperature, accelerometer, compass, etc.).

3. The *system applications* (sys apps) media-independent services such as playing a network game.

4. The *SDR* functions which include RF sensing and SDR applications.

5. The *cognition* functions (symbol grounding for system control, planning, and learning).

6. The *local effector* functions (speech synthesis, text, graphics, and multimedia displays).

These functional components are embodied on an iCR platform, a hardware realization of the six functions. To support the capabilities described in the prior chapters, these components go beyond SDR in critical ways. First, the user interface goes well beyond buttons and displays. The traditional user interface has been partitioned into a substantial user sensory subsystem and a set of local effectors. The user sensory interface includes buttons (the haptic interface) and microphones (the audio interface) to include acoustic sensing that is directional, capable

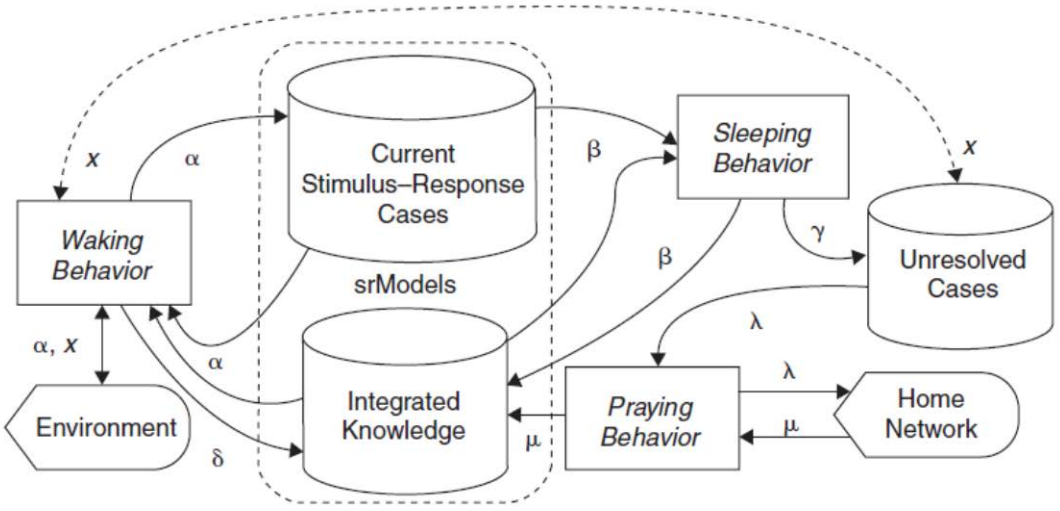
of handling multiple speakers simultaneously, and able to include full motion video with visual scene perception. In addition, the audio subsystem does not just encode audio for

(possible) transmission; it also parses and interprets the audio from designated speakers, such as the <User/>, for a high-performance spoken natural language (NL) interface. Similarly, the text subsystem parses and interprets the language to track the user's information states, detecting plans and potential communications and information needs unobtrusively as the user.

2) **What is behaviour? Explain the various modes of behaviour. (L-3, CO-4)**

Behaviors in the CRA

CRA entails three modes of behavior: waking, sleeping, and praying. Behavior that lasts for a specific time interval is called a behavioral epoch. The axiomaticrelationships among these behaviors are expressed in the topological maps of Figure.



α: action cycle; δ: in cremental machnine learning;
β: nonincremental machine learning; γ: learning conflicts;
λ: RKRL/KQML requests for external assistance;
μ: authoritative assistance; *x*: attempt to resolve problems.

Wak ing Behavior

Waking behavior is optimized for real-time interaction with the user, isochronous control of SWR assets, and real-time sensing of the environment. The conduct of the waking behavior is informally referred to as the awake-state, although it is not a specific system state, but a set of behaviors. Thus, referring to Figure, the awake-state cognition-actions map the environment interactions to the current stimulus–response cases. These cases are the dynamic subset of the embedded serModels. Incremental ML maps these interactions to integrated knowledge, the persistent subset of the serModels.

Sleeping and Dreaming Behaviors

Cognitive PDAs (CPDAs) detect conditions that permit or require sleep and dreaming. For example, if the PDA predicts or becomes aware of a long epoch of low utilization (such as overnight hours), then the CPDA may autonomously initiate sleeping behavior. Sleep occurs during planned inactivity, for example, to recharge batteries. Dreaming behavior employs energy to retrospectively examine experience since the last period of sleep. In the CRA, all sleep includes dreaming. In some situations, the CPDA may request permission to enter sleeping/dreaming behavior from the user (e.g., if predefined limits of aggregate experience are reached). Regular sleeping/dreaming limits the combinatorial explosion of the process of assimilating aggregated experience into the serModels needed for rea ltime behavior during the waking behaviors. During the dreaming epochs, the CPDA processes experiences from the waking behavior using non incremental ML algorithms. These algorithms map current cases and new knowledge into integrated knowledge . A conflict is a context in which the user overrode a CPDA decision about which the CPDA had little or no uncertainty. Map may resolve the conflict. If not, it will place the conflict on a list of unresolved conflicts.

Prayer Behavior

Attempts to resolve unresolved conflicts via the mediation of the PDA’s home network may be called prayer behavior, referring the issue to a completely trusted source with substantially superior capabilities. The unresolved-conflicts list is mapped () to RXML queries to the PDA’s home CN expressed in XML,OWL, KQML, RKRL, RXML, or a mix of declared knowledge types. Successful resolution maps network responses to integrated

knowledge . Many research issues surround the successful download of such knowledge, including the set

of support for referents in the unresolved-conflicts lists and the updating of knowledge in the CPDA needed for full assimilation of the new knowledge or procedural fix to the unresolved conflict. The prayer behavior may not be reducible to finite-resource introspection, and thus may be susceptible to the “partialness” of TC, even though the CPDA and CWN enforce watchdog timers.

3**) With neat architecture, explain the cognitive radio components. (L-1, CO-4)**

There is no unanimous dentition of which requirements a radio must satisfy in order to be considered software denned. Depending on the level of software recon durability, some authors and organizations have established a division between, for example, Software Capable, Software Programmable and Software Denned Radios. For the sake of the simplicity, all of these will from now on be referred to as Software Denned Radios, since, from the security point of view, they mostly share common threats and problems. It is useful to categorize the types of software present in Software Denned Radios, as per the Wireless Innovation Forum's guidelines, since this categorization is widely accepted and commonly referred to in the scienti_c environment. Following that, it is possible to classify the software in SDRs as Radio Operating Environment (ROE): consists of the core framework, the operating system, device drivers, middleware, installer, and any other software fundamental to the operation of the radio platform; Radio Applications (RA): software which controls behavior of the RF function of the radio. This includes any software de_ning the air interface and the modulation and communication protocols, as well as software used to manage or control the radio in a network environment;

Service Provider Applications (SPA): software used to support network and other service providers' support for the user of the radio. It includes voice telephone calls, data delivery, paging, instant messaging service, emergency assistance, and geolocation; User Applications (UA): application software not falling into any of the above categories.

General SDR-related security threats

One of the potential hazards for SDRs lies in the possibility of tampering with their hardware. Since these hazards apply to all wireless systems and are not unique to the new features that SDRs bring, the focus of this section is on the other types of threats: the ones stemming out from the software's recon_gurability. Main threats to recon_gurability come from faulty and buggy software { hence, the deployed schemes need to protect the system from download and usage of improper software. In general, security-enabling mechanisms for SDRs can be divided into hardware-based and software-based ones, each with their own advantages and disadvantages. Hardware-based mechanisms include hardware modules for monitoring the SDR's recon_gurable parameters. However, unlike the SDRs that they are securing, these mechanisms themselves are typically not easily recon_gurable, and updating the security parameters or policies may be problematic and expensive. Software-based mechanisms, in their turn, rely on deploying the tamper-resistance techniques, providing safe and secure authentication, communication security and integrity, as well as safe algorithms for downloading, updating and distributing the software. The potential vulnerability of such schemes is the openness to malicious medications. Chunxiao et al. [9] present a security architecture based on separation of the application environment and the ROE, so that the compromise of one does not abet the other. Furthermore, SDR recon_guration parameters produced by the application environment are varied against security policies before they are executed in theradio environment. So, in cases where the application environment is tampered with and becomes malicious, it cannot infect the radio environment, and thus the RF characteristics can be ensured to be in compliance with the desired policies. For software classi_cation, the authors have used the Wireless Innovation Forum's guidelines, as was described before, where, on top of the ROE, RA, and SPA they dine the User Application Environment (UAE) as the environment (OS) where UA are executed. The authors proceed to de_ne a new separate layer called Secure Radio Middleware (SRM) { a layer implemented below UAE, which includes the most security-critical components, namely RA and ROE. SRM is composed of: Bypass: the component

in charge of non-critical operations; Memory Management Unit: the unit that controls the behavior of the OSM; Virtualized Hardware: the layer where all the radio applications are performed; Security Policy Monitor: the component that tries to decide a normal value or range for the radio parameters and compare them to the ones that the OS passes to Virtualized Hardware, leading to initialization of the appropriate recovery mechanisms in cases of violation.

As the authors themselves note, their implementation has several constraints. Since a desktop PC has been used as a testbed, the implementation does not reect the performance in the potential real-life scenarios, where platforms will typically be far more resource-constrained. Furthermore, their architecture does not incorporate mechanisms for encryption/decryption, information integrity, access control and secure radio software download, which are issues that need to be addressed separately. Brawerman et al. [5] propose a lightweight version of the Secure Socket Layer (SSL) protocol. SSL provides bulk encryption, end point authentication, and data integrity protection. For encryption, symmetric key algorithms are used, whereas for authentication, client and server can mutually authenticate each other. Light SSL redesigns the SSL protocol in order to decrease the computational complexity of the performed operations and to perform most of the cryptography at the server side, thus making it suitable for power-constrained devices such as SDR terminals. The authors have de_ned several possible attacks, and the corresponding defense features employed within the protocol, namely: Access control: countered by the authentication mechanism; Masquerade attack: attacker emulates the manufacturer server or a client, countered by the use of mutual authentication; Con_dentiality: secrecy of information is ensured by establishing secure connections; Replay: attacker re-transmits messages after a certain time period, countered by using timestamps;

Attack type	Contribution	Attacker's special characteristics	Proposed defense scheme
PUEA: emulating characteristics of a primary user to acquire exclusive spectrum rights	[7]	Altering its transmission power, modulation mode and frequency; injecting false data to the localization system	3-step mechanism: verification of signal characteristics, RSS measurement, localization of the signal
	[8]	Applying the estimation techniques to enhance its performance	Assumes that emulating the channel features is not feasible for the attacker. Invariants of communication channels are used as means of differentiating between the PUE attackers from legitimate PUs
	[24]	-	Novel physical layer authentication mechanism, which incorporates cryptographic and wireless link signatures
	[11] (proposed)	Ability to emulate any of the PU's transmission characteristics	Location integrity checking as means of deciding on the credibility of a user
Byzantine: providing wrong data to other nodes in collaborative spectrum sensing	[32]	Two operating modes: causing False Alarm attack, or causing False Alarm & Misdetection	Each user is attributed a suspicious level, turned into a trust value, but also a consistency value
	[25]	Two types of attacks: false-positive and false-negative. The attackers are assumed to be able to estimate the channel occupancy with 100% precision	Double-defense mechanism: the correlations between the reported RSS values using correlation filters are observed and the suspicious nodes are outlined; weight-combining data fusion rule is used
	[27]	Hit-and-run attacker: able to estimate its current suspicious level and adapt its attacking scheme	Novel reputation algorithm - the user is permanently excommunicated once his reputation value is below a threshold
OFA: disrupting CR's learning mechanism	[28]	-	Set of general guidelines, e.g., Multi-Objective Programming module verifies all the reconfigured parameters in each iteration
Lion attack: multi-layer attack with the goal of causing DoS at the transport layer	[13]	-	Set of general guidelines for reducing the efficiency of the attack
Attacks on CCC	[34]	two types of attacks: DoS attack in multi-hop networks, and the greedy MAC layer behavior	-
	[29]	-	Authentication of communicating Cognitive Radionodes as the key security feature
Spectrum trading security issues	[35]	Attacker decreases the QoS while declaring that it remains the same	Once it observes illegal behavior, PU decreases the amount of spectrum shared with SU, thus reducing its overall utility
802.22-specific	[2]	Identification of the possible attacks: DoS; Replay; Jamming in QPs; PUEA; Threats to WMBs; Attacks on Self-Coexistence mechanism	Security sublayer deals with some of the vulnerabilities, mainly through: Privacy Key Management v2; message authentication codes; Advanced Encryption Standard

