Dynamic Cellular Cognitive System

Ying Wang

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
Electrical Engineering

Charles W. Bostian (Chair)
Allen B. MacKenzie
Claudio da Silva
Michael Hsiao
Tonya Smith-Jackson

September 25, 2009
Blacksburg, Virginia

Keywords: Cognitive Radio, Cognitive Radio Network, Signal Classification, Synchronization, Dynamic Spectrum Accessing, Channel Allocation, Software Defined Radio,

Copyright 2009, Ying Wang

Dynamic Cellular Cognitive System

Ying Wang

(Abstract)

Dynamic Cellular Cognitive System (DCCS) serves as a cognitive network for white space devices in TV white space. It is also designed to provide quality communications for first responders in area with damaged wireless communication infrastructure. In DCCS network, diverse types of communication devices interoperate, communicate, and cooperate with high spectrum efficiency in a Dynamic Spectrum Access (DSA) scenario. DCCS can expand to a broad geographical distribution via linking to existing infrastructure. DCCS can quickly form a network to accommodate a diverse set of devices in natural disaster areas. It can also recover the infrastructure in a blind spot, for example, a subway or mountain area. Its portability and low cost make it feasible for commercial applications.

This dissertation starts with an overview of DCCS network. DCCS defines a cognitive radio network and a set of protocols that each cognitive radio node inside the network must adopt to function as a user within the group. Multiple secondary users cooperate based on a fair and efficient scheme without losing the flexibility and self adaptation features. The basic unit of DCCS is a cell. A set of protocols and algorithms are defined to meet the communication requirement for intra-cell communications.

DCCS includes multiple layers and multiple protocols. This dissertation gives a comprehensive description and analysis of building a DCCS network. It covers the network architecture, physical and Medium Access Control (MAC) layers for data and command transmission, spectrum management in DSA scenario, signal classification and synchronization and describes a working prototype of DCCS.

Two key technologies of intra-cell communication are spectrum management and Universal Classification and Synchronization (UCS). A channel allocation algorithm based on calculating the throughput of an available is designed and the performance is analyzed. UCS is conceived as a self-contained system which can detect, classify, and synchronize with a received signal and extract all parameters needed for physical layer demodulation. It enables the accommodation of non-cognitive devices and improves communication quality by allowing a cognitive receiver to track physical layer changes at the transmitter.

Inter-cell communications are the backhaul connections of DCCS. This dissertation discusses two approaches to obtaining spectrum for inter-cell communications. A temporary leasing approach focuses on the policy aspects, and the other approach is based on using OFDMA to combine separate narrowband channels into a wideband channel that can meet the inter-cell communications throughput requirements.

A prototype of DCCS implemented on GNU radio and USRP platform is included in the dissertation. It serves as the proof of concept of DCCS.

Acknowledgements

I would like to express profound gratitude to my advisor, Dr. Charles W. Bostian, for his invaluable encouragement, support, guidance, throughout my study in Virginia Tech. He sees my potential and teaches me how to use it. He bears with my low tide and believes in me to move on. He encourages me to face challenges and trusts me to conquer the difficulties. He gives me the space to experiment, and points me in the right direction when I get lost. He opens doors to the professional world for me and has equipped me with the knowledge to explore. His character, integrity and wisdom are the "good will hunting" for me. I am honored to be his student and what I have learned from Dr. Bostian is the fortune for my life.

I would like to thank Ms. Judy Hood for her kindness, patience, and enormous help to me. She makes me feel at ease every day in the lab. I would like to thank Dr. Allen B. MacKenzie, Dr. Claudio da Silva, Dr. Michael Hsiao, and Dr. Tonya Smith-Jackson for their precious suggestions and support in my research.

It has been a great pleasure working with my lab mates. I would like to thank Qinqin Chen for her company in my three years of study. Without her, there won't be so many joys and achievements. We are the complement and we make a good team. I would like to thank every one of my lab mates. It is they who make the long hours in the lab a fun journey, who make every deadline a finish line.

I am as ever, especially indebted to my parents, for their unlimited love and support throughout my life.

Table of Contents

Table of Figures

List of Tables

Chapter 1 Introduction

1.1 Cognitive Radio Network Motivation

The research work of this dissertation was first motivated by the challenge of providing quality communications for first responders in areas with damaged wireless communication infrastructure. During the research, we realized that the designed Dynamic Cellular Cognitive System (DCCS) can not only be used for public safety purposes, but also as a cognitive radio network for unlicensed device communications in TV white space. This dissertation is written in a way to treat DCCS as a general solution to form a network of unlicensed devices to access the spectrum as secondary users. Public safety remains an important application for DCCS.

Cognitive radio and cognitive radio network technology have become hot research topics in recent years. Fast development of wireless communication technology has brought an explosion in popularity of technologies in the overcrowded unlicensed bands. The demand for higher wireless communications data rates has increased continually, as has the demand for greater mobility range of communication devices. The lack of interoperability among different types of devices has prevented effective communications in some cases. Cognitive radio and cognitive radio networks can fulfill the requirements for efficient utilization of spectrum and more flexible and robust network architecture while promoting interoperability.

Public safety is an important application for cognitive radios and cognitive radio networks. DCCS was first designed for public safety purposes. In natural disaster areas, infrastructure is often damaged and communication devices relying on this infrastructure are not able to function correctly. DCCS will set up a communication network at the same time that the first responders arrive. DCCS can accommodate many different types of devices, and provide a variety of services. These features are beneficial for first responders in emergency situations.

TV white space is a more general application for DCCS. White space devices access the spectrum as unlicensed devices; competition among devices might cause inefficient utilization of

the spectrum, and also cause interference for TV receivers. DCCS was designed not to rely on fixed infrastructure to provide communications. DCCS provides self organization among the nodes, and it is designed to set up a network that can use the spectrum more efficiently.

Standard communication systems like the Global System for Mobile Communications (GSM) or Universal Mobile Telecommunications System (UMTS) are based on fixed infrastructure, pre-defined channel allocation schemes and sets of protocols. Relying on fixed base stations reduces robustness in emergently situations. Predefined channel allocation schemes might generate a scarcity of spectrum in some unusual conditions. Wi-Fi and the cellular system developed independent and utilize different spectra and devices. This results in a lack of interoperability among devices and a waste of spectrum. The DCCS system is designed and implemented to solve all of these challenges.

DCCS defines a cognitive radio network and a set of protocols that each CR node inside the network must adopt to function as a user within the group. A CR adapts to channel conditions using the process of sensing an existing wireless channel, configuring a radio's operation to accommodate the perceived channel, and evaluating what happens when a change is made. Additionally, multiple secondary users cooperate by sharing spectrum information and self organize to manage resources based on a fair and efficient scheme without losing the flexibility and self adaptation features of CR.

The objective of DCCS is to provide infrastructure-independent high quality cognitive communications. It causes minimal interference to primary users and to other secondary users. It can accommodate different types of devices. Cognitive radio nodes in DCCS are self-organized to achieve fairness among users and optimal use of resources.

1.2 Challenges for Building a Cognitive Radio Network

Recent progress in cognitive radio technologies enables new advances in cognitive radio networks. From a hardware point of view, different Software Defined Radio (SDR) platforms have been developed. For example, the Universal Software Radio Peripheral (USRP) [1] and Lyrtech stand-alone advanced development platforms[2] are widely used in universities and

research organizations. From a software point of view, new waveform and new communication standards are constantly added to SDR software packages like GNU radio[3], OSSIE[4] etc. From the spectrum point of view, the FCC has given its permission for unlicensed devices to access white space as secondary users[5]. However, there are challenges in regulating and organizing cognitive radios and forming a reliable and efficient cognitive radio network. Compared to traditional radios, cognitive radios have considerable autonomy in accessing channels, choosing modulation types and determining data rates, etc. in a cognitive radio network. More coordination between transmitters and receivers is needed, and multi user schemes must avoid unnecessary competition between radios. Some of the challenges are briefly described below.

1. How can a cognitive radio avoid unacceptable interference to primary users?
2. How much interoperability should a cognitive radio provide to the currently applied standard wireless networks?
3. What are the channel resources, and how can they be fairly distributed among multiple competing cognitive radios?

In DCCS, we tried to address these three questions and provide our best solution for each of them. Facing the first challenge, we need to discuss an important feature in cognitive radio, Dynamic Spectrum Access (DSA). In Section 1.3, we discuss the essential challenge in DSA. To answer the second question, it is necessary have a basic understanding of the current wireless communication standards. Thus, in Section 1.4, we give a brief description of these standards. The current implementation of DCCS doesn't provide interfaces to many of the standard networks, but it is important to have the information in mind and to add new interfaces in the future. To answer the third question, it is necessary to have background knowledge of the available spectrum that is suitable for a cognitive radio network to access. TV white space, the 2.4 GHz ISM band and 5GHz ISM band are accessible by cognitive radios. TV white space (50MHz – 800MHz) is accessible by unlicensed devices as secondary users. The 2.4 GHz ISM band is also available for unlicensed users. However, Wi-Fi, Bluetooth, and other standard communication devices have made this band crowded. For 5GHz ISM band, which is also available to unlicensed devices, the transmission range is significantly shorter. This limits the

usefulness of cognitive radios. In Section 1.5, we address the most suitable spectrum for cognitive radio networks – white space, and some potential problems in accessing this band.

1.3 Dynamic Spectrum Access

One of the great features of CR is DSA. The demand for capacity in both cellular networks and Wireless Local Area Networks (WLANs) is increasing quickly. The purpose of DSA is to increase the spectrum efficiency by accessing vacant channels without interference to the primary users. To avoid interference with primary users, a cognitive radio needs to measure the occupancy of the spectrum it is accessing before it transmits any signal and the measurement of the spectrum is not a simple task.

" Unless a cognitive radio can measure the effect of its transmission on all possible receivers, taking a useful interference temperature measurement may not be feasible"[6]

Compared to other intelligent communication technologies like the smart antenna, CR mitigates interference by sensing the spectrum and using an idle channel. In 2004, when Qualcomm analyzed the feasibility of using CR in cellular wireless communication, one point made was that, in order to accurately avoid interference, a cognitive transmitter is required to measure the effect of its transmission on all possible receivers. On one hand, it is not an easy task for the transmitter to sense the environment of the entire set of possible receivers when the receiver distribution is geographically large. Even if the transmitter is able to do so, it might be difficult to find the optimal solution for both the transmitter and the receiver without any interference to primary users in such a large area and under such complicated conditions. Besides, CR is a self observing, self learning and self decision making radio. When it is performing as the sole secondary user, it is efficient and can reach optimal utilization of the resources. However when multiple secondary users exist, the competition among the secondary users wastes resources.

1.4 A Brief Description of Wireless Standards

Most wireless communications standards can be categorized as belonging to one of two different branches, cellular networks and wireless local area networks. Cellular networks include 1G, 2G, 3G, and the current developing LTE network. Wireless local area networks (WLAN) includes Wi-Fi, WiMAX, and Bluetooth etc. We will follow the two branches and give a brief description of wireless communications standards.

On the cellular network side, all networks except 1G are digital based. In 1G, multiplexing is FDMA, and each user is allocated a certain channel; thus, the spectrum efficiency is fairly low in a 1G network. For 2G and 3G, based on different air interfaces and multiplexing scheme, there are two basic types of standards.

One type is based on CDMA technology, and the standard is called IS 95 or cdmaOne. The uplink of IS 95 is OQPSK, and downlink is QPSK. Its multiplexing is based on CDMA using 64 bit Walsh code. The operating band of IS 95 includes 450MHz, 950MHz, 1800MHz, and 2100MHz in different areas. Based on IS 95, CDMA 2000 is considered to be the 3G version of IS 95. It provides more bandwidth and thus has a larger capacity. With more improvement based on CDMA 2000, EVDO and EVDV continue to improve the capacity. For EVDO, capacity can reach a downlink data rate of 4.9Mbits/s on average and 14.7Mbits/s on peak.

The other type of a 2G network is GSM which occupies the largest market in cellular networks so far. GSM uses GMSK as the basic air interface and TDMA as the multiplexing scheme. The biggest advantage of GMSK is its spectrum efficiency. GMSK uses minimum shift keying, MSK, and thus occupies the least possible bandwidth for a FSK modulation type. In a 200 kHz channel, the throughput can reach 270kbits/s. The voice channel is coded into a 13.3k bits/s waveform. Using TDMA as the multiplexing, each time period is divided into 8 slots, and the slots can be allocated to different users. Adding a Gaussian filter before the MSK modulation makes it more spectrum efficient and yields GMSK. The constant envelope also makes it more power efficient and reduces the complexity of Automatic Gain Control (AGC) design in the handset and base station. However, the drawback of GMSK is its Bit Error Rate (BER) performance. A Gaussian filter doesn't meet the Nyquist criteria[7], and thus it introduces Inter Symbol Interference (ISI) into the channel. For voice transmission, as long as the BER is under

10^{-3}bits error/s, it is not perceivable by human beings, and it is not difficult for GSM to deal with. This error rate is not acceptable for data. The later 2.5G GPRS and 2.75G EDGE improved this drawback and made GSM more suitable to transmit both data and voice. GPRS use the same air interface as GSM, and has a more flexible TDMA scheme by employing unused time slots of GSM to improve capacity. EDGE uses 8PSK, and offers almost triple the capacity of GSM. The 3G version of GSM is Universal Mobile Telecommunications System (UMTS), combined with High Speed Packet Access (HSPA). The core network of UMTS is built on the core network of GSM, without much change in hardware or software. The air interface of UMTS is Adaptive Modulation and Coding (AMC), and it adaptively changes coding schemes and modulation types among 8PSK, or 16QAM or 64QAM etc., according to the channel conditions. The multiplexing scheme is CDMA. This greatly increases the user capacity in one cell. An important issue in UMTS is power control, because of the critical near-far problem of CDMA. Although UMTS has a strict power control strategy, the battery life of UMTS handsets is still much shorter compared to that of GSM devices. The frequency bands of GSM, GPRS, EDGE, and UMTS are all 850 MHz, or 1900 MHz in U.S.

The developing LTE standard adopts OFDM as the downlink and SC-FDMA as the uplink multiplexing scheme. The air interface can be 16QAM or 64QAM. MIMO is also used to improve capacity and performance. SC-FDMA has a low Peak-to-Average Power Ratio (PAPR), and that makes it easy to implement on handset. The objective of LTE is to reach 100M bits/s for the downlink data rate and 50M bits/s for the uplink data rate.

Another important class of wireless communication networks is WLAN. Wi-Fi is the most common and widely used WLAN standard. Strictly speaking, WiMAX doesn't completely belong to WLAN because it is intended for building city-wide broadband communications networks. Because of its similarities to many aspects of Wi-Fi, we placed it in this category. Wi-Fi is standardized in IEEE802.11. The most popular version is IEEE 802.11g. It uses the ISM band located at 2.4GHz. Each channel is 20 MHz wide, and there are 3 non-overlapping channels. The air interfaces can be BPSK, QPSK, 8PSK, 16QAM and 64QAM, etc. The multiplexing scheme is OFDM and CSMA-CA. IEEE 802.11g provides up to 54 M bit/s data rate. It has good resistance to multipath fading; thus it is robust in indoor environments. The

typical coverage radius of Wi-Fi is around 10 meters. The random accessing of the channel makes it difficult for a node far away from an Access Point (AP) to access the channel. WiMAX, on the other hand, uses a scheduled scheme to access the channel, which is Scalable – OFDMA (SOFDMA)[8], and this greatly increases the coverage of the APs. WiMAX has been considered as a wireless backhaul technology for 2G, 3G, and 4G networks in both developed and poor nations[9]. Bluetooth is another widely used wireless network. Bluetooth uses frequency-hopping as the spread spectrum method. It is also operated at ISM band, with 79 channels. The data rate of Bluetooth can reach up to 3M bits/s, and the coverage for Class 1 Bluetooth can reach up to 100 meters.

1.5 Background Knowledge of TV White Space

TV broadcast spectrum at locations "where channels are not being used for authorized services, including broadcast television, broadcast auxiliary services such as wireless microphones, and private land mobile radio (primarily public safety) is often referred to as the "TV white spaces" "[10]. On June 12, 2009, most of the analog TV transmitters were turned off as a part of the transition from analog TV to digital TV. The spectrum that used for transmitting analog TV signals is now vacant, and that greatly increases the available amount of white space. An example of white space is shown in Figure 1. The areas formerly covered by the analog TV towers become white space, and qualified white space devices can transmit in this area using the spectrum for broadcast services.

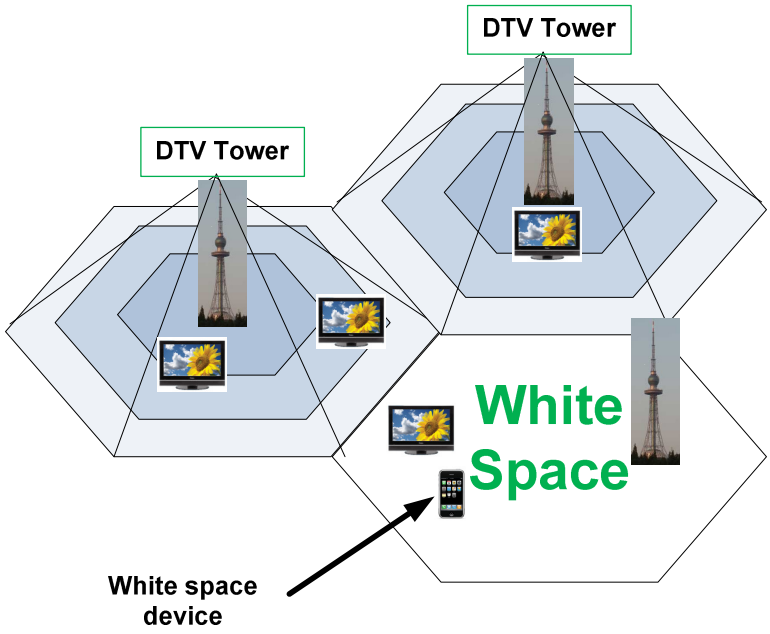

Figure 1: An example of white space

The spectral location of white space, which falls in several non-contiguous bands from 50MHz – 800MHz, determines that signals in this band can be easily transmitted a fairly long distance and can penetrate building walls, features which are attractive for broadband communications.