Currently, there are two dominant open-source architectures for SDRs: GNU Radio, which is particularly appealing to academic community due to the relative simplicity of use and compatibility with low-cost off-the-shelf SDR platforms such as Universal Software Radio Peripheral (USRP), and Software Communications Architecture, which is the architecture adopted by the Wireless Innovation Forum. GNU Radio is an open-source software toolkit that, coupled with hardware equipment such as USRP, allows for a complete platform for building Software Defined Radios. GNU Radio can also be used as a stand-alone simulation environment. Most of GNU Radio's applications are written in Python, whereas C++ is used for implementing signal processing blocks. Python commands are used to control all of the USRP's software defined parameters, such as transmission power, gain, frequency, antenna selection, etc. GNU Radio is built on two main structural entities: signal processing blocks and flow graphs. Blocks are structured to have a certain number of input and output ports, consisting of small signal-processing components. When the blocks are appropriately connected, a flow graph is made. Hill et al. [14] have analyzed threats related to GNU Radio-based SDR systems. By considering the GNU Radio Software Applications, written in C++, as the Radio Applications (RA), and the Python functions as the Radio Operating Environment (ROE), the authors identify the following shortcomings related to the ROE of GNU Radio: • At the moment, there is no embedded functionality for verification, i.e., securing the SDR device from being reconfigured by a malicious code; • There are risks related to the execution of models in the graph. Since a single address space is shared among all the software modules, there is a possibility for the malicious user to alter the data in the whole address space. To counter this, the authors propose restricting each module to only be able to access its dedicated address space; 37 • There is the possibility of a buffer overflow, stemming from the use of the shared buffer. Mechanisms for restricting the amount of data that can be written to the buffer are needed. They also define three possible attacks, depending on the parameter targeted: • Modulation attack: improper change of the modulation format; • Frequency attack: jamming attack where an impostor is transmitting on the frequencies that it is not allowed to; • Output power attack: where an attacker can continuously

transmit at high power, forcing other users to increase their power level, which leads to increased battery drain. The authors go on to suggest that GNU Radio ROE has to provide mechanisms for evaluating and enforcing policies for specifying the operating constraints of the SDRs, defined by the network administrators and regulators. Software Communications Architecture (SCA) was originally defined by the United States government with the purpose of securing waveform portability and improving software reuse. Built originally for the United States military's Joint Tactical Radio System (JTRS) program, it has been accepted as a communication standard in military services of many other countries, as well as by the commercial organizations such as Wireless Innovation Forum. It is an always-evolving standard, with first version dating back to 2000, that provides standardized set of methods for installing, managing, and uninstalling new waveforms, therefore maintaining interoperability between various SDR systems. Security is a very important aspect of radios featuring SCA. The architecture provides the foundation to solve issues such as programmable cryptographic capability, certificate management, user identification and authentication, key management, and multiple independent levels of classification. Manufacturers and users are embracing the approach, albeit at a relatively slow rate. For example, the Security Supplement to the JTRS SCA [20] requires that the SDR devices "shall only accept cryptographic algorithms/algorithm packages signed by National Security Agency (NSA)", that "NSA shall digitally sign all Security Policy XML files", and that "the operating system invocation method shall be a NSA digitally signed script". However, SDR middleware and tools vendors supporting JTRS customers do not yet support digital signature features within their products, although they generally express openness to including such features in future releases. Similarly, user and manufacturer representatives 38 in the Wireless Innovation Forum's Public Safety Special Interest Group are trying to identify alternatives to digital signatures before committing to such an approach, largely due to perceptions regarding the complexity of the Public Key Infrastructure (PKI) technology

4) **Explain the various components of cognitive Cycle architecture. (L-1, CO-4)**

The Cognition Cycle

The CRA comprises a set of design rules by which the cognitive level of information services may be achieved by a specified set of components in a way that supports the cost-effective evolution of increasingly capable implementations over time [1]. The cognition subsystem of the architecture includes an inference hierarchy and the temporal organization and flow of inferences and control states—the cognition cycle.

The Cognition Cycle

The cognition cycle developed for CR1 [14] is illustrated in Figure. This cycle implements the capabilities required of iCR in a reactive sequence. Stimuli enter the CR as sensory interrupts, dispatched to the cognition cycle for a response. Such an iCR continually observes (senses and perceives) the environment, orients itself, creates plans, decides, and then acts. In a single-processor inference system, the CR's flow of control may also move in the cycle from observation to action. In a multiprocessor system, temporal structures of sensing, preprocessing, reasoning, and acting may be parallel and complex. Special features

synchronize the inferences of each phase. The tutorial code all works on a single processor in a rigid inference sequence defined in Figure . This process is called the "wake epoch" because the primary reasoning activities during this large epoch of time are reactive to the environment. We will refer to "sleep epochs" for power-down conditions, "dream epochs" for performing computationally intensive pattern recognition and learning, and "prayer epochs" for interacting with a higher authority such as network infrastructure.

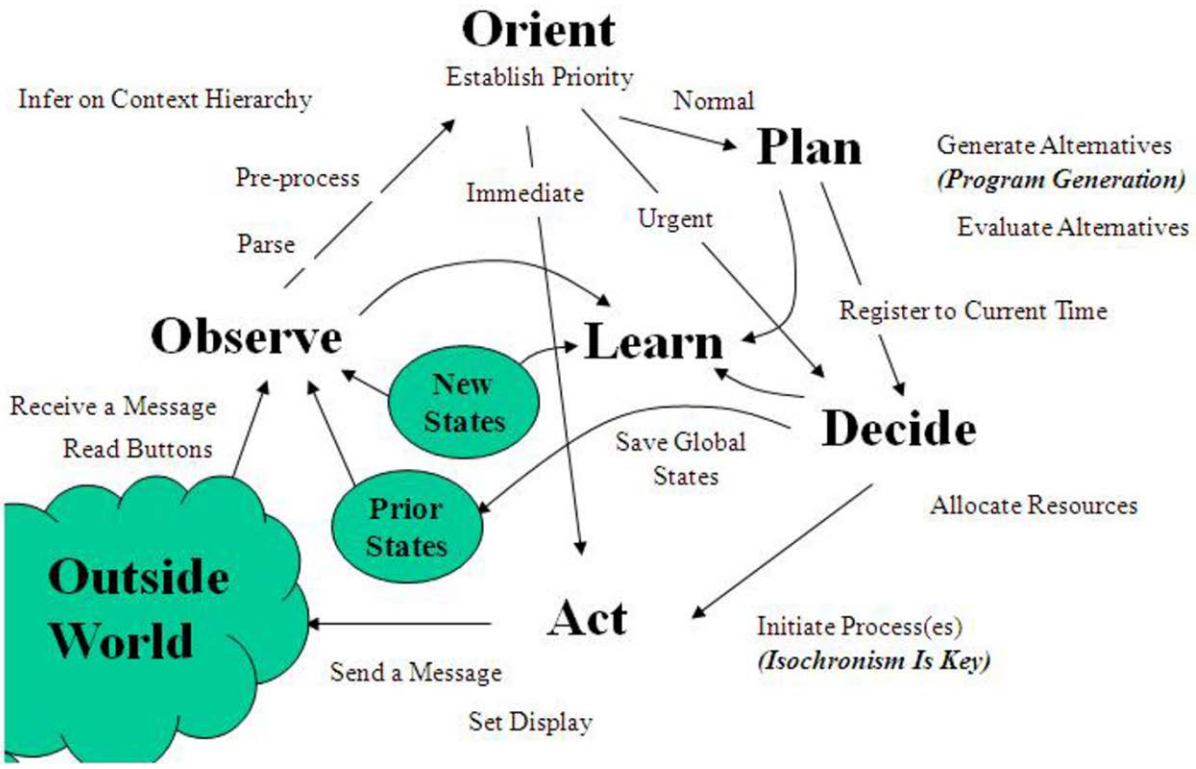


Figure : Simplified cognition cycle. The observe, orient, plan, decide, act (OOPDA) loop is a primary cycle; learning, planning, and sensing the outside world are crucial phases of the larger OOPDA-loop

During the wake epoch, the receipt of a new stimulus on any of a CR's sensors or the completion of a prior cognition cycle initiates a new primary cognition cycle. The CR observes its environment by parsing incoming information streams. These can include monitoring and speech-to-text conversion of radio broadcasts (e.g., the Weather Channel, stock ticker tapes, etc.). Any RF-LAN or other shortrange wireless broadcasts that provide services awareness information may be also parsed. In the observation phase, a CR also reads location, temperature, and lightlevel sensors, among other parameters, to infer the user's communications context.

Observe (Sense and Perceive)

The iCR senses and perceives the environment (via “observation phase” code) by accepting multiple stimuli in many dimensions simultaneously and by binding these stimuli—all together or more typically in subsets—to prior experience so that it can subsequently detect time-sensitive stimuli and ultimately generate plans for action.

Thus, iCR continuously aggregates experience and compares prior aggregates to the current situation. A CR may aggregate experience by remembering everything. This may not seem like a very smart thing to do until you calculate that all the audio, unique images, and e-mails the radio might experience in a year takes up only a few hundred gigabytes of memory, depending on image detail. So the computational architecture for remembering and rapidly correlating current experience against everything known previously is a core capability of the CRA. A *novelty* detector identifies new stimuli, using the new aspects of partially familiar

stimuli to identify incremental-learning primitives. In the six-component (user SP, environment, effectors, SDR, sys apps, and cognition) functional view of the architecture defined in section *AACR Node Functional Components*, the observe phase comprises both the user SP and the environment (RF and physical) sensor subsystems. The subsequent orient phase is part of the cognition component in this model of architecture.

Orient

The orient phase determines the significance of an observation by binding the observation to a previously known set of stimuli of a “scene.” The orient phase contains the internal data structures that constitute the equivalent of the short-term memory (STM) that people use to engage in a dialog without necessarily remembering everything with the same degree of long-term memory (LTM). Typically people need repetition to retain information over the long term. The natural environment supplies the information redundancy needed to instigate transfer from STM to LTM. In the CRA, the transfer from STM to LTM is mediated by the sleep cycle in which the contents of STM since the last sleep cycle are analyzed both internally and with respect to existing LTM.

Matching of current stimuli to stored experience may be achieved by “stimulus recognition” or by “binding.” The orient phase is the first collection of activity in the cognition component.

Stimulus Recognition

Stimulus recognition occurs when there is an exact match between a current stimulus and a prior experience. The CR1 prototype is continually recognizing exact matches and recording the number of exact matches that occurred along with the time measured in the number of cognition cycles between the last exact match. By default, the response to a given stimulus is to merely repeat that stimulus to the next layer up the inference hierarchy for aggregation of the raw stimuli. But if the system has been trained to respond to a location, a word, an RF condition, a signal on the power bus, or some other parameter, it may either react immediately or plan a task in reaction to the detected stimulus. If that reaction were in error, then it may be trained to ignore the stimulus, given the larger context, which consists of all the stimuli and relevant internal states, including time. Sometimes, the orient phase causes an action to be initiated immediately as a “reactive” stimulus–response behavior. A power failure, for example, might directly invoke an act that saves the data (the “immediate” path to the act phase in Figure). A nonrecoverable loss of signal on a network might invoke reallocation of resources (e.g., from parsing input to searching for alternative RF channels.

This may be accomplished via the path labeled “urgent” in Figure.

Binding

Binding occurs when there is a nearly exact match between a current stimulus and a prior experience and very general criteria for applying the prior experience to the current situation are met. One such criterion is the number of unmatched features of the current scene. If only one feature is unmatched and the scene occurs at a high level such as the phrase or dialog level of the inference hierarchy, then binding is the first step in generating a plan for behaving in the given state similar to the last occurrence of the stimuli. In addition to number of features that match exactly, which is a kind of hamming

code, instance-based learning (IBL) supports inexact matching and binding. Binding also determines the priority associated

with the stimuli. Better binding yields higher priority for autonomous learning, whereas less-effective binding yields lower priority for the incipient plan.

Plan

Most stimuli are dealt with “deliberatively” rather than “reactively.” An incoming network message would normally be dealt with by generating a plan (in the plan phase, the “normal” path). Such planning includes plan generation. In research quality or industrial-strength CRs, formal models of causality must be embedded into planning tools. The plan phase should also include reasoning about time. Typically, reactive responses are preprogrammed or defined by a network (i.e., the CR is “told” what to do), whereas other behaviors might be planned. A stimulus may be associated with a simple plan as a function of planning parameters with a simple planning system. Open source planning tools enable the embedding of planning subsystems into the CRA, enhancing the plan component. Such tools enable the synthesis of RF and information access behaviors in a goal-oriented way based on perceptions from the visual, audio, text, and RF domains as well as RA rules and previously learned user preferences.

Decide

The decide phase selects among the candidate plans. The radio might have the choice to alert the user to an incoming message (e.g., behave like a pager) or to defer the interruption until later (e.g., behave like a secretary who is screening calls during an important meeting).

Act

Acting initiates the selected processes using effector modules. Effectors may access the external world or the CR’s internal states.

Externally Oriented Actions

Access to the external world consists primarily of composing messages to be spoken into the local environment or expressed in text form locally or to another CR or CN using the Knowledge Query and Manipulation Language (KQML), Radio Knowledge

Representation Language (RKRL), Web Ontology Language (OWL), Radio eXtensible Markup Language (RXML), or some other appropriate knowledge interchange standard.

Internally Oriented Actions

Actions on internal states include controlling machine-controllable resources such as radio channels. The CR can also affect the contents of existing internal models, such as adding a model of stimulus–experience–response (serModel) to an existing internal model structure [12]. The new concept itself may assert-related concepts into the scene. Multiple independent sources of the same concept in a scene reinforce that concept for that scene. These models may be asserted by the <Self/> to encapsulate experience. The experience may be reactively integrated into RXML knowledge structures as well, provided the reactive response encodes
them properly.

Learning

Learning is a function of perception, observations, decisions, and actions. Initial learning is mediated by the observe phase perception hierarchy in which all SP are continuously matched against all prior stimuli to continually count occurrences and to remember time since the last occurrence of the stimuli from primitives to aggregates.

Learning also occurs through the introduction of new internal models in response to existing models and case-based reasoning (CBR) bindings. In general, there are many opportunities to integrate ML into AACR. Each of the phases of the cognition cycle offers multiple opportunities for discovery processes, such as <Histogram/>, as well as many other ML approaches. The architecture includes internal reinforcement via counting occurrences and via serModels, so ML with uncertainty is also supported [12, 31].

Finally, a learning mechanism occurs when a new type of ser Model is created in response to an action to instantiate an internally generated ser Model. For example, prior and current internal states may be compared with expectations to learn about the effectiveness of a communications mode, instantiating a new mode specific ser Model.

Self-monitoring

Each of the prior phases must consist of computational structures for which the execution time may be computed in advance. In addition, each phase must restrict its computations to not consume more resources (time _ allocated processing capacity) than the precomputed upper bound. Therefore, the architecture has some prohibitions and some data set requirements needed to obtain an acceptable degree of stability of behavior for CRs as self-referential self-modifying systems. Since first-order predicate calculus (FOPC) used in some reasoning systems is not decidable, one cannot in general compute in advance how much time an FOPC expression will take to run to completion. There may be loops that will preclude this, and even with loop detection, the time to resolve an expression may be only loosely approximated as an exponential function of some parameters (such as the number of statements in the FOPC database of assertions and rules). Therefore, unrestricted FOPC is not allowed.

Similarly, unrestricted For, Until, and While loops are prohibited. In place of such loops are bounded iterations in which the time required for the loop to execute is computed or supplied independent of the computations that determine the iteration control of the loop. This seemingly unnatural act can be facilitated by next-generation compilers and computer-aided software engineering (CASE) tools. Because self-referential self-modifying code is prohibited by structured design and programming practices, no such tools are available on the market today. But CR is inherently self-referential and self-modifying, such tools most likely will emerge, perhaps assisted by the needs of CR and the architecture framework of the cognition cycle.

Finally, the cognition cycle itself cannot contain internal loops. Each iteration of the cycle must take a defined amount of time, just as each frame of a 3G air interface takes 10 milliseconds. As CR computational platforms continue to progress, the amount of computational work done within the cycle will increase, but under no conditions should explicit or implicit loops be introduced into the cognition cycle that would extend it beyond a given cycle time.

Retrospection

The assimilation of knowledge by ML can be computationally intensive, so, as previously stated in Section 14.3.1 and further discussed in Section 14.5.4, CR has sleep and prayer epochs that support ML. A sleep epoch is a relatively long period of time (e.g., minutes to hours) during which the radio will not be in use, but has sufficient electrical power for processing. During the sleep epoch, the radio can run ML algorithms without detracting from its ability to support its user's needs. ML algorithms may integrate experience by aggregating statistical parameters. The sleep epoch may re-run stimulus–response sequences with new learning parameters in the way that people dream. The sleep cycle could be less anthropomorphic, however, employing a genetic algorithm to explore a rugged fitness landscape,potentially improving the decision parameters from recent experience.

Reaching Out

Learning opportunities not resolved in the sleep epoch can be brought to the attention of the user, the host network, or a designer. We refer to elevating complex problems to an infrastructure support as a prayer epoch.

5) **Explain natural language encapsulation and Radio Procedure Knowledge Encapsulation (L-1, CO-4)**

The advent of software defined radio (SDR) technology offers a more sophisticated form of processing resources than prior radio technology. Although the initial development of SDR technology was almost exclusively for military applications, as the field has matured its scope has broadened to include commercially-oriented perspectives (e.g., the SDR Forum and now both the standards for designing SDRs (e.g., the Software Communications Architecture and representative, open source implementations reflect industry-standard, objectoriented software practices. With this enhanced capability, however, comes the burden of developing the software and selecting configurations applicable for the various scenarios the SDR may encounter. One technology that promises to not only utilize this processing capability but to also provide an autonomous and flexible architecture that is applicable to a wide array of operational scenarios is the cognitive radio (CR) . The ultimate vision of CR technology—denoted by Mitola as the "ideal cognitive radio (iCR)"—encompasses many facets of intelligent behavior such as context awareness, adaptation of action due to stimulus and prior information, reasoning including inferring information not explicitly stated, learning, natural language processing, and planning. A growing research community is investigating the means for taking advantage of the processing resources in SDR platforms to develop the iCR; to date, most researchers choose to focus on one or a few of these facets of intelligence. As a result, literature on the subject defines the term CR in a variety ways, usually in a narrow, application-specific manner. As an example CR application, the Federal Communications Commission (FCC) is considering a case where vacant portions of the TV broadcast bands could be shared with unlicensed devices with sufficient intelligence to detect the licensed users and avoid causing harmful interference to those users a related effort by the IEEE 802.22 standards committee seeks to create a technical standard for a network of these devices . While this definition of CR does include a radio with some awareness of the spectrum and some ability to adapt operating behavior based upon that information, this definition otherwise duplicates conventional radio technology with procedural-style

specification of the radio's behavior. The DARPA XG program aims to demonstrate opportunistic spectrum access of otherwise idle spectrum under a range of conditions. An important component of that application is a policy checking entity that determines whether or not the dynamic spectrum access adheres to a policy. Their current approach employs a Prolog-based policy reasoner to evaluate such queries. Other researchers such as Berlmann et al. proposed policybased reasoning to check a broader range of CR behaviors. Work by Neel et al. applied game theory principles to design distributed algorithms for adaptive behaviors. Rondeau et al.proposed genetic algorithms for optimizing the settings of the many control parameters available to CRs. Both addressed the problem of adaptation in CRs, and could also be viewed as addressing a learning component. The research group of Kokar investigated how to create CRs with self awareness of their own capabilities via an ontology framework and how to replace procedural- 2 of 7 style radio control constructs with machine reasoning techniques. The use of an ontology as a knowledge representation mechanism is central to this paper's approach too; references provide tutorial information on the topic of ontologies. On a related note, some of the authors of this paper took an ontology-based approach in providing CRs with context awareness; for example, the term radio channel has many possible meanings, and in order to reason about the availability of a radio channel the CR must know what definition applies in a given context . Recognizing the possible applications and approaches for introducing machine reasoning to radio systems as discussed in this section, this paper explores a CR incorporating a differential-response capability by augmenting the existing SDR processing paradigm using knowledge representation concepts such as ontologies and rules. Sections that follow note the importance of a knowledge-driven differential-response capability not found in prior work and describe components necessary to instantiate it. Finally, the paper closes by describing a simulated prototype CR with this capability and how it can achieve goals despite facing conflicts that would have thwarted conventional radios. REQUIREMENTS FOR KNOWLEDGE-DRIVEN DIFFERENTIAL-RESPONSE As a motivating example, this paper considers the problem of a CR attempting to gain access to a portion of the radio

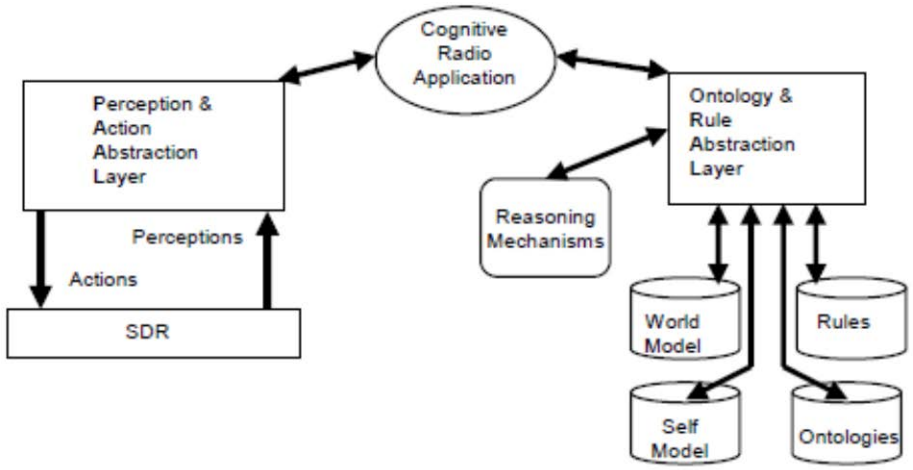
band governed by radio beacons at one or more locations. A number of beacon-based protocols are possible to facilitate dynamic spectrum access; this example assumes a policy regime in which both positive and negative control beacons are employed [18]. In order for the CR to be able to access a radio channel C, two conditions must be fulfilled: 1) the CR must be within radio range of a positive beacon station from which it receives a coded beacon message authorizing access to channel C, and 2) the CR must not simultaneously be in range of a negative beacon station from which it receives a coded beacon message denying access to C. Although there are a number of security, protocol, and radio engineering issues that must be addressed in the design and implementation of such beacons, it is not necessary to describe those aspects of the beacon in order to appreciate the value of differential-response capability. For both conventional radios and the CRs described in the previous section, if the radio's location is such that conditions (1) and (2) are not satisfied, it cannot access the channel, and—more fundamentally—it cannot reason about why the goal of channel access has failed or what alternative conditions would permit overcoming the failure. Note that there are essentially two ways in which the goal of using channel C can be thwarted: a) no positive beacon signal for access to C is received, and b) both positive and negative beacon signals for access to C are received. The cases in which no beacon signals are detected at all, or in which only negative beacon signals are received can be viewed as being subsumed under the other cases. For example, in the case where only negative signals are received, the radio needs to somehow get a positive signal, which is case (a). If it manages to solve that problem and it still is receiving one or more negative signals, then it is now in case (b) (otherwise it has solved the problem of gaining access). The situations covered by cases (a) and (b) can be used to elaborate upon the notion of knowledge-driven differential-response. In both cases a radio will fail in its goal of getting access to channel C. Depending upon its functionality, a conventional radio might be able to distinguish between the two cases of failure, in the sense that it goes into a different internal state depending upon the circumstances. Through a user-interface it may be able to give an indication of its current state. However, even if the concrete indicators conveyed by the

radio are different in the two cases, this is still not a knowledge-driven differentialresponse. First of all, one can say, with some justice, that in both cases the radio is really doing exactly the same thing: upon failure to access a channel convey the current internal state to the user. That is an accurate description of the radio's actions, because that is exactly how the radio is programmed to behave. Secondly, even if, for the sake of argument, the responses are deemed to be different, they do not illustrate knowledge-driven differential-response. The reason is that the radio being in a distinct internal state is an irreducible and non-analyzable cause of its taking whatever action it takes. From a formal point of view, we may say that the radio behaves as a finite-state machine. The particulars of the internal state and the way in which those particulars relate to aspects of the external world do not enter into an explanation of the radio's behavior. The latter is at least part of what is required in order for any agent to be capable of cognition, and this is what we mean by knowledge-driven differential-response. So, returning to the example scenario, what could a CR do in case (b) as opposed to case (a)? A CR would have representations of policy conditions (1) and (2) and it could also have a representation of a beacon signal conflict, that is, a situation in which two or more conflicting beacon signals are received. It would also have the knowledge, expressed in a rule, that when a goal cannot be achieved due to such a conflict one can request a move to a location where only the desired beacon is in range. Depending on the radio's mobility capabilities, it could then act in a number of ways depending on the circumstances. For example, using its inherent signal strength detection capability it could guide its user to a region where only the desired beacon signal is received. 3 of 7 Furthermore, a general set of requirements can be extrapolated from this example: 1. Knowledge and Reasoning Requirements Knowledge requirements for CR are of two sorts: conceptual and rule-based. Conceptual knowledge includes knowing the meanings of fundamental notions in a domain of interest as well as fundamental principles relating those concepts. In current practice, this kind of knowledge is said to be ontological and is formally encoded in knowledge representation frameworks known as ontology languages. An ontology is a formal representation of the key concepts and principles of a domain of interest.