A group of companies called the Wireless Innovation Alliance, including Google, Microsoft, HP, Philips, Dell, EarthLink, Samsung and Intel, believes that there are huge markets and great research opportunities if white space can be opened up to unlicensed devices for broadband communications. Most broadcast service providers, represented by the National Association of Broadcasters, are not in favor of this. Their argument is that introducing unlicensed devices into the TV band might cause disruption to millions of TV receivers. After a long debate, the Federal Communications Commission (FCC) agreed that white space can be accessed by unlicensed devices that meet a set of requirements for not causing interference to TV receivers. The sensing threshold for some wireless devices is required to be as low as -107dBm as mentioned in[11].

Successfully avoiding interference to primary users is not a trivial task. Two main difficulties are spectrum sensing and the hidden terminal problem. The purpose of spectrum sensing is to detect the presence of signals in a frequency span, usually by measuring the RF energy that is there. When the signal energy is low or the signal is weak compared to the noise level, it is difficult to determine the availability of a channel. A hidden terminal is one that can be heard by

or interfere with one end of a link but not the other. In the TV white space case, a hidden terminal occurs when, for example, the devices are in the shadow of a mountain from the TV tower, which is illustrated in Figure 2. By spectrum sensing, white space device A will consider itself to be in a white space; however, it is not supposed to transmit because that will cause interference to TV receiver C. White space device B, on the other hand, is in a much better position to observe the spectrum. In this case, cooperative sensing is beneficial.

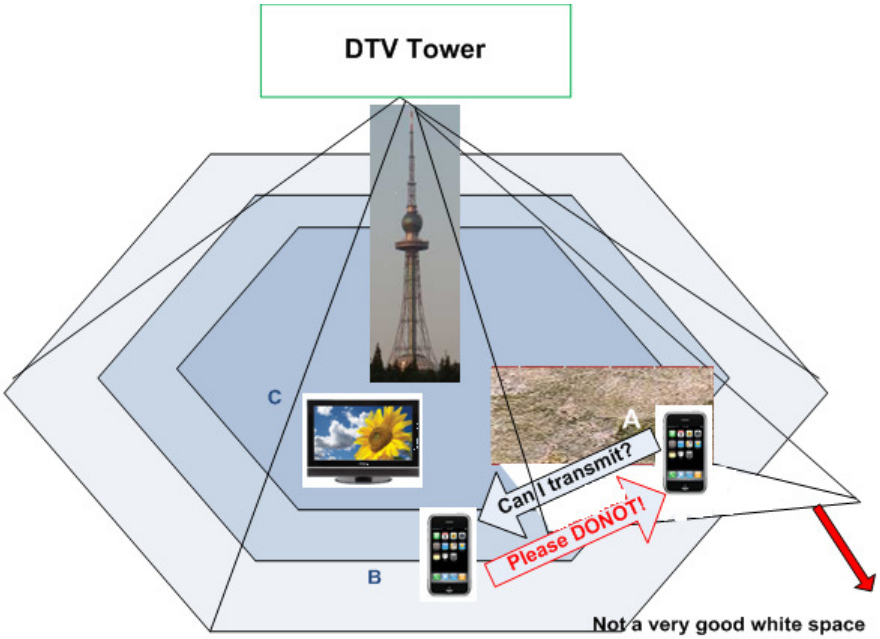

Figure 2: Example of hidden terminal

Since white space devices access the spectrum as unlicensed devices, competition among devices might cause inefficient utilization of the spectrum and also cause interference for TV receivers. Self organization among the nodes can provide a solution to set up a network that can use the spectrum more efficiently. If all the communication devices in DCCS are accessing TV white space, then DCCS serves as a network for white space devices. The cooperation of networked communication devices in DCCS greatly avoids possible interference to the TV receivers and makes the transmission safer. The self-organization feature of the DCCS network avoids vicious competition among the white space devices and uses the spectrum more efficiently. By a loose multiplexing scheme and modulation standards requirement, it avoids the cost of rebuilding new devices and accommodates as many legacy devices as possible.

The recent FCC auction of some of the TV white space (700MHz), which seeks to reallocate some of the former UHF TV channels to provide broadband communications for public safety, is consistent with national priorities focusing on homeland security and broadband and our commitment to ensure that emergency first responders have access to reliable and interoperable communications. DCCS can be implemented in the 700MHz public safety broadband in a way that will satisfy the FCC requirements, can achieve a nationwide reliable broadband system, and can be compatible with legacy public safety device.

1.6 Organization

This dissertation is organized into 8 chapters. Chapter 1 introduces the motivation of cognitive radio networks and DCCS, the challenges in cognitive radio networks, and addresses background knowledge about wireless communication standards and TV white spaces.

Chapter 2 first describes a DCCS application scenario. DCCS can cover the areas where infrastructure is not available. Subsequently, Chapter 2 discusses the architecture of DCCS. DCCS is a cellular based structure. Within a cell, the connections are centralized, and among cells, the connections are flat. Lastly, Chapter 2 gives an overview of the basic elements in a DCCS network, including Potential Picocell Cognitive Node (PPCN), Picocell Cognitive Node (PCN), and CMT (Cognitive Mobile Terminal).

Chapter 3 addresses intra-cell communications. It discussed the cell formation, which can also be understood as the initialization of the network. It also describes the communication methods within the cell. And last, Chapter 3 presents the algorithm for cell maintenance, including cell splitting and cell merging. The key technologies for intra-cell communications are channel allocation and signal classification. These two issues are presented in two separate chapters (Chapter 4 and Chapter 5).

Chapter 4 addresses the spectrum management and channel allocation algorithms for intra-cell communications. In a DCCS system, channel access has three steps: surveying the spectrum,

requesting the spectrum, and optimally allocating the spectrum. Thus, Chapter 4 covers three aspects: spectrum sensing, multi-user scheme, and channel allocation.

Chapter 5 is the description of the universal signal classification and synchronization (UCS) system. The purpose of this system is to automatically extract the features from the received signal needed for physical layer demodulation and to demodulate the signal. In a DSA scenario, this feature is important. When the transmitter changes the channel and the modulation types, it may not have an opportunity to tell the receiver, so having a receiver smart enough to by itself extract the information it needs is an attractive feature. UCS is an independent system that can be implemented on DCCS. This chapter is not integrated with the other chapters in this dissertation, but reproduces the text of a co-authored journal paper.

Chapter 6 is the description of inter-cell communication. In this chapter, three problems are discussed. Spectrum utilization is the first problem. Neighborhood discovery protocols are the second problem, and physical and MAC layer protocol for inter-cell communications is the third problem.

Chapter 7 describes the prototype and testing of DCCS. This chapter includes the prototype description, performance analysis, test cases, demonstrations, and hardware settings. The software structures, detailed flow graphs and GUIs for both Picocell Cognitive Node (PCN) and Cognitive Mobile Terminal (CMT) are described. Several important components and algorithms' performance analyses are discussed; the channel allocation algorithm, for example. The test cases for PCN and CMT functions are listed with expected results and experimental results. Hardware settings in the demonstration are also described.

Chapter 8 gives a summary of the research result of DCCS and discusses future work in continuing design and implementing DCCS.

Chapter 2 System Overview

In this chapter, we provide a brief overview of a DCCS network application, DCCS network architecture and an introduction to the important components in DCCS. We will present the scenarios that a DCCS network fits, discuss the architecture of DCCS, explain the functions of each component in a DCCS network, and characterize the basic DSA connections in DCCS.

Some of the material in this chapter is taken verbatim from my patent application[12].

Before the overview, I first would like to define the DCCS network. In the name "Dynamic Cellular Cognitive System", "cognitive" indicates that DCCS is a network built around cognitive radios. A cognitive radio can sense the environment and use the results as part of the information needed to make optimal decisions. No matter how intelligent a cognitive radio is, it can only detect the channel conditions in the specific spot where it is located. But the nature of radio wave propagation ensures that channel conditions do change spatially. Thus, cooperative sensing is not only necessary to enhance the performance of spectrum sensing, it is essential to the very concept of spectrum sensing. "Cellular" in the name of DCCS indicates the cellular structure of the network, as well as its centralized cooperative scheme in spectrum sensing within one cell.