Rulebased knowledge, which is typically represented using some formal rule language, can be thought of as the bridge that relates conceptual knowledge to the problem-solving needs of a particular application. Knowledge and reasoning go hand-in-hand. A piece of knowledge that is not somehow related to other pieces of knowledge through inference (i.e., reasoning) is essentially useless. Since the knowledge and reasoning requirements, including rules, are fundamental to the CR, the topic is addressed in a separate section ("Role of Knowledge Representation") of the paper. 2. Perceptual Requirements It is clear from the example scenario that a CR must be able to recognize sensory inputs, or patterns thereof, as being or representing something in its environment The concept of self-perception is in this category too; broadly speaking, a CR must be able to recognize certain internal states and processes as representing certain facts about itself. For example, a CR should be able to perceive its current rate of power consumption as a property belonging to itself just as a human perceives a bodily sensation such as pain as something that is internal. 3. Action Requirements A CR needs to be able to initiate action based upon the conclusions it reaches. For example, if a CR decides that it should attempt to communicate with another trusted host on behalf of its user during some emergency, then it must not only have knowledge of the communication protocol but also have the ability to execute the protocol. The knowledge satisfying the above requirements also forms the basis of a world model and a self model, both of which are created and maintained within each CR. ARCHITECTURE The discussion thus far can be encapsulated in a proposed architecture for a CR device, as shown in fig. 1. In the left side of the figure we see that the goal of augmenting SDR processing structures is accomplished in this architecture by means of the Perception & Action Abstraction Layer (PAAL). The Perception and Action Abstraction Layer (PAAL) is defined in terms of certain standard radio concepts and is used to characterize device observables and actions in a platform-independent knowledge representation. Figure 1. CR Architecture. This is a key layer if one wants to allow for reuse of the cognitive portion of the architecture with different conventional radio implementations. That is, different radios could use different signal processing algorithms at a very low level that have no bearing

on how a cognitive radio application perceives an instance of, for example, a certain kind of waveform. The PAAL makes it possible for a device to interpret its sensory input in perceptual terms that can be used to drive a CR's world and self models. Going in the other direction, it also makes it possible for a CR to do things by exporting SDR primitive actions in a platform independent format. The right side of the figure shows the components involved in augmenting an SDR architecture to allow for cognitive capabilities including the previously mentioned World Model and Self Model components. The next section decribes the Rules, Ontologies, and Reasoning Mechanisms components. The remaining component is another abstraction layer, the Ontology & Rule Abstraction Layer. This layer serves a purpose that is symmetric to PAAL. It allows ontology and rule concepts to be represented in a platform-independent standard. This is important if one wants to allow the same radio implementation to be used with alternate ontology and rule reasoning platforms. Just as radio notions such as signal and waveform should have meaning independent of any particular radio implementation, so too notions such as concept, and rule should have meaning independent of any particular implementations. As an overview of how the proposed architecture works to augment existing SDR implementations, one can consider the case where the radio senses some waveform. We asSDR Perceptions Actions Perception & Action Abstraction Layer Cognitive Radio Application World Rules Model Ontologies Self Model Ontology & Rule Abstraction Layer Reasoning Mechanisms 4 of 7 sume that the radio's SDR interface can be used to program a wrapper around its existing methods so that such an event triggers a method defined in terms of PAAL that allows an appropriate instance of a signal object (as defined in the ontology) to be constructed and deposited into the world model. Conversely, suppose the reasoner concludes that a certain action, such as evacuating a channel, should be taken. From its self model it knows that it is capable of taking such an action. Then, by virtue of the PAAL layer, the ontological element that represents that action will be linked to a method that can invoke the radio's native interface with a call to perform that action (or perform some procedure). ROLE OF KNOWLEDGE REPRESENTATION This section elaborates on the types and forms of knowledge that a

CR must have in order to exhibit a knowledgedriven differential-response. Conceptual knowledge is the kind of knowledge that ontologies are intended to represent. Conceptual knowledge is typically either analytic or axiomatic in nature. These types of knowledge are both thought of as representing necessary truths, but for different reasons. A piece of knowledge is analytic if it expresses or follows from the meaning of concepts. For example, it is useful to talk about radios that can be moved from place to place (without impairing their operational capabilities). The concept of a mobile radio would therefore be defined as a radio that has this property. Representing this definition in an ontology would make it possible for it to be applied in a formal reasoning system. Axiomatic conceptual knowledge, on the other hand, expresses fundamental conceptual relationships that are not based on meaning alone. For example, the fundamental principles that radio waves are a form of electromagnetic energy and that they travel at the speed of light might be considered axioms within an ontology of radio knowledge. What might be considered an axiom from the point of view of one ontological domain, however, might be considered a derived piece of knowledge (e.g., a theorem) from the point of view of a more fundamental domain. This relativity of what is an axiom has been demonstrated many times in the history of science. Kepler's laws of planetary motion, for example, had the status of independent axioms when initially formulated, but were later shown to be consequences of Newton's general laws of motion. Rules are also important; they may be thought of as theorems that are worth committing to memory, so to speak, because, 1) they are useful in an application of interest and, 2) the computational cost of deriving them from axioms on demand is prohibitive. For example, it is known that certain frequencies of radio signals are likely to degrade because of atmospheric conditions, and mathematical laws governing this phenomenon can be derived from first principles. However, for any application in which this kind of knowledge is critical, it is highly likely that even a human expert would depend upon known rules for calculating such attenuation rather than performing an analysis based on the



reasoning process to derive. Therefore, statements that relate the existence of a policy in a region to actions that need to be taken (or avoided) in order to be in compliance with that policy might often be worth committing to memory in the form of rules. Ontologies also enable reasoning. From a theoretical point of view the kind of reasoning afforded by ontologies differs from rule-based reasoning. Subsumption reasoning is one example. Thus, as discussed above, from the fact that R is a radio and has the property of being mobile, and the definition of mobile radio, one can infer that R is a mobile radio. In practice, subsumption reasoning can be implemented using an underlying rule-based approach, but that is not necessary. SCENARIO USING PROTOTYPE SIMULATION We have implemented a prototype simulation environment capable of handling the beacon signal conflict scenario we outlined above. The ontological knowledge is expressed in OWL [20]. We use Jena [19] as our ontology API and we also use the rule language provided with Jena for representing rules. The inference mechanisms are also Jenabased. The PAAL is implemented by linking the ontology API with our own interface to a simple software defined radio emulation in Java. The simulation enables one or more CRs and one or more beacons to be represented in a two dimensional space. The CRs can be mobile. Currently this means that they are

associated with a user who can move around in the simula- 5 of 7 tion environment. As a radio is moved and as the various components of the environment change,

6) **Describe the design rules in detail for cognitive radio. (L-3, CO-4)**

Depending on transmission and reception parameters, there are two main types of cognitive radio:

- *Full Cognitive Radio* (Mitola radio), in which every possible parameter observable by a wireless node (or network) is considered
- *Spectrum-Sensing Cognitive Radio*, in which only the radio-frequency spectrum is considered.

Other types are dependent on parts of the spectrum available for cognitive radio:

- *Licensed-Band Cognitive Radio*, capable of using bands assigned to licensed users (except for unlicensed bands, such as the U-NII band or the ISM band. The IEEE 802.22working group is developing a standard for wireless regional area network (WRAN), which will operate on unused television channel
- *Unlicensed-Band Cognitive Radio*, which can only utilize unlicensed parts of the radio frequency (RF) spectrum One such system is described in the IEEE 802.15Task Group 2 specifications,[6] which focus on the coexistence of IEEE 802.11 and Bluetooth.
- *Spectrum mobility*: Process by which a cognitive-radio user changes its frequency of operation. Cognitive-radio networks aim to use the spectrum in a dynamic manner by allowing radio terminals to operate in the best available frequency band, maintaining seamless communication requirements during transitions to better spectrum.
- *Spectrum sharing*: Spectrum sharing cognitive radio networks allow cognitive radio users to share the spectrum bands of the licensed-band users. However, the cognitive

radio users have to restrict their transmit power so that the interference caused to the licensed-band users is kept below a certain threshold.

- *Sensing-based Spectrum sharing*: In sensing-based spectrum sharing cognitive radio networks, cognitive radio users first listen to the spectrum allocated to the licensed users to detect the state of the licensed users. Based on the detection results, cognitive radio users decide their transmission strategies. If the licensed users are not using the bands, cognitive radio users will transmit over those bands. If the licensed users are using the bands, cognitive radio users share the spectrum bands with the licensed users by restricting their transmit power

Technology

Although cognitive radio was initially thought of as a software-defined radio extension (full cognitive radio), most research work focuses on spectrum-sensing cognitive radio (particularly in the TV bands). The chief problem in spectrum-sensing cognitive radio is designing high-quality spectrum-sensing devices and algorithms for exchanging spectrum-sensing data between nodes. It has been shown that a simple energy detector cannot guarantee the accurate detection of signal presence,[9] calling for more sophisticated spectrum sensing techniques and requiring information about spectrum sensing to be regularly exchanged between nodes. Increasing the number of cooperating sensing nodes decreases the probability of false detection.

Filling free RF bands adaptively, using OFDMA, is a possible approach. Timo A. Weiss and Friedrich K. Jondral of the University of Karlsruhe proposed a spectrum poolingsystem, in which free bands (sensed by nodes) were immediately filled by OFDMA subbands. Applications of spectrum-sensing cognitive radio include emergency-network andWLAN higher throughput and transmission-distance extensions. The evolution of cognitive radio toward cognitive networks is underway; the concept of cognitive networks is to intelligently organize a network of cognitive radios.

Functions

The main functions of cognitive radios are:

- *Power Control*: Power control is usually used for spectrum sharing CR systems to maximize the capacity of secondary users with interference power constraints to protect the primary users.
- *Spectrum sensing*: Detecting unused spectrum and sharing it, without harmful interference to other users; an important requirement of the cognitive-radio network is to sense empty spectrum. Detecting primary users is the most efficient way to detect empty spectrum. Spectrum-sensing techniques may be grouped into three categories:
- *Transmitter detection*: Cognitive radios must have the capability to determine if a signal from a primary transmitter is locally present in a certain spectrum. There are several proposed approaches to transmitter detection:
- Matched filter detection
- Energy detection: Energy detection is a spectrum sensing method that detects the presence/absence of a signal just by measuring the received signal power.[14] This signal detection approach is quite easy and convenient for practical implementation. To implement energy detector, however, noise variance information is required. It has been shown that an imperfect knowledge of the noise power (noise uncertainty) may lead to the phenomenon of the SNR wall, which is a SNR level below which the energy detector can not reliably detect any transmitted signal even increasing the observation time.[15] It[16] has also been shown that the SNR wall is not caused by the presence of a noise uncertainty itself, but by an insufficient refinement of the noise power estimation while the observation time increases.
- Cyclostationary-feature detection: These type of spectrum sensing algorithms are motivated because most man-made communication signals, such as BPSK, QPSK,AM, OFDM, etc. exhibit cyclostationary behavior.[17] However, noise signals (typically white noise) do not exhibit cyclostationary behavior. These detectors are robust against noise variance uncertainty. The aim of such detectors is to exploit the

cyclostationary nature of man-made communication signals buried in noise. Cyclostationary detectors can be either single cycle or multicycle cyclostatonary.

- *Wideband spectrum sensing*: refers to spectrum sensing over large spectral bandwidth, typically hundreds of MHz or even several GHz. Since current ADC technology cannot afford the high sampling rate with high resolution, it requires revolutional techniques, e.g., compressive sensing and sub-Nyquist sampling.[18]
- *Cooperative detection*: Refers to spectrum-sensing methods where information from multiple cognitive-radio users is incorporated for primary-user detection[19]
- *Interference-based detection*
- *Null-space based CR*: With the aid of multiple antennas, CR detects the null-space of the primary-user and then transmits within the null-space, such that its subsequent transmission causes less interference to the primary-user
- *Spectrum management*: Capturing the best available spectrum to meet user communication requirements, while not creating undue interference to other (primary) users. Cognitive radios should decide on the best spectrum band (of all bands available) to meet quality of service requirements; therefore, spectrum-management functions are required for cognitive radios. Spectrum-management functions are classified as:
  - *Spectrum analysis*
  - *Spectrum decision*[20]

The practical implementation of spectrum-management functions is a complex and multifaceted issue, since it must address a variety of technical and legal requirements. An example of the former is choosing an appropriate sensing threshold to detect other users, while the latter is exemplified by the need to meet the rules and regulations set out for radio spectrum access in international (ITU radio regulations) and national (telecommunications law) legislation.

Versus intelligent antenna (IA)

An intelligent antenna (or smart antenna) is an antenna technology that uses spatial beam-formation and spatial coding to cancel interference; however, applications are emerging for extension to intelligent multiple or cooperative-antenna arrays for application to complex communication environments. Cognitive radio, by comparison, allows user terminals to sense whether a portion of the spectrum is being used in order to share spectrum with neighbor users. The following table compares the two:

Point	Cognitive radio (CR)	Intelligent antenna (IA)
Principal goal	Open spectrum sharing	Ambient spatial reuse
Interference processing	Avoidance by spectrum sensing	Cancellation by spatial precoding/post-coding
Key cost	Spectrum sensing and multi-band RF	Multiple- or cooperative-antenna arrays
Challenging algorithm	Spectrum management tech	Intelligent spatial beamforming/coding tech
Applied techniques	Cognitive software radio	Generalized dirty paper coding and Wyner-Ziv coding
Basement approach	Orthogonal modulation	Cellular based smaller cell

Point	Cognitive radio (CR)	Intelligent antenna (IA)
Competitive technology	Ultra-wideband for greater band utilization	Multi-sectoring (3, 6, 9, so on) for higher spatial reuse
Summary	Cognitive spectrum-sharing technology	Intelligent spectrum reuse technology

Note that both techniques can be combined as illustrated in many contemporary transmission scenarios.Cooperative MIMO (CO-MIMO) combines both techniques.

Applications

CR can sense its environment and, without the intervention of the user, can adapt to the user's communications needs while conforming to FCC rules in the United States. In theory, the amount of spectrum is infinite; practically, for propagation and other reasons it is finite because of the desirability of certain spectrum portions. Assigned spectrum is far from being fully utilized, and efficient spectrum use is a growing concern; CR offers a solution to this problem. A CR can intelligently detect whether any portion of the spectrum is in use, and can temporarily use it without interfering with the transmissions of other users.[22] According to Bruce Fette, "Some of the radio's other cognitive abilities include determining its location, sensing spectrum use by neighboring devices, changing frequency, adjusting output power or even altering transmission parameters and characteristics. All of these capabilities, and others yet to be realized, will provide wireless spectrum users with the ability to adapt to real-time spectrum conditions, offering regulators, licenses and the general public flexible, efficient and comprehensive use of the spectrum".

Simulation of CR networks

At present, modeling & simulation is the only paradigm which allows the simulation of complex behavior in a given environment's cognitive radio networks. Network simulators likeOPNET, NetSim, MATLAB and NS2 can be used to simulate a cognitive radio network. Areas of research using network simulators include:

1. Spectrum sensing & incumbent detection
2. Spectrum allocation
3. Measurement and modeling of spectrum usage
4. Efficiency of spectrum utilization

Unit V

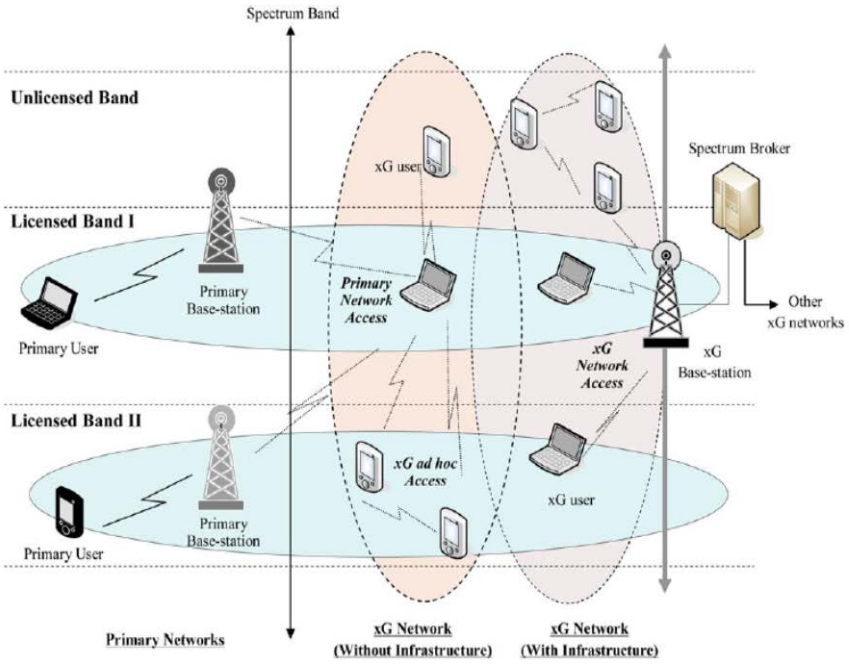
Next Generation Wireless Networks

Part-A

1) Define XG? What is the need of it? (L-1, CO-5)

Today's wireless networks are characterized by a fixed spectrum assignment policy. However, a large portion of the assigned spectrum is used sporadically and geographical variations in the utilization of assigned spectrum ranges from 15% to 85% with a high variance in time. The limited available spectrum and the inefficiency in the spectrum usage necessitate a new communication paradigm to exploit the existing wireless spectrum opportunistically. This new networking paradigm is referred to as NeXt Generation (xG) Networks as well as Dynamic Spectrum Access (DSA) and cognitive radio networks

2) Draw the XG network architecture. (L-1, CO-5)



3) Mention the functions of cognitive radios in xG network. (L-1, CO-5)

A ‘‘Cognitive Radio’’ is a radio that can change its transmitter parameters based on interaction with the environment in which it operates

4) What is Spectrum sensing? (L-1, CO-5)

A cognitive radio monitors the available spectrum bands, captures their information, and then detects the spectrum holes.

5) What is Spectrum analysis: (L-1, CO-5)

The characteristics of the spectrum holes that are detected through spectrum sensing are estimated.

6) What is Spectrum decision? (L-1, CO-5)

A cognitive radio determines the data rate, the transmission mode, and the bandwidth of the transmission. Then, the appropriate spectrum band is chosen according to the spectrum characteristics and user requirements. Once the operating spectrum band is determined, the communication can be performed over this spectrum band. However, since the radio environment changes over time and space, the cognitive radio should keep track of the changes of the radio environment.

7) Define Reconfigurability. (L-1, CO-5)

Reconfigurability is the capability of adjusting operating parameters for the transmission on the fly without any modifications on the hardware components. This capability enables the cognitive radio to adapt easily to the dynamic radio environment.

8) Define Operating frequency? (L-1, CO-5)

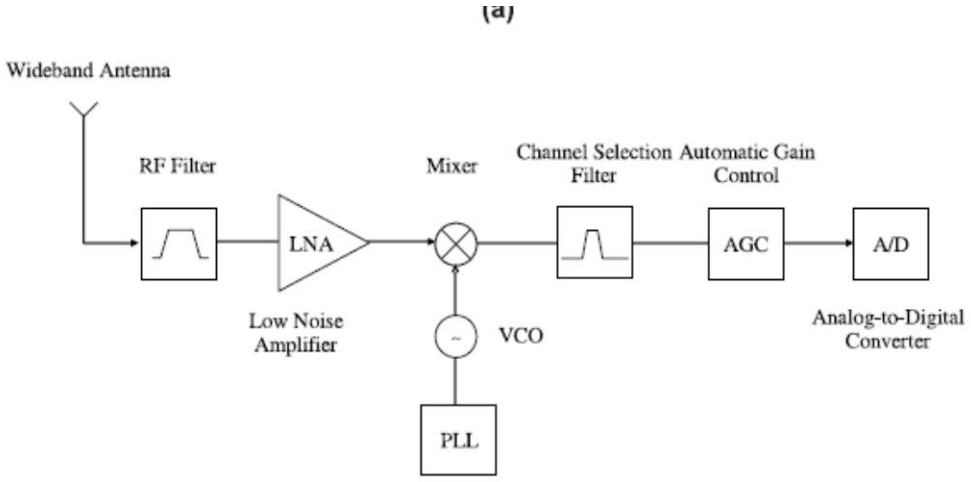
A cognitive radio is capable of changing the operating frequency. Based on the information about the radio environment, the most suitable operating frequency can be determined and the communication can be dynamically performed on this appropriate operating frequency

.

9) Define cognitive capability. (L-1, CO-5)

Cognitive capability refers to the ability of the radio technology to capture or sense the information from its radio environment. This capability cannot simply be realized by monitoring the power in some frequency band of interest but more sophisticated techniques are required in order to capture the temporal and spatial variations in the radio environment and avoid interference to other users. Through this capability, the portions of the spectrum that are unused at a specific time or location can be identified. Consequently, the best spectrum and appropriate operating parameters can be selected.

10) Draw the physical architecture of cognitive radio. (L-1, CO-5)



11) List out some main components of wideband RF front-end architecture. (L-1, CO-5)

- RF filter: The RF filter selects the desired band by bandpass filtering the received RF signal.
- Low noise amplifier (LNA): The LNA amplifies the desired signal while simultaneously minimizing noise component. • Mixer: In the mixer, the received signal is mixed with locally generated RF frequency and converted to the baseband or the intermediate frequency (IF).

- Voltage-controlled oscillator (VCO): The VCO generates a signal at a specific frequency for a given voltage to mix with the incoming signal. This procedure converts the incoming signal to baseband or an intermediate frequency.
- Phase locked loop (PLL): The PLL ensures that a signal is locked on a specific frequency and can also be used to generate precise frequencies with fine resolution.
- Channel selection filter: The channel selection filter is used to select the desired channel and to reject the adjacent channels. There are two types of channel selection filters [52]. The direct conversion receiver uses a low-pass filter for the channel selection. On the other hand, the super heterodyne receiver adopts a bandpass filter.
- Automatic gain control (AGC): The AGC maintains the gain or output power level of an amplifier constant over a wide range of input signal levels.

12) What are the steps involved in cognitive cycle? (L-1, CO-5)

1. Spectrum Analysis
2. Spectrum Decision
3. Spectrum Management

13) Define Holding Time. (L-1, CO-5)

The activities of primary users can affect the channel quality in xG networks. Holding time refers to the expected time duration that the xG user can occupy a licensed band before getting interrupted. Obviously, the longer the holding time, the better the quality would be. Since frequent spectrum handoff can decrease the holding time, previous statistical patterns of handoff should be considered while designing xG networks with large expected holding time 120 sec

14) Define path loss. (L-2, CO-5)

The path loss increases as the operating frequency increases. Therefore, if the transmission power of an xG user remains the same, then its transmission range decreases at higher frequencies. Similarly, if transmission power is increased to compensate for the increased path loss, then this results in higher interference for other users.

15) Define spectrum handoff. (L-1, CO-5)

In XG networks, spectrum mobility arises when current channel conditions become worse or a primary user appears. Spectrum mobility gives rise to a new type of handoff in XG networks that we refer to as spectrum handoff. The protocols for different layers of the network stack must adapt to the channel parameters of the operating frequency. Moreover, they should be transparent to the spectrum handoff and the associated latency.

16) **What is CCC? (L-1, CO-5)**

Many spectrum sharing solutions, either centralized or distributed, assume a CCC for spectrum sharing. It is clear that a CCC facilitates many spectrum sharing functionalities such as transmitter receiver handshake [40], communication with a central entity [7], or sensing information exchange. However, due to the fact that xG network users are regarded as visitors to the spectrum they allocate, when a primary user chooses a channel, this channel has to be vacated without interfering. This is also true for the CCC.

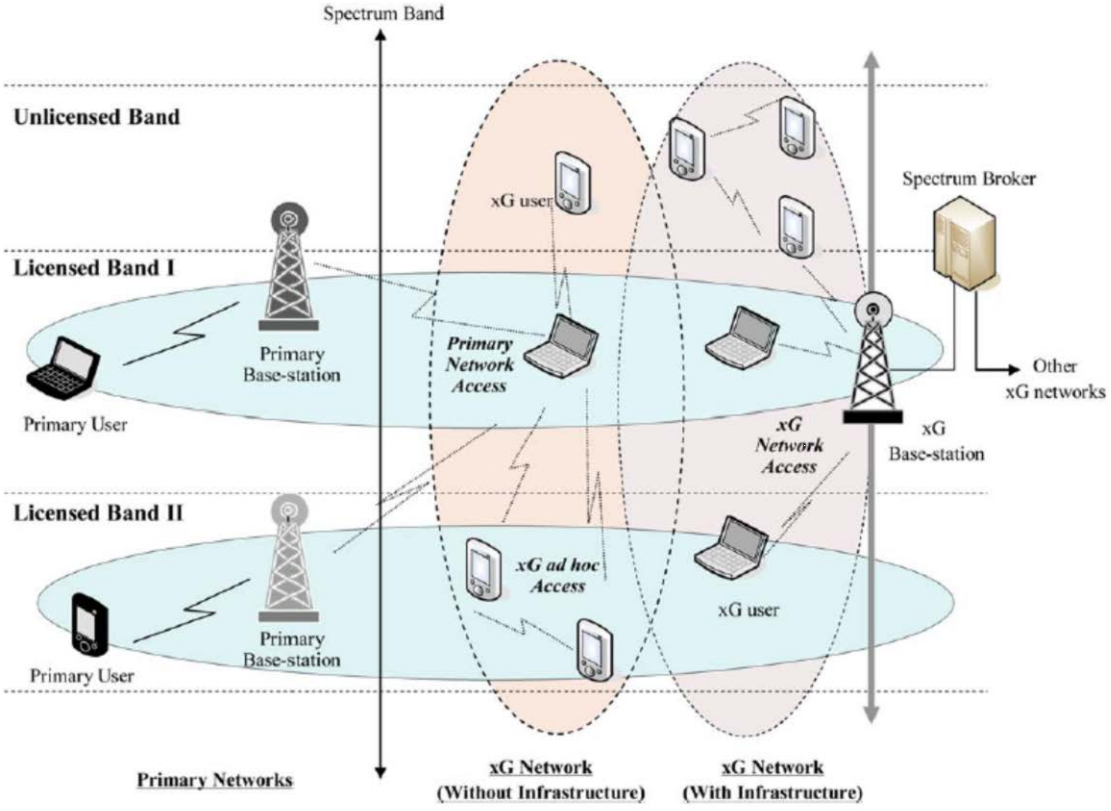
PART-B

1) Explain the XG network communication components and their interactions with diagram. (L-1, CO-5)

Existing wireless network architectures employ heterogeneity in terms of both spectrum policies and communication technologies [3]. Moreover, some portion of

the wireless spectrum is already licensed to different purposes while some bands remain unlicensed. For the development of communication protocols, a clear description of the xG network architecture is essential. In this section, the xG network architecture is presented such that all possible scenarios are considered.