Channel conditions vary with time and space, and the user status in the system also changes. Thus, to employ a channel efficiently, a flexible and robust network needs to be defined. "Dynamic" in DCCS indicates both the features of dynamically accessing the channel and of dynamically adjusting the network structure based on response to channel conditions that vary in time and space.

Thus, the definition of DCCS is a dynamic, cellular, cognitive network that will provide high quality cognitive communications among CR nodes with reasonable computation complexity while causing minimal interference to primary and other secondary users.

2.1 Scenario of DCCS Application

In general, this network can be used in a scenario as shown in Figure 3. In some areas where infrastructure is not available, DCCS can set up a communication network at the same time that radio nodes arrive. For example, in natural disaster areas, where the infrastructure is broken, as first responders carrying DCCS devices enter the disaster areas, the DCCS network will automatically form. This network can cover the area where communications are needed. Some of the nodes near the edge of the non-infrastructure area will have access to infrastructure. Then, the entire network can be connected to the outside world through the nodes that have the access to the infrastructure networks. DCCS in this area can support multiple types of devices, in order to accommodate more possible devices within the network. Each member of DCCS accesses the spectrum as a secondary user.

In Figure 3, nodes represented by green stars are the PCNs that have access to the infrastructure. Because of their connection to the infrastructure, the entire network is connected to infrastructure through them. As mentioned in Section 1.6, PCN stands for Picocell Cognitive Node. The word “Picocell” is borrowed from the GSM standards. There are five different types of cells defined in GSM, and picocells are the small cells whose coverage diameter is a few dozen meters. In DCCS, the coverage of a PCN is usually larger than this, but still smaller than a macro or micro cell in GSM.

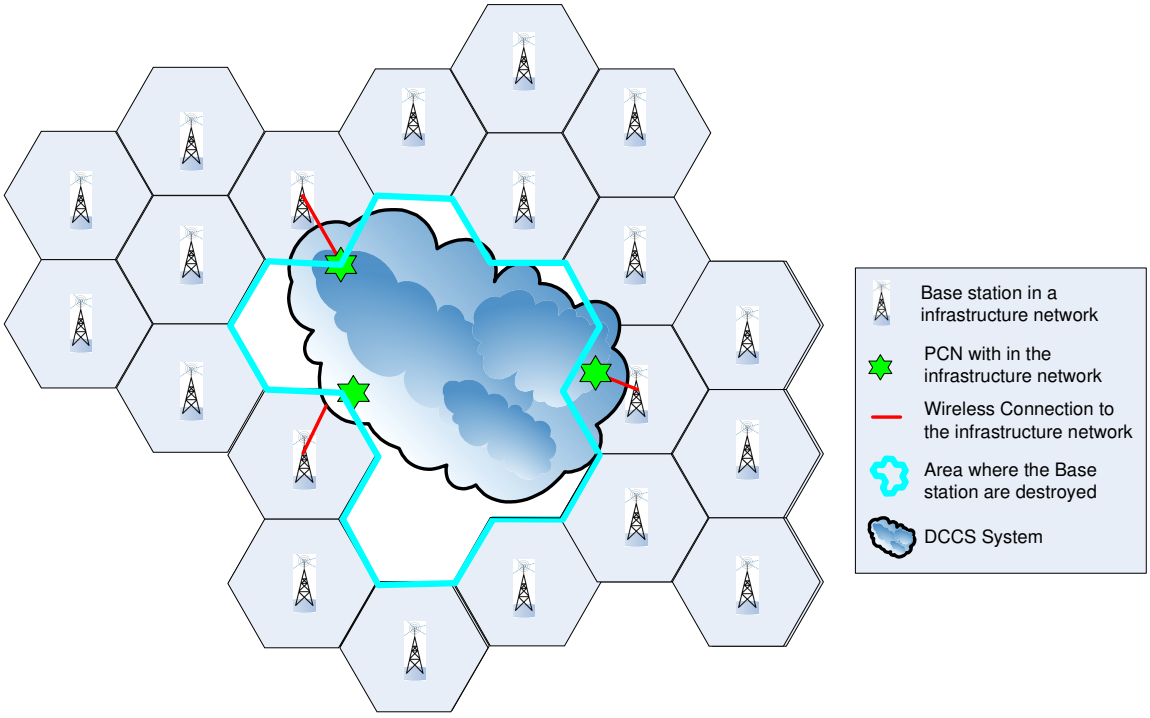

Figure 3: DCCS application scenario

2.2 DCCS Architecture

From the architecture point of view, DCCS is a mobile ad hoc network (MANET). A MANET is a special type of wireless ad hoc network defined as "a self-configuring network of mobile devices connected by any number of wireless links"[13, 14]. Each device in a MANET can move independently and potentially change the links within the network on a regular basis. [15]

There are two differences between DCCS and other MANETs. The first difference is that DCCS nodes access radio channels as secondary users. The second difference is that some of the nodes in DCCS can function as both base stations and mobile terminals. This feature gives DCCS more flexibility and adaptability.

Before I describe the detailed architecture, there are several concepts that need to be clarified.

Picocell: A Picocell is a wireless communication system typically covering a small area, for example, indoor environment or in an aircraft. A picocell in DCCS is analogous to a picocell in GSM network.

PCN: Picocell CR Node, the CR node that performs as a base station in a cell.

Cell: The coverage of a PCN.

CMT: Cognitive Mobile Terminal, CR nodes that do not perform as a PCN in the cells.

PPCN: Potential PCN, CR nodes that have the ability to become the PCN, including both those that have already switched to PCNs and the ones that stay in CMT status because there is already a PCN in the same cell.

Sub-channel: The estimated spectrum accessing range of DCCS is divided into a number of sub-channels. For purposes of this dissertation, the bandwidth of each sub-channel is 100 kHz. The total number of channel is N, and the value of N depends on the application scenario. Each channel is numbered in sequence from 0 to N-1. In the following chapters, sub-channel, channel, or sub-carrier mean essentially the same thing.

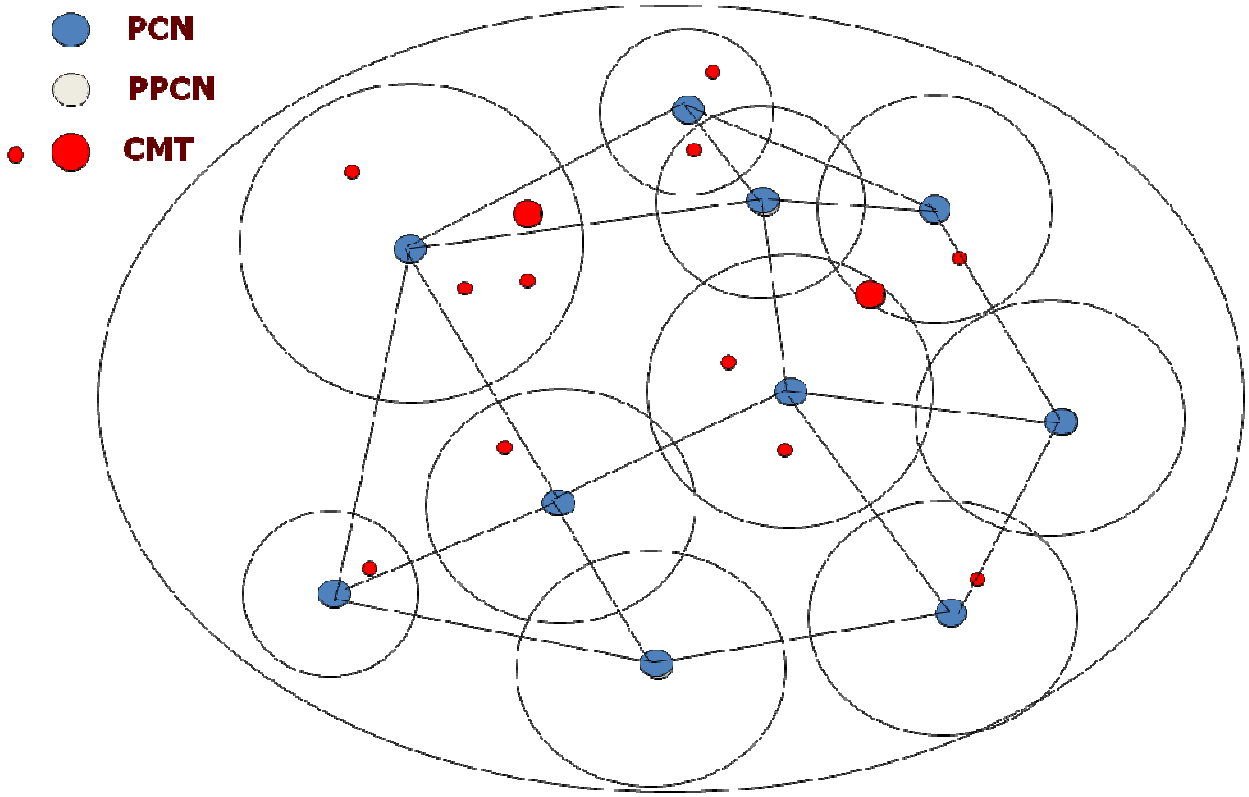

Figure 4: Concept level DCCS architecture design[12]

Figure 4 shows a concept level design of DCCS architecture. PCN serves as the base station in DCCS. CMTs within the coverage of a PCN are associated with the PCN. A PCN, its associated CMTs, and the link among them compose a cell. The communication links among PCNs serve as backhaul connections of DCCS. [12]

All PCNs are transformed from PPCNs. A CMT node can either be transformed from a PPCN or else it can only serve as a CMT because of the hardware and other restrictions. When a PCN comes to an area, it will first send out registering messages to request registering with an existed PCN as a CMT. If another existing PCN responds to the request and sends back a respond message, then this PPCN sends back an acknowledge message to confirm the registration and to become a CMT. If no existing PCN responds to the registration request messages, then this PPCN will transform to PCN status, connect to other PCNs and perform a PCN's role.