The components of the xG network architecture, as shown in Fig. 6, can be classified in two groups as the primary network and the xG network. The basic elements of the primary and the xG network are defined as follows



• Primary network: An existing network infrastructure is generally referred to as the primary network, which has an exclusive right to a certain spectrum band. Examples include the common cellular and TV broadcast networks. The components of the primary network are as follows:

– Primary user: Primary user (or licensed user) has a license to operate in a certain spectrum band. This access can only be controlled by the primary base-station and should not be affected by the operations of any other unlicensed users. Primary users do not need any modification or additional functions for coexistence with xG base-stations and xG users.

– Primary base-station: Primary base-station (or licensed base-station) is a fixed infrastructure network component which has a spectrum license such as base-station transceiver system (BTS) in a cellular system. In principle, the primary base-station does not have any xG capability for sharing spectrum with xG users. However, the primary base-station may be requested to have both legacy and xG protocols for the primary network access of xG users, which is explained below.

• xG network: xG network (or cognitive radio network, Dynamic Spectrum Access network, secondary network, unlicensed network) does not have license to operate in a desired band. Hence, the spectrum access is allowed only in an opportunistic manner. xG networks can be deployed both as an infrastructure network and an ad hoc network as shown in Fig. 6. The components of an xG network are as follows:

– xG user: xG user (or unlicensed user, cognitive radio user, secondary user) has no spectrum

license. Hence, additional functionalities are required to share the licensed spectrum band.

– xG base-station: xG base-station (or unlicensed base-station, secondary base-station) is a fixed infrastructure component with xG capabilities. xG base-station provides single hop connection to xG users without spectrum access license. Through this connection, an xG user can access other networks.

– Spectrum broker: Spectrum broker (or scheduling server) is a central network entity that plays a role in sharing the spectrum resources among different xG networks. Spectrum broker can be connected to each network and can serve as a spectrum information manager to enable coexistence of multiple xG networks [10,32,70].

The reference xG network architecture is shown in Fig. 6, which consists of different types of networks: a primary network, an infrastructure based xG network, and an ad-hoc xG network. xG networks are operated under the mixed spectrum environment that consists of both licensed and unlicensed bands. Also, xG users can either communicate with each other in a multihop manner or access the base-station. Thus, in xG networks, there are three different access types as explained next:

• xG network access: xG users can access their own xG base-station both on licensed and unlicensed spectrum bands.

• xG ad hoc access: xG users can communicate with other xG users through ad hoc connection on both licensed and unlicensed spectrum bands.

• Primary network access: The xG users can also access the primary base-station through the licensed band.

According to the reference architecture shown in Fig. 6, various functionalities are required to support the heterogeneity in xG networks. In Section 3.1, we describe the xG network functions to support the heterogeneity of the network environment. Moreover, in Sections 3.2and 3.3, we overview xG network applications and existing architectures

xG network functions

As explained before, xG network can operate in both licensed and unlicensed bands. Hence, the functionalities required for xG networks vary according to whether the spectrum is licensed or unlicensed. Accordingly, in this section, we classify the xG network operations as xG network on licensed band and xG network on unlicensed band. The xG network functions are explained in the following sections according to this classification.

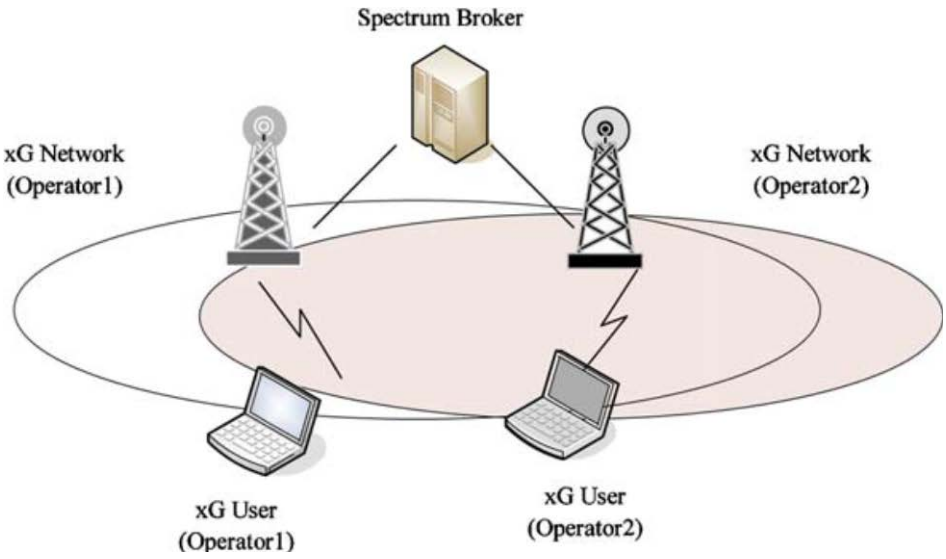
xG network on licensed band

As shown in Fig. 1, there exist temporally unused spectrum holes in the licensed spectrum band.Hence, xG networks can be deployed to exploit these spectrum holes through cognitive communication techniques. This architecture is depicted in Fig. 7, where the xG network coexists with the primary network at the same location and on the

same spectrum band. There are various challenges for xG networks on licensed band due to the existence of the primary users. Although the main purpose of the xG network is to determine the best available spectrum, xG functions in the licensed band are mainly aimed at the detection of the presence of primary users. The channel capacity of the spectrum holes depends on the interference at the nearby primary users. Thus, the interference avoidance with primary users is the most important issue in this architecture. Furthermore, if primary users appear in the spectrum band occupied by xG users, xG users should vacate the current spectrum band and move to the new available spectrum immediately, called spectrum handoff.

xG network on unlicensed band

Open spectrum policy that began in the industrial scientific and medical (ISM) band has caused an impressive variety of important technologies and innovative uses. However, due to the interference among multiple heterogeneous networks, the spectrum efficiency of ISM band is decreasing. Ultimately, the capacity of open spectrum access and the quality of service they can offer, depend on the degree to which a radio can be designed to allocate spectrum efficiently. xG networks can be designed for operation on unlicensed bands such that the efficiency is improved in this portion of the spectrum. The xG network on unlicensed band architecture is illustrated in Fig. 8. Since there are no license holders, all network entities have the same right to access the spectrum bands. Multiple xG networks coexist in the same area and communicate using the same portion of the spectrum. Intelligent spectrum sharing algorithms can improve the efficiency of spectrum usage and support high QoS.



In this architecture, xG users focus on detecting the transmissions of other xG users. Unlike the licensed band operations, the spectrum handoff is not triggered by the appearance of other primary users. However, since all xG users have the same right to access the spectrum, xG users should compete with each other for the same unlicensed band. Thus, sophisticated spectrum sharing methods among xG users are required in this architecture. If multiple xG network operators reside in the same unlicensed band, fair spectrum sharing among these networks is also required.

xG network applications

xG networks can be applied to the following cases:

Leased network: The primary network can provide a leased network by allowing opportunistic access to its licensed spectrum with the agreement with a third party without sacrificing the service quality of the primary user [56]. For example, the primary network can lease its spectrum access right to a mobile virtual network operator (MVNO). Also the primary network can provide its spectrum access rights to a regional community for the purpose of broadband access. Cognitive mesh network: Wireless mesh networks are emerging as a cost-effective technology for providing broadband connectivity [4]. However, as the network density increases and the applications require higher

throughput, mesh networks require higher capacity to meet the requirements of the applications. Since the cognitive radio technology enables the access to larger amount of spectrum, xG networks can be used for mesh networks that will be deployed in dense urban areas with the possibility of significant contention [38]. For example, the coverage area of xG networks can be increased when a meshed wireless backbone network of infrastructure links is established based on cognitive access points (CAPs) and fixed cognitive relay nodes (CRNs) [6]. The capacity of a CAP, connected via a wired broadband access to the Internet, is distributed into a large area with the help of a fixed CRN. xG networks have the ability to add temporary or permanent spectrum to the infrastructure links used for relaying in case of high traffic load.

Emergency network: Public safety and emergency networks are another area in which xG networks can be implemented [41]. In the case of natural disasters, which may temporarily disable or destroy existing communication infrastructure, emergency personnel working in the disaster areas need to establish emergency networks. Since emergency networks deal with

the critical information, reliable communication should be guaranteed with minimum latency. In addition, emergency communication requires a significant amount of radio spectrum for handling huge volume of traffic including voice, video and data. xG networks can enable the usage of the existing spectrum without the need for an infrastructure and by maintaining communication priority and response time.

Military network: One of the most interesting potential applications of an xG network is in a military radio environment [47]. xG networks can enable the military radios choose arbitrary, intermediate frequency (IF) bandwidth, modulation schemes, and coding schemes, adapting to the variable radio environment of battlefield. Also military networks have a strong need for security and protection of the communication in hostile environment. xG networks could allow military personnel to perform spectrum handoff to find secure spectrum band for themselves and their allies.

2) With physical architecture, discuss about the uses, characteristics and objectives of cognitive radio in xG network. (L-1, CO-5)

Cognitive radio technology is the key technology that enables an xG network to use spectrum in a dynamic manner. The term, cognitive radio, can formally be defined as follows [20]:

''Cognitive Radio'' is a radio that can change its transmitter parameters based on interaction with the environment in which it operates.

From this definition, two main characteristics of the cognitive radio can be defined [27,58]:

- *Cognitive capability*: Cognitive capability refers to the ability of the radio technology to capture or sense the information from its radio environ- ment. This capability cannot simply be realized by monitoring the power in some frequency band of interest but more sophisticated techniques are required in order to capture the temporal and spatial variations in the radio environment and avoid interference to other users. Through this capability, the portions of the spectrum that are unused at a specific time or location can be identified. Consequently, the best spectrum and appropriate operating parameters can be selected.
- *Reconfigurability*: The cognitive capability pro- vides spectrum awareness whereas reconfigu- rability enables the radio to be dynamically programmed according to the radio environ- ment. More specifically, the cognitive radio can be programmed to transmit and receive on a variety of frequencies and to use diff erent trans- mission access technologies supported by its hardware design [34].

The cognitive radio concept was first introduced in [45,46], where the main focus was on the radio knowledge representation language (RKRL) and how the cognitive radio can enhance the flexibility of personal wireless services. The cognitive radio is regarded as a small part of the physical world to use and provide information from environment.

The ultimate objective of the cognitive radio is to obtain the best available spectrum

through cogni- tive capability and reconfigurability as described

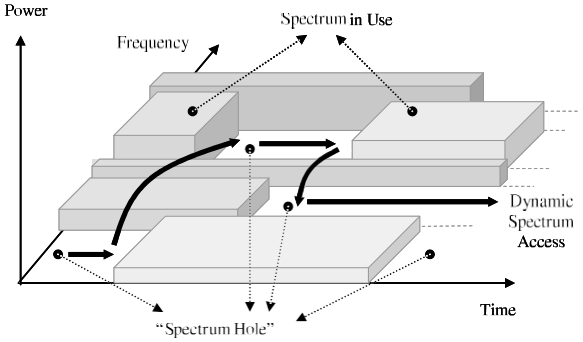


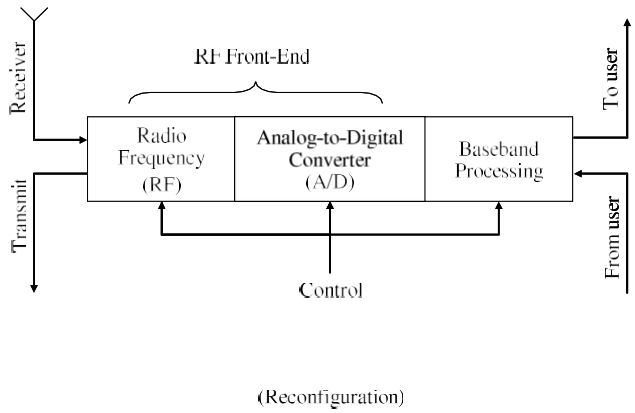
Fig. 3. Spectrum hole concept.

before. Since most of the spectrum is already assigned, the most important challenge is to share the licensed spectrum without interfering with the transmission of other licensed users as illustrated in Fig. 3. The cognitive radio enables the usage of temporally unused spectrum, which is referred to as *spectrum hole* or *white space* [27]. If this band is further used by a licensed user, the cognitive radio moves to another spectrum hole or stays in the same band, altering its transmission power level or modulation scheme to avoid interference as shown in Fig. 3.

In the following subsections, we describe the physical architecture, cognitive functions and recon- figurability capabilities of the cognitive radio technology.