The process of exchanging command messages is a handshaking process. Because of the uncertainty of the channel availability before detection, it is necessary to have a handshake to guarantee that the command and negotiations between the two sides of communication are well understood. If started from a CMT, the three-way handshake includes request, response and confirmation. If started from a PCN, it is only a two-way handshake, including inform and confirmation. For example, if CMT A requests a channel to communicate with CMT B, the PCN needs not only to respond to the CMT A, but also to inform CMT B, and the message exchange between PCN and CMT B is a two-way handshake.

The cells of two PCNs are allowed to overlap. A soft handoff similar to that in a cellular network is performed. The term 'handoff' in cellular network refers to the process of transferring an ongoing call or data transmission from one cell to another. In DCCS, it is defined as the process of transferring an ongoing call or data session from one cell connected to the PCN to another. There are two scenarios in which a handoff is necessary:

- When the node is moving away from the area covered by one cell and entering the area covered by another cell, the call is transferred to the second cell in order to avoid call termination when the node gets outside the range of the first cell.

- When the capacity for connecting new calls of a given cell is used up, and an existing or new link, which is located in an area overlapped by another cell, is transferred to that cell in order to free-up some capacity in the first cell for other users, who can only be connected within that cell.

Handoff is defined to have occurred only when communications are transferred from one cell to another cell. The channel reallocation within a cell, for example, with the communications being interrupted and then resuming due to a primary user's return is not part of the handoff problem. Thus, handoff in DCCS can also be called inter-cell handoff. The inter-cell handoff can be used to maintain the link because the subscriber is moving out of the area covered by the source cell and entering the area of the target cell.

One of the current drawbacks for a low cost software defined radio is its latency. This drawback determines that if a hard handoff is used in DCCS, the interval between the break and make is perceptible to users. Thus, we adopted the soft handoff in the DCCS system. In DCCS, a soft handoff may keep connections to more than two PCNs at the same time. All channels will receive signals and combine them to produce the best quality signal passed to the user. Because DCCS is based in part on communication by secondary users, this process is more complicated than a normal cellular network soft handoff because the channels might be interrupted by primary users.

2.3 Important Components in DCCS

Figure 5 shows the important components to implement a DCCS network. In an area where the infrastructure of an IP network is not available, CRs collaborating in an efficient, effective manner can serve as an ad-hoc communication network. One of the advantages of a cognitive radio is its ability to perform multiple roles because of its cognition and reconfiguration ability. Thus, some cognitive radio nodes in DCCS can serve as user terminals, or they can serve as mini-base-stations. This capability is one of the foundation pieces necessary to build a flexible network architecture in DCCS. A PPCN is the cognitive radio node in DCCS that has the ability to perform multiple roles. When it is performing as a mini-base-station, it is a PCN. When it is performing as a mobile terminal, it is a CMT. In some areas, when the IP network infrastructure

is not available, DCCS creates a framework such as shown in Figure 5. When some other the elements in DCCS have the access to the infrastructure network, these elements can bridge the entire DCCS network to be able to communicate to other IP networks.[12]

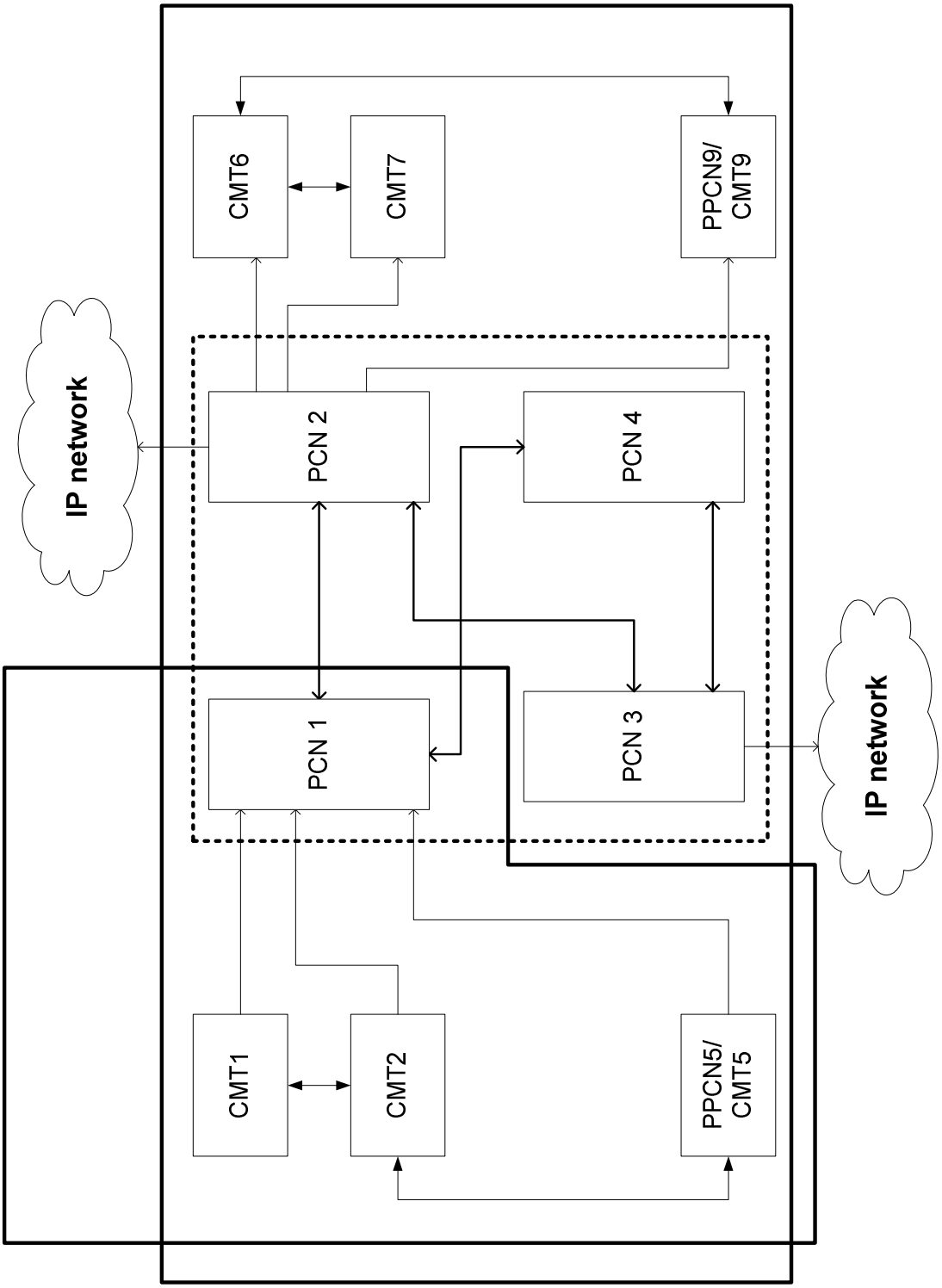

Figure 5: DCCS components and application example[12]

A PCN serves as the base station in DCCS network. The picocell concept appears in some standard communication networks. A picocell is analogous to an access point in a Wi-Fi network, and to a picocell in a GSM network. In cellular networks, picocells are typically used to extend coverage to indoor areas or some area with large traffic in data transmission. It is natural to extend the concept to apply in areas where infrastructure is damaged or not available; for example, New Orleans after Katrina or a region that has experienced an earthquake, etc. These scenarios are also the ones that have pressing needs for communications. A PCN in DCCS serves not only as a base station providing uplink and downlink for CMTs, but also as an intra-cell management and DSA controlling function. It also has the ability to function as either a base station or a mobile terminal. The intra-cell management and DSA controlling function can accommodate as many types of radios as possible, and allow a group of secondary users to access a channel without causing interference to primary users and other secondary users outside of the DCCS system. The function of switching between base station and mobile terminal can efficiently organize the network dynamically. A PPCN only changes to a PCN when a PCN is needed in its area. Because a PCN consumes more power than a CMT, in this way, power can be used efficiently and the network can last longer in some extreme situations. [12]

The size of a cell is dependent on the PCN's transmission power and current battery storage, as well as on the distributions of CMTs or other registered radios. A cell can be understood in two ways: elements based or geographically based. For elements based, it is defined as the distribution of all the nodes that register with a PCN, which means both the PCN and CMT can correctly receive and demodulate command messages and data packages. Geographically, it is defined as the area where the received signal sent from a PCN has an SNR larger than a certain threshold. Because of the mobility of the PCNs and CMTs, the former definition is more effective and timely. [12]

A PCN has the ability to identify the received signal, synchronize to the signal, and demodulate the signal automatically. These capabilities accommodate different types of radios and channel variations. The radios that a PCN can accommodate include cognitive radios, white space devices, broadband communication devices and some other legacy radios. In an emergency situation, the more types of radios that can be accommodated, the higher the probability that the

system will meet the communications needs. The configurability of cognitive radios can provide interoperability among different radios. The identification and classification function of a PCN can provide the means to identify and correctly communicate with different type of radios. A cognitive terminal can reconfigure itself to accommodate the more complicated protocols designed for DCCS.

A PCN provides the spectrum management within its cell for CMTs. All of the CMTs in the cell can be secondary users. The detailed information about spectrum management will be discussed in Section 4.3. By coordinating with a PCN, CMTs can all access the spectrum in an efficient way without causing interference with primary users. PCNs serve as a digital gateway providing interoperability among different types of radios including, but not limited to, FRS radio, public safety radio, cell phone, and broadband devices, all of which normally cannot cooperate with each other. PCNs use a power control optimization algorithm to control the power of each node within its cell to realize frequency reuse among different cells.[12]

The connections among PCNs are called inter-cell communications. Because of the bandwidth demand, Wi-Fi or WiMAX are used here. The inter-cell connections are dynamic mobile ad-hoc connections. Some of the protocols developed for ad-hoc networks apply to inter-cell connections. Examples include proactive routing, reactive routing, situation aware routing, hybrid routing or flow orientated routing. Among them, we choose situation aware routing because of two advantages: (1) Advantage depends on number of nodes activated; (2) Reaction to traffic demand depends on gradient of traffic volume. These two advantages fit the requirement of a dynamic number and distribution of PCNs. [12]

Channel allocation and power control are the two primary enabling technologies. Because of the cellular structure based network, increased capacity arises because the same radio frequency can be reused in a different cell for a completely different transmission. This technology is the same as that used in a traditional cellular telephone network. PCNs use power control to adjust the range of the transmitted signals and therefore guarantee the quality of cells' frequency reuse.

In DCCS, there are two types of communication; data transmission and command message transmission. Command message transmission is also called control message transmission or message exchange. Frequency reuse describes only the frequencies used for data transmission, not to command message transmission. A command message is for exchanging information between a PCN and a CMT or among PCNs to coordinate on spectrum utilization, routing topology and cell management. Thus, when a CMT or a PCN first joins the network, it can send out a request message to set up the connection with a PCN without knowing the prior cell frequency allocation information.

DCCS's topology adapts based on an algorithm that allows adjacent cells to merge into one. This functionality depends on the area of the cells, the number of users in the cells, and the geographic distribution of the cells (Section 3.3).

In addition to inter-cell communication, DCCS manages communication between PCNs and CMTs within a given cell, defined as intra-cell communication. Protocols governing intra-cell communication define how three-way handshaking among a PCN and CMTs can allow CMTs to access spectrum as secondary users without causing interference to primary users and users outside of the DCCS system. Three-way handshaking includes a CMT sending out a request, a PCN respond to the request and a CMT confirmation by sending an acknowledge message. A CMT is required to act under guidance from a PCN except when a CMT is allocated a channel. A CMT can cognitively set/change its modulation types according to the channel environments. A PCN allocates channels either for communication between two CMTs within the cell or for communication between a CMT in the cell and a CMT outside of the cell via PCN itself as a forwarding point. It also serves as gateway for two incompatible radios that do not directly have compatibility.[12]

When more users appear in the network, there will be more PCNs in the network, and the network will be expanded. By the bridges, a local DCCS system can connect to the IP network and reach a wider range of locations. Multiple DCCS systems can also be connected to each other with the assistance of the bridges. The existing infrastructure network in the area surrounding DCCS could be a cellular network or other type of network system with base

stations. If because of a natural disaster or for other reasons, some base stations are destroyed, the cells served by these base stations will lose communication. When the first responders arrive, there is no communication access. To solve the problems brought about by this situation, we implement DCCS. These first responders carrying DCCS devices, including PPCNs and CMTs will arrive and will be distributed to the places where they are needed. As we mentioned above, a DCCS system is set up at the same time when the responders arrive. The PCNs that are located in the edge of the disaster area then have access to the national network, they perform as the bridge, and the DCCS system performs as an island. Then, the people who are besieged in the disaster area, even those who just have a walkie-talkie, can not only communicate with each other, but also can set up a nationwide connection; their voice messages will be coded by the associated PCN and forwarded to the desired receiver.

Another example for DCCS application would be implementing DCCS in the 700 MHz public safety band. The FCC's decision to create a nationwide, broadband, interoperable public safety network in legacy wideband 767MHz-773MHz and 797MHz-803MHz can provide the bandwidth DCCS needs for the broadband intra cell communication and for the backbone connection. DCCS provides the real time set up of a reliable broadband communication system which can accommodate multiple communication devices. The FCC proposed three principles for ensuring effective public safety use of the 700MHz band: nationwide, competitive equipment market; and flexibility to meet the needs of regional communities. Implementation of DCCS in 700MHz with bridging to broadband communication would fit this situation. It would present an opportunity to put into place a regulatory framework that would ensure the availability of effective spectrum in the 700MHz band for interoperable, public safety use.

Chapter 3 Intra-cell Communication

In this chapter, we describe the behavior of a unit of DCCS --- the cell, including the formation of, communication within, and maintenance of a cell. The structure of the chapter is shown in Figure 6.

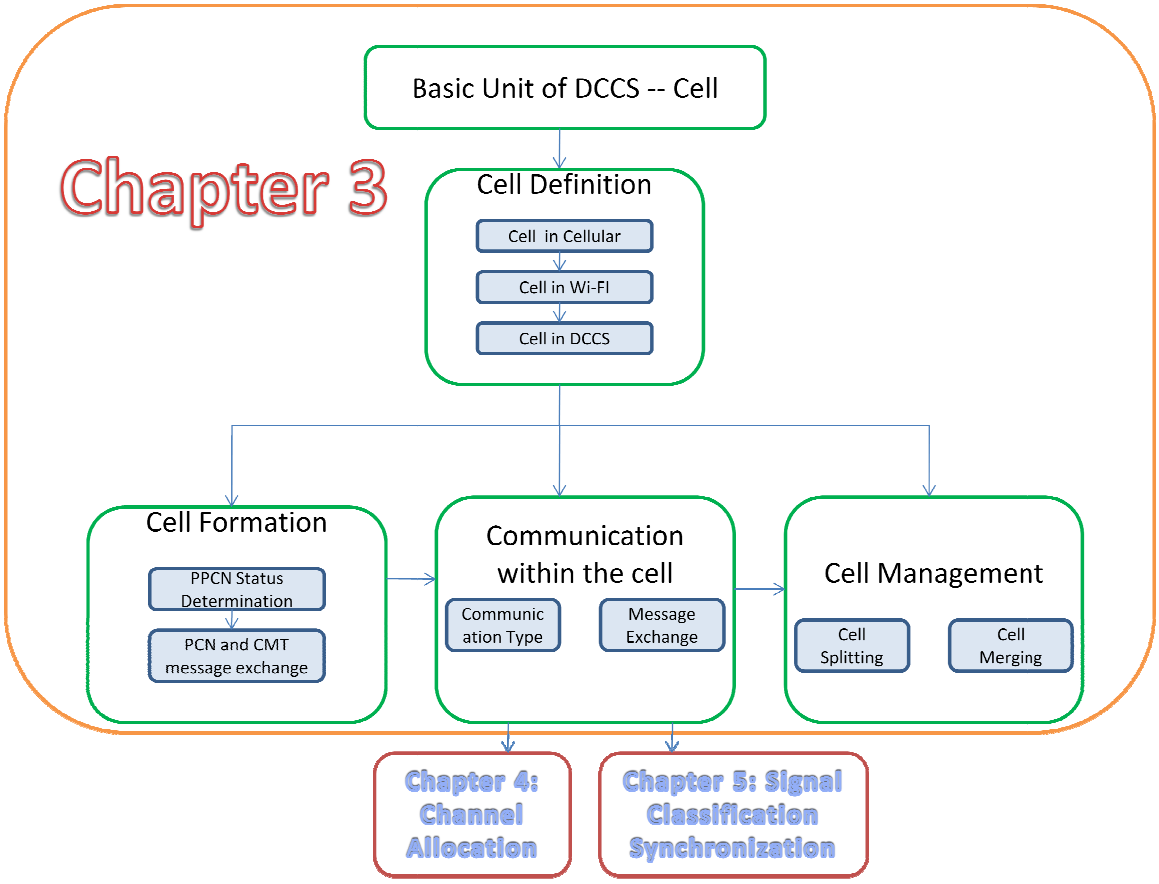

Figure 6: Structure of chapter 3

A cell in DCCS is not defined geographically but by the signal propagation environment. Thus the cell boundaries could vary both with signal strength and with ambient noise. If a communication device sets up communication with a specific PCN, it is called a Mobile Terminal (MT) of the cell. A MT could be a cognitive radio, which is called CMT, or just a communication device. All of the MTs and together with the PCN constitute the cell. Each effective PCN has an associated cell. A cell does not necessary contain any MTs. MTs and the

PCN of a cell are called components of a cell. Each of the current terminals registered with a PCN is associated with the PCN, and is called the PCN's MT. The PCN is called a MT's PCN if the MT is associated with it.