Physical architecture of the cognitive radio

A generic architecture of a cognitive radio trans- ceiver is shown in Fig. 4(a) [34]. The main compo- nents of a cognitive radio transceiver are the radio front-end and the baseband processing unit. Each component can be reconfigured via a control bus



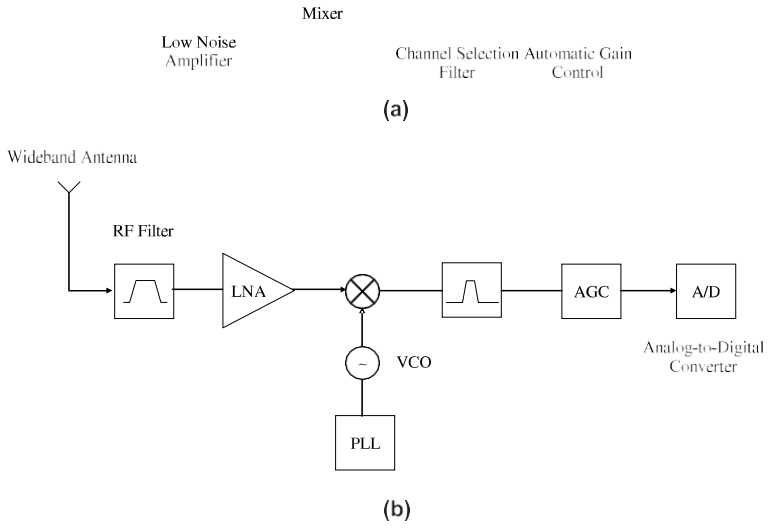


Fig. 4. Physical architecture of the cognitive radio [12,34]: (a) Cognitive radio transceiver and (b) wideband RF/analog front-end architecture.

to adapt to the time-varying RF environment. In the RF front-end, the received signal is amplified, mixed and A/D converted. In the baseband process- ing unit, the signal is modulated/demodulated and encoded/decoded. The baseband processing unit of a

cognitive radio is essentially similar to existing transceivers. However, the novelty of the cognitive radio is the RF front-end. Hence, next, we focus on the RF front-end of the cognitive radios.

The novel characteristic of cognitive radio trans- ceiver is a wideband sensing capability of the RF front-end. This function is mainly related to RF hardware technologies such as wideband antenna, power amplifier, and adaptive filter. RF hardware for the cognitive radio should be capable of tuning to any part of a large range of frequency spectrum. Also such spectrum sensing enables real-time measurements of spectrum information from radio environment. Generally, a wideband front-end archi- tecture for the cognitive radio has the following struc- ture as shown in Fig. 4(b) [12]. The components of a cognitive radio RF front-end are as follows:

- *RF filter*: The RF filter selects the desired band by bandpass filtering the received RF signal.
- *Low noise amplifier (LNA)*: The LNA amplifies the desired signal while simultaneously minimiz- ing noise component.
- *Mixer*: In the mixer, the received signal is mixed with locally generated RF frequency and con- verted to the baseband or the intermediate fre- quency (IF).
- *Voltage-controlled oscillator (VCO)*: The VCO generates a signal at a specific frequency for a given voltage to mix with the incoming signal. This procedure converts the incoming signal to baseband or an intermediate frequency.
- *Phase locked loop (PLL)*: The PLL ensures that a signal is locked on a specific frequency and can also be used to generate precise frequencies with fine resolution.
- *Channel selection filter*: The channel selection fil- ter is used to select the desired channel and to reject the adjacent channels. There are two types of channel selection filters [52]. The *direct conver- sion receiver* uses a low-pass filter for the channel selection. On the other hand, the *superheterodyne receiver* adopts a bandpass filter.
- *Automatic gain control (AGC)*: The AGC main- tains the gain or output power level of an ampli- fier constant over a wide range of input signal levels.

In this architecture, a wideband signal is received through the RF front-end, sampled by the high speed analog-to-digital (A/D) converter, and mea- surements are performed for the detection of the licensed user signal. However, there exist some limi- tations on developing the cognitive radio front-end. The wideband RF antenna receives signals from var- ious transmitters operating at different power levels, bandwidths, and locations. As a result, the RF front- end should have the capability to detect a weak sig- nal in a large dynamic range. However, this capabil- ity requires a multi-GHz speed A/D converter with high resolution, which might be infeasible [12,13].

The requirement of a multi-GHz speed A/D con- verter necessitates the dynamic range of the signal to be reduced before A/D conversion. This reduction can be achieved by filtering strong signals. Since strong signals can be located anywhere in the wide spectrum range, tunable notch filters are required for the reduction [12]. Another approach is to use multiple antennas such that signal filtering is per- formed in the spatial domain rather than in the fre- quency domain. Multiple antennas can receive signals selectively using beamforming techniques [13].

As explained previously, the key challenge of the physical architecture of the cognitive radio is an accurate detection of weak signals of licensed users over a wide spectrum range. Hence, the implementa- tion of RF wideband front-end and A/D converter are critical issues in xG networks.

Cognitive capability

The cognitive capability of a cognitive radio enables real time interaction with its environment to determine appropriate communication parame- ters and adapt to the dynamic radio environment. The tasks required for adaptive operation in open spectrum are shown in Fig. 5 [27,46,58], which is referred to as the *cognitive cycle*. In this section, we provide an overview of the three main steps of the cognitive cycle: *spectrum sensing*, *spectrum anal- ysis*, and *spectrum decision*. The details and the related work of these functions are described in Sec- tions 4 and 5.

The steps of the cognitive cycle as shown in Fig. 5 are as follows:

1. *Spectrum sensing*: A cognitive radio monitors the available spectrum bands, captures their infor- mation, and then detects the spectrum holes.

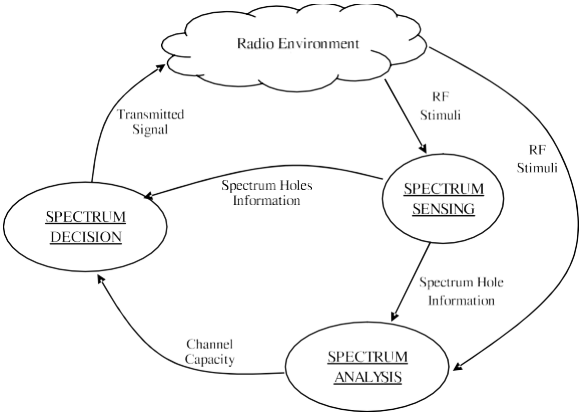


Fig. 5. Cognitive cycle.

2. *Spectrum analysis*: The characteristics of the spectrum holes that are detected through spec- trum sensing are estimated.
3. *Spectrum decision*: A cognitive radio determines the data rate, the transmission mode, and the bandwidth of the transmission. Then, the appro- priate spectrum band is chosen according to the spectrum characteristics and user requirements.

Once the operating spectrum band is determined, the communication can be performed over this spec- trum band. However, since the radio environment changes over time and space, the cognitive radio should keep track of the changes of the radio envi- ronment. If the current spectrum band in use becomes unavailable, the *spectrum mobility* function

that will be explained in Section 6, is performed to provide a seamless transmission. Any environmen- tal change during the transmission such as primary user appearance, user movement, or traffic variation can trigger this adjustment.

Reconfigurability

Reconfigurability is the capability of adjusting operating parameters for the transmission on the fly without any modifications on the hardware com- ponents. This capability enables the cognitive radio to adapt easily to the dynamic radio environment. There are several reconfigurable parameters that can be incorporated into the cognitive radio [20] as explained below:

- *Operating frequency*: A cognitive radio is capable of changing the operating frequency. Based on the information about the radio environment,

  the most suitable operating frequency can be determined and the communication can be dynamically performed on this appropriate oper- ating frequency.
- *Modulation*: A cognitive radio should reconfigure the modulation scheme adaptive to the user requirements and channel conditions. For exam- ple, in the case of delay sensitive applications, the data rate is more important than the error rate. Thus, the modulation scheme that enables the higher spectral efficiency should be selected. Con- versely, the loss-sensitive applications focus on the error rate, which necessitate modulation schemes with low bit error rate.
- *Transmission power*: Transmission power can be reconfigured within the power constraints. Power control enables dynamic transmission power con- figuration within the permissible power limit. If higher power operation is not necessary, the cog- nitive radio reduces the transmitter power to a lower level to allow more users to share the spec- trum and to decrease the interference.
- *Communication technology*: A cognitive radio can also be used to provide interoperability among diff erent communication systems.

The transmission parameters of a cognitive radio can be reconfigured not only at the beginning of a transmission but also during the transmission. According to the spectrum characteristics, these parameters can be reconfigured such that the cognitive radio is switched to a different spectrum band, the transmitter and receiver parameters are reconfigured and the appropriate communication protocol parameters and modulation schemes are used.

3) Explain the each components and its functionality of xG network architecture. (L-1, CO-5)

Existing wireless network architectures employ heterogeneity in terms of both spectrum policies and communication technologies [3]. Moreover, some portion of the wireless spectrum is already licensed to different purposes while some bands remain unlicensed. For the development of commu- nication protocols, a clear description of the xG net- work architecture is essential. In this section, the xG network architecture is presented such that all pos- sible scenarios are considered.

The components of the xG network architecture, as shown in Fig. 6, can be classified in two groups as the *primary network* and the *xG network*. The basic

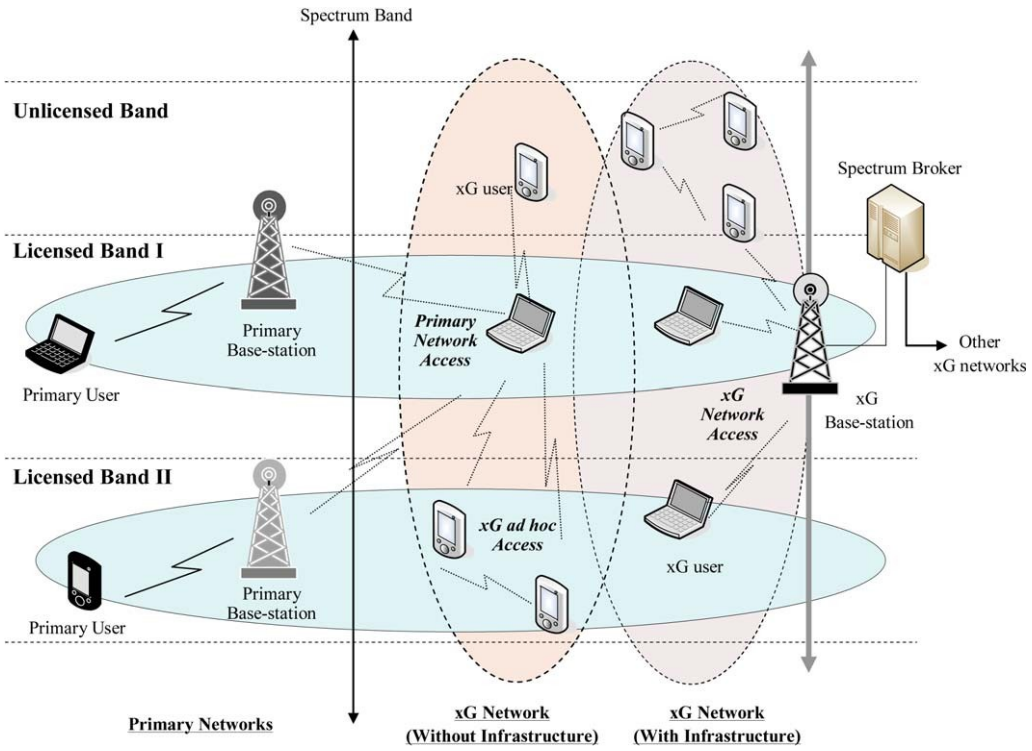


Fig. 6. xG network architecture.

elements of the primary and the xG network are defined as follows:

- *Primary network*: An existing network infrastruc- ture is generally referred to as the primary net- work, which has an exclusive right to a certain spectrum band. Examples include the common cellular and TV broadcast networks. The compo- nents of the primary network are as follows:
  - *Primary user*: Primary user (or licensed user) has a license to operate in a certain spectrum band. This access can only be controlled by the primary base-station and should not be aff ected by the operations of any other unli- censed users. Primary users do not need any modification or additional functions for coex- istence with xG base-stations and xG users.

- *Primary base-station*: Primary base-station (or licensed base-station) is a fixed infrastructure network component which has a spectrum license such as base-station transceiver system (BTS) in a cellular system. In principle, the primary base-station does not have any xG capability for sharing spectrum with xG users. However, the primary base-station may be requested to have both legacy and xG protocols for the *primary network access* of xG users, which is explained below.

• *xG network*: xG network (or cognitive radio net- work, Dynamic Spectrum Access network, sec- ondary network, unlicensed network) does not have license to operate in a desired band. Hence, the spectrum access is allowed only in an oppor- tunistic manner. xG networks can be deployed both as an infrastructure network and an ad hoc network as shown in Fig. 6. The components of an xG network are as follows:

- *xG user*: xG user (or unlicensed user, cognitive radio user, secondary user) has no spectrum license. Hence, additional functionalities are required to share the licensed spectrum band.
- *xG base-station*: xG base-station (or unli- censed base-station, secondary base-station) is a fixed infrastructure component with xG capabilities. xG base-station provides single hop connection to xG users without spectrum access license. Through this connection, an xG user can access other networks.
- *Spectrum broker*: Spectrum broker (or sched- uling server) is a central network entity that plays a role in sharing the spectrum resources among diff erent xG networks. Spectrum broker can be connected to each network and can serve as a spectrum information manager to enable coexistence of multiple xG networks [10,32,70].

The reference xG network architecture is shown in Fig. 6, which consists of different types of net- works: a primary network, an infrastructure based xG network, and an ad-hoc xG network. xG net- works are operated under the mixed spectrum envi- ronment that consists of both licensed and unlicensed bands. Also, xG users can either commu- nicate with each other in a multihop manner or access the base-station. Thus, in xG networks,

there are three different access types as explained next:

- *xG network access*: xG users can access their own xG base-station both on licensed and unlicensed spectrum bands.
- *xG ad hoc access*: xG users can communicate with other xG users through ad hoc connection on both licensed and unlicensed spectrum bands.
- *Primary network access*: The xG users can also access the primary base-station through the licensed band.

According to the reference architecture shown in Fig. 6, various functionalities are required to sup- port the heterogeneity in xG networks. In Section 3.1, we describe the xG network functions to sup- port the heterogeneity of the network environment. Moreover, in Sections 3.2 and 3.3, we overview xG network applications and existing architectures.

xG network functions

As explained before, xG network can operate in both licensed and unlicensed bands. Hence, the func- tionalities required for xG networks vary according to whether the spectrum is licensed or unlicensed. Accordingly, in this section, we classify the xG net- work operations as *xG network on licensed band* and *xG network on unlicensed band*. The xG network functions are explained in the following sections according to this classification.

xG network on licensed band

As shown in Fig. 1, there exist temporally unused spectrum holes in the licensed spectrum band. Hence, xG networks can be deployed to exploit these spectrum holes through cognitive communi- cation techniques. This architecture is depicted in

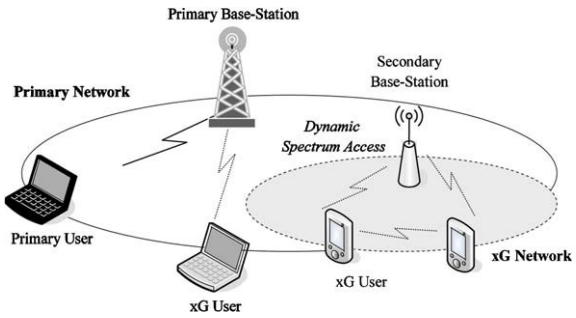


Fig. 7. xG network on licensed band.

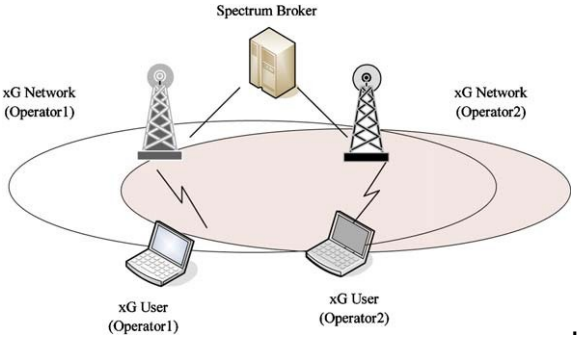
Fig. 7, where the xG network coexists with the pri- mary network at the same location and on the same spectrum band.

There are various challenges for xG networks on licensed band due to the existence of the primary users. Although the main purpose of the xG net- work is to determine the best available spectrum, xG functions in the licensed band are mainly aimed at the detection of the presence of primary users. The channel capacity of the spectrum holes depends on the interference at the nearby primary users. Thus, the interference avoidance with primary users is the most important issue in this architecture. Furthermore, if primary users appear in the spec- trum band occupied by xG users, xG users should vacate the current spectrum band and move to the new available spectrum immediately, called *spec- trum handoff*.

xG network on unlicensed band

Open spectrum policy that began in the industrial scientific and medical (ISM) band has caused an impressive variety of important technologies and innovative uses. However, due to the interference among multiple heterogeneous networks, the spec- trum efficiency of ISM band is decreasing. Ulti- mately, the capacity of open spectrum access, and the quality of service they can offer, depend on the degree to which a radio can be designed to allocate spectrum efficiently.

xG networks can be designed for operation on unlicensed bands such that the efficiency is improved in this portion of the spectrum. The *xG network on unlicensed band* architecture is illustrated in Fig. 8. Since there are no license holders, all network entities have the same right to access the spectrum bands. Multiple xG networks coexist in the same area and communicate using the same por- tion of the spectrum. Intelligent spectrum sharing



. xG network on unlicensed band.

algorithms can improve the efficiency of spectrum usage and support high QoS.

In this architecture, xG users focus on detecting the transmissions of other xG users. Unlike the licensed band operations, the spectrum handoff is not triggered by the appearance of other primary users. However, since all xG users have the same right to access the spectrum, xG users should compete with each other for the same unlicensed band. Thus, sophisticated spectrum sharing methods among xG users are required in this architecture. If multiple xG network operators reside in the same unlicensed band, fair spectrum sharing among these networks is also required.

xG network applications

xG networks can be applied to the following cases:

Leased network: The primary network can pro- vide a leased network by allowing

opportunistic access to its licensed spectrum with the agreement with a third party without sacrificing the service quality of the primary user [56]. For example, the primary network can lease its spectrum access right to a mobile virtual network operator (MVNO). Also the primary network can provide its spectrum access rights to a regional community for the pur- pose of broadband access.

Cognitive mesh network: Wireless mesh networks are emerging as a cost-effective technology for pro- viding broadband connectivity [4]. However, as the network density increases and the applications require higher throughput, mesh networks require higher capacity to meet the requirements of the applications. Since the cognitive radio technology enables the access to larger amount of spectrum, xG networks can be used for mesh networks that l be deployed in dense urban areas with the pos- sibility of significant contention [38]. For example, the coverage area of xG networks can be increased when a meshed wireless backbone network of infra- structure links is established based on cognitive access points (CAPs) and fixed cognitive relay nodes (CRNs) [6]. The capacity of a CAP, connected via a wired broadband access to the Internet, is distrib- uted into a large area with the help of a fixed CRN. xG networks have the ability to add tempo- rary or permanent spectrum to the infrastructure links used for relaying in case of high traffic load.

Emergency network: Public safety and emergency networks are another area in which xG networks can be implemented [41]. In the case of natural disasters, which may temporarily disable or destroy existing communication infrastructure, emergency personnel working in the disaster areas need to establish *emer- gency networks*. Since emergency networks deal with the critical information, reliable communication should be guaranteed with minimum latency. In addition, emergency communication requires a sig- nificant amount of radio spectrum for handling huge volume of traffic including voice, video and data. xG networks can enable the usage of the existing spec- trum without the need for an infrastructure and by maintaining communication priority and response time.

Military network: One of the most interesting potential applications of an xG network is in a mil- itary radio environment [47]. xG networks can enable the military radios choose

arbitrary, interme- diate frequency (IF) bandwidth, modulation schemes, and coding schemes, adapting to the vari- able radio environment of battlefield. Also military networks have a strong need for security and pro- tection of the communication in hostile environ- ment. xG networks could allow military personnel to perform spectrum handoff to find secure spectrum band for themselves and their allies.

4) Discuss about spectrum management. (L-1, CO-5)

In xG networks, the unused spectrum bands will be spread over wide frequency range including both unlicensed and licensed bands. These unused spec- trum bands detected through spectrum sensing show different characteristics according to not only the time varying radio environment but also the spectrum band information such as the operating frequency and the bandwidth.

Since xG networks should decide on the best spectrum band to meet the QoS requirements over all available spectrum bands, new spectrum man- agement functions are required for xG networks, considering the dynamic spectrum characteristics. We classify these functions as *spectrum sensing*, *spectrum analysis*, and *spectrum decision*. While *spectrum sensing*, which is discussed in Section 4, is primarily a PHY layer issue, *spectrum analysis* and *spectrum decision* are closely related to the upper layers. In this section, spectrum analysis and spectrum decision are investigated.

Spectrum analysis

In xG networks, the available spectrum holes show different characteristics which vary over time. Since the xG users are equipped with the cognitive radio based physical layer, it is important to under- stand the characteristics of different spectrum bands. Spectrum analysis enables the characteriza- tion of different spectrum bands, which can be exploited to get the spectrum band appropriate to the user requirements.

In order to describe the dynamic nature of xG networks, each spectrum hole should be character- ized considering not only the time-varying radio environment and but also

the primary user activity and the spectrum band information such as operat- ing frequency and bandwidth. Hence, it is essential to define parameters such as interference level, channel error rate, path-loss, link layer delay, and holding time that can represent the quality of a par- ticular spectrum band as follows:

- *Interference*: Some spectrum bands are more crowded compared to others. Hence, the spectrum band in use determines the interference character- istics of the channel. From the amount of the interference at the primary receiver, the permissi- ble power of an xG user can be derived, which is used for the estimation of the channel capacity.
- *Path loss*: The path loss increases as the operating frequency increases. Therefore, if the transmis- sion power of an xG user remains the same, then its transmission range decreases at higher fre- quencies. Similarly, if transmission power is increased to compensate for the increased path loss, then this results in higher interference for other users.
- *Wireless link errors*: Depending on the modula- tion scheme and the interference level of the spec- trum band, the error rate of the channel changes.
- *Link layer delay*: To address diff erent path loss, wireless link error, and interference, diff erent types of link layer protocols are required at dif- ferent spectrum bands. This results in diff erent link layer packet transmission delay.
- *Holding time*: The activities of primary users can aff ect the channel quality in xG networks. Hold- ing time refers to the expected time duration that the xG user can occupy a licensed band before getting interrupted. Obviously, the longer the holding time, the better the quality would be. Since frequent spectrum handoff can decrease the holding time, previous statistical patterns of handoff should be considered while designing xG networks with large expected holding time.

Channel capacity, which can be derived from the parameters explained above, is the most important factor for spectrum characterization. Usually SNR at the receiver has been used for the capacity estima- tion. However, since SNR considers only local

observations of xG users, it is not enough to avoid interference at the primary users. Thus, spectrum haracterization is focused on the capacity estima- tion based on the interference at the licensed receiv- ers. The interference temperature model [21] given in Section 4.3 can be exploited for this approach. The interference temperature limit indicates an upper bound or cap on the potential RF energy that could be introduced into the band. Consequently, using the amount of permissible interference, the maximum permissible transmission power of an xG user can be determined.

In [63], a spectrum capacity estimation method has been proposed that considers the bandwidth and the permissible transmission power. Accord- ingly, the spectrum capacity, *C*, can be estimated as follows:be selected for the current transmission considering the QoS requirements and the spectrum characteris- tics. Thus, the spectrum management function must be aware of user QoS requirements.

Based on the user requirements, the data rate, acceptable error rate, delay bound, the transmission mode, and the bandwidth of the transmission can be determined. Then, according to the decision rule, the set of appropriate spectrum bands can be chosen. In [73], five spectrum decision rules are presented, which are focused on fairness and communication cost. However, this method assumes that all chan- nels have similar throughput capacity. In [36], an opportunistic frequency channel skipping protocol is proposed for the search of better quality channel.

5).**Explain in detail about challenges in spectrum sharing. (L-1, CO-5)**

The followings are the open research issues for efficient spectrum mobility in xG networks.

- At a particular time, several frequency bands may be available for an xG user. Algorithms are required to decide the best available spectrum based on the channel characteristics of the avail- able spectrum and the requirements of the appli- cations that are being used by an xG user.
- Once, the best available spectrum is selected, the next challenge is to design new mobility and con- nection management approaches to reduce delay and loss during

spectrum handoff .

- When the current operational frequency becomes busy (this may happen if a licensed user starts to use this frequency) in the middle of a communi- cation by an xG user, then applications running on this node have to be transferred to another available frequency band. However, the selection of new operational frequency may take time. Novel algorithms are required to ensure that applications do not suff er from severe perfor- mance degradation during such transitions.
- Spectrum handoff may occur due to reasons other than the detection of the primary user. Thus, there exist various other spectrum handoff schemes in xG networks. If an xG user moves from one place to another, spectrum handoff may occur just because the available spectrum bands change. Thus the desired spectrum handoff scheme should integrate inter-cell handoff . Apart from this, spectrum handoff between diff erent networks, referred to as vertical handoff is also likely to occur in xG networks. Under such a diverse environment, it is essential that spectrum handoff scheme takes all the above mentioned possibilities into consideration.
- Spectrum mobility in time domain: xG networks adapt to the wireless spectrum based on available bands on the spectrum. Since these available channels change over time, enabling QoS in this environment is challenging. The physical radio should ‘‘move’’ through the spectrum to meet the QoS requirements.
- Spectrum mobility in space: The available bands also change as a user moves from one place to another. Hence, continuous allocation of spec- trum is a major challenge. in xG networks.