In a standard cellular network, a cell is defined as a unit to govern frequency reuse. The frequency settings inside the cell are fixed by pre-allocation, and the successful reuse of the frequency depends on the propagation attenuation. Thus a cell in a standard cellular network is basically defined geographically with some adjustment for user influence; for example, the breathing of a cell [16]. A cell is defined as the range where the signal after propagation attenuation is higher than a certain threshold. Depending on the base station antenna, the shape of the cell can be round or sectorial. The bandwidth allocated to a terminal within the cell is predefined and fixed.

In an IEEE 802.11, e.g. Wi-Fi, network, a cell is limited to a much smaller size. The size of the cell also impacts the speed of the network, varying from 11Mbps to 54Mbps [17]. Because of a CSMA-CA based MAC layer design, the distance between the terminal and the access point has important effects on the performance. This random accessing feature determines the relatively short radius of a Wi-Fi cell compared to that of a standard cellular network. IEEE802.16 (WiMAX) combines these two and provides a city wide wideband network.

In both of these cases, the base station, or the access point, is not moving. The structure is relatively stable and it is possible to define the cell geographically. However, in DCCS, PPCNs are allowed to move, and the conversion from PCN to CMT is performed when necessary. Thus the cell size in DCCS is connection oriented. A cell is formed because it is needed in the network, and the range of a cell is variable depending on the location, number of users, and number of available channels, mobility, power, and path loss. Thus it is necessary to discuss both the cell formation and channel maintenance.

In cell formation, the initiation process of a cell and the components that constitute a cell will be described. The components of the cell are responsible for different aspects of the intra-cell communications and inter-cell communications. The dynamic structure of a cell allows new MTs

to join at any time. The update information for both topology and spectrum re-allocation must be distributed among each of the components in an efficient and effective way.

The diversity of the radio devices accommodated by the DCCS system determines that special design is required for both radios and protocols to provide interoperability and organization. Intra-cell communication includes the regulation and compatibility of communication signals, the devices' behaviors, as well as spectrum management. In this chapter, we will describe the protocols and design requirements for radios. Two of the key issues, spectrum management and signal classification, will be addressed in Chapter 4 and Chapter 5

As a system that can be used in a wide range of situations, especially public safety, one need of DCCS is to expect the unexpected. Thus, having a robust maintenance scheme is important. When the environment changes, or a dramatic part of the cell components changes, we need to consider how to maintain the communication. In Section 3.3, we will discuss how to combine and split the cell in order to keep the communication when the above situations take place.

3.1 Cell formation

In this section, we focus on solving three problems to understand the process for a cognitive radio to join DCCS under all possible conditions.

1. How does a PPCN determine its status when first joining an existing network?
2. How does a PCN behave to perform its responsibility for cell formation?
3. How does a CMT register with its PCN?

The definition for a node's "joining a network" in this chapter specifically means the process for the stated node transferring its status from being outside of the network to being part the network.

3.1.1 PPCN Status Determination

A PPCN is a node that can perform both as a base station or a mobile terminal. The status of a PPCN depends on the neighboring nodes' behavior. The objective is to use the resources in an efficient way. Thus for a PPCN to join the DCCS network, the principle is that if a PPCN can

detect other PCNs' existence and fall into the cell range of the discovered PCN, its joining status is CMT; otherwise, its joining status is PCN.

The simplest way to determine the existence of a reachable PCN is to send out a joining message and wait for a possible response. Suppose node A is a PPCN ready to join a DCCS network. The message it sends out should include message type, identification number of node A, and position of node A if GPS is applicable. The format of the message is shown in Table 1.

Table 1: PPCN status determination message

DST-ID	SRC-ID	MSG-TYPE	MSG-STATUS	RX-ID	Ch-Num
FFFF	CMT#	RQST	PPCN	FFFF	#########(462662600)
CMT#	PCN#	RSPD	PPCN	CMT#	#########
PCN#	CMT#	ACKW	PPCN	PCN#	#########

This message can be seen as a type of beacon message. The propagation of the message power is shown in Figure 7. In this example, the modulation of this message is dbpsk, symbol rate is 100k, power is 0dBm and squared root raised cosine pulse shaping is used. The detailed parameters can be adjusted in the application. The listed parameters are used in the prototype which will be described in Chapter 7. The received power of this message will decrease while the distance between this PPCN and an existed PCN increases. If a PCN is within the useful communications range, then, the PPCN will stay as CMT and registered with the PCN. Otherwise, it will become a PCN itself.

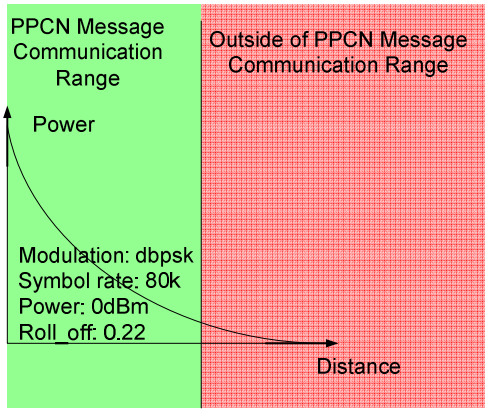

Figure 7: PPCN request message propagation

Because channel occupancy is secondary user based, the scheme of accessing the channel is DSA based CSMA. The entire possible available channel band is divided into multiple sub-channels. Node A automatically chooses a sub-channel, senses the energy in this sub-channel. If no energy is detected, repeated request message will be sent out periodically until energy is detected or response message is received or a certain time limit is reached. If energy is detected before a response message is received, node A will change to another randomly chosen channel and start the process over again. If a response message is received, node A sends back an acknowledge message indicating that it is associating with the PCN and node A stays as a CMT. If multiple response messages are received, only the strongest received response message will be considered. For PCN B, the strategy is to go around all the sub-channels periodically. In the circulation process, if a message is received, then PCN B will respond to this message and finish the three-way handshaking. In the prototype, PCN B stays in each channel for 50 ms, and 10 sub-channels are used. The total circulation time for PCN B is 0.5 s. Figure 8 shows the three-way hand shaking for PPCN status determination if a PCN exists in the appropriate communication range.

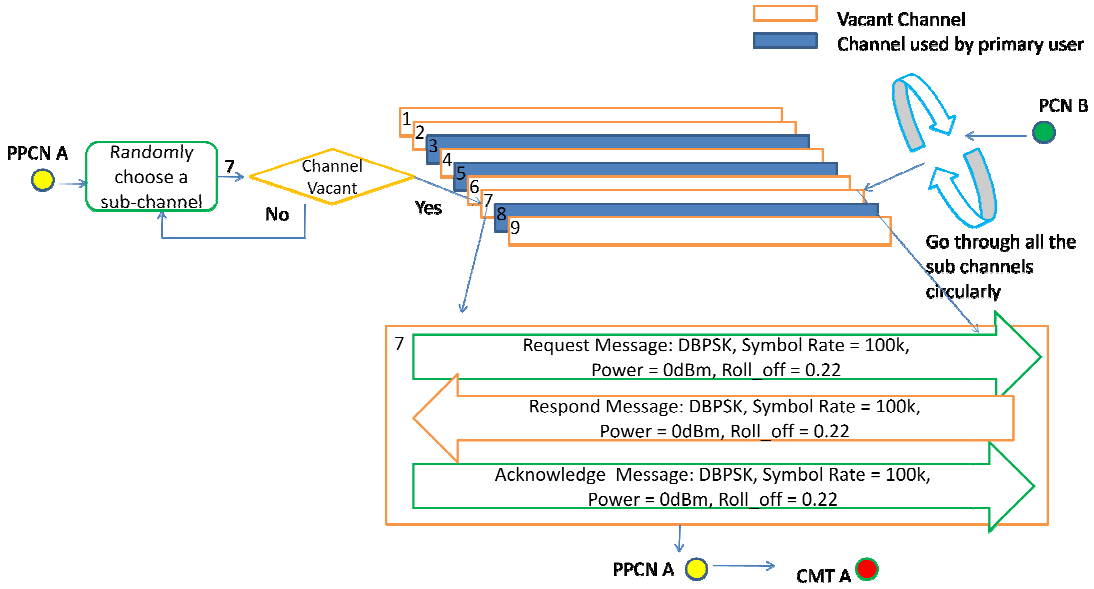

Figure 8: PPCN status determination

If for a period $\tau = 1$s, no respond message is successfully received, it is assumed that no PCNs are within the cell range and node A switches its status to a PCN. The duration of the period $\tau = 1s$ is chosen in our DCCS prototype. Other values can be chosen based on application

scenarios. As a PCN, node A accesses the spectrum for inter-cell communications and sends out messages at a higher power looking for connections with neighboring PCN nodes. Meanwhile it begins to process and manage intra-cell communication. The details of connections with other PCNs will be introduced in Chapter 6 and the management of intra-cell communication will be described in Chapter 4. The experimental result will be described in Chapter 5 and Chapter 7.

3.1.2 PCN and CMT Interaction in Cell Formation

For an existing PCN, its role in cell forming is to register those CMTs which request permission to join the network. The interaction between them is done through three-way hand-shaking message exchanges. The format of the three-way hand shaking messages is shown in Table 2.

Table 2: Request message three-way hand-shaking

DST-ID	SRC-ID	MSG-TYPE	MSG-STATUS	RX-ID	Ch-Num
FFFF	CMT#	RQST	REGI	FFFF	#########(462662600)
CMT#	PCN#	RSPD	REGI	CMT#	#########
PCN#	CMT#	ACKW	REGI	PCN#	#########

The first column represents the identification of the expected receivers. It can be an IP address, or a predefined ID number of the devices. “FFFF” means that no intended receivers are specified, and any PCN that receives the message should send out the respond message. The second column represents the identification of the transmitter. The third column shows the steps in the three-way hand shaking processes. RQST is for request message, RSPD is for respond, and ACKW is the acknowledgement for confirmation. The fourth column is the function of the message; REGI means it is used for PCN to register a CMT. The fifth column represents the ID of the intended receiver and it is not used in the registration period. This column is not used in the cell forming stage. The last column indicates the center frequency to transmit this message. This information seems redundant here because, if the receiver can correctly receive a message, it should have known the center frequency. However, as we will show in the implementation part, because of the inaccuracy of the oscillator in a USRP, there is always a frequency offset between the desired frequency and the real frequency. Having this information an automatic calibration among devices is easier to accomplish.

3.2 Cell Communication

The types of modulation types for intra-cell communication are described in Table 3.

Table 3: Modulations used for intra-cell communication

Cognitive Radios	**Legacy Radios**			
	Public safety Radio	FRS radio	GSM or CDMA (other types of cellular radio)	Wi-Fi devices
MPSK, QAM, FM, AM,FSK etc	FM, FSK	FM	Predefined modulation types, GMSK etc based.	Predefined modulation types, OFDM based

Any node with data to communicate must first to go through the command communication with the associated PCN to request a channel. If both sides of the communication are within one cell, the PCN will allocate the channel, and the two sides of the communication will stay in the allocated channel and choose their preferred modulation format. The channel allocation scheme will be introduced in Section 4.3. If the two sides of the communication are in different cells, a channel will still be allocated, and the PCN will serve as the gateway to forward the data. Intra-cell communications between CMTs work cognitively, which means that the CMTs calculate the optimal resource utilization and modulation scheme. Intra-cell communication includes the communication between two CMTs within the cell, and between a CMT and the PCN in the cell. The communication could be narrow band modulations including MPSK, FM, AM, FSK, or it could be OFDM based wideband communication. Each CMT chooses the mode that best fits the current environment and individual transmission requirements. The inter-cell communications between CMTs in different cells are coordinated through PCNs. PCN's execute inter-cell communication using wideband transmission schemes such as Wi-Fi or WiMAX. Wi-Fi and WiMAX or other broadband communications are preferred because the sums of inter-cell communication payloads are relatively large. Among them, WiMAX is preferred because it covers longer distance with a scheduled MAC layer principle.

One responsibility of a PCN is to manage spectrum and optimally allocate channels. In Section 4.3, the detailed description of the algorithm will be shown. Every time CMT needs to transmit data, it needs to request a channel from PCN. Based on the data records of the channels over time, PCN will allocate the channel.

A PCN as a powerful CR node has the ability to classify signals and perform synchronization; thus, it can accommodate multiple modulation types. In Chapter 5, signal universal classification and synchronization (UCS) will be described. During the communication, if one side of the communication needs to change the modulation type without informing the other side of the new setting, the receiver side will lose tracking of the signal. Then, UCS will be automatically be launched and locate, identify, synchronize and demodulate the signal.

A more important application for UCS in DCCS is the auto-interoperability. DCCS is designed to accommodate different types of devices. In some emergency situations, more types of accommodated devices mean more chances to obtain communication.

In Figure 9, the function of spectrum management and UCS in PCNs is shown. Both of them are key technologies of PCNs. Detailed descriptions of these two systems appear in the next two chapters.

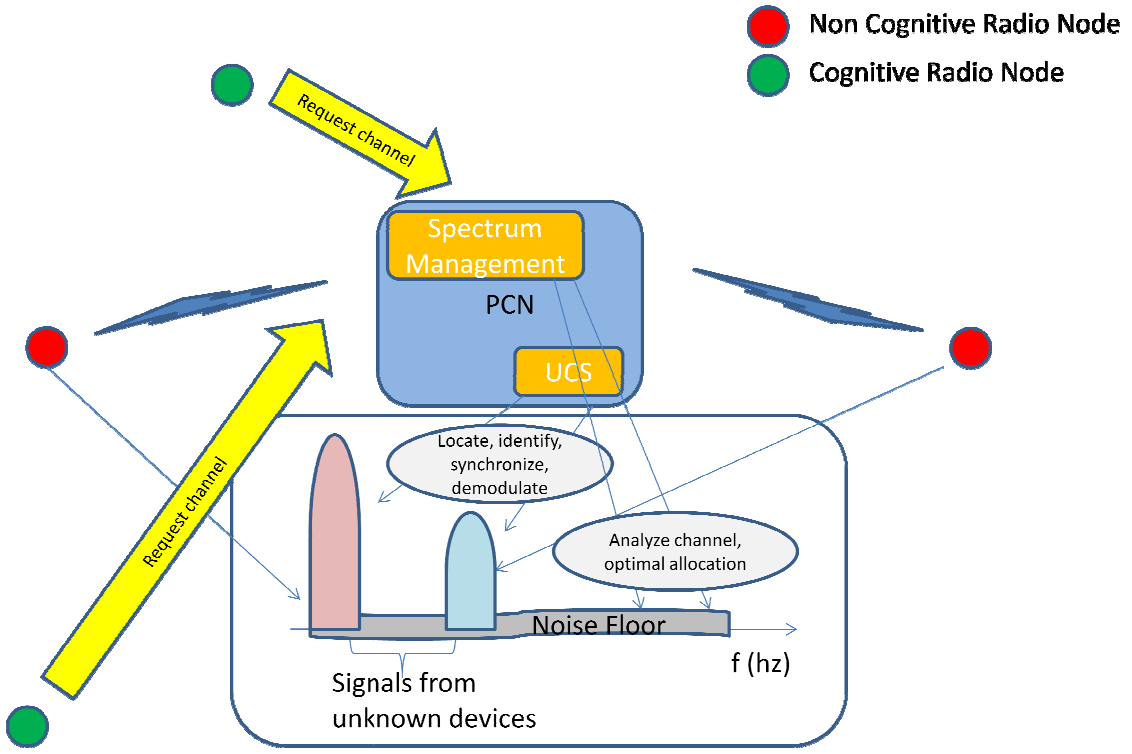

Figure 9: Functions of spectrum management and UCS in PCN

PCN is the manager of intra-cell communications. The management or control messages format is the same as it is in the cell forming stage and it is shown in Table 4.

Table 4: Intra- cell communication control message

	DST-ID	SRC-ID	MSG-TYPE	MSG-STATUS	RX-ID	Ch-Num
CHAN (request channel)	PCN#	CMT#	RQST	CHAN	CMT#	FFFFFFFFF
	CMT#	PCN#	RSPD	CHAN	CMT#	#########
	CMT#	PCN#	RSPP	CHAN	CMT#	#########
	PCN#	CMT#	ACKW	CHAN	CMT#	#########
	PCN#	CMT#	ACKP	CHAN	CMT#	#########
RESU (resume communication)	PCN#	CMT#	RQST	RESU	CMT#	FFFFFFFFF
	CMT#	PCN#	RSPD	RESU	CMT#	#########
	PCN#	CMT#	ACKW	RESU	CMT#	#########
TERM(terminate communication)	FFFF	CMT#	RQST	TERM	CMT#	#########
	CMT#	PCN#	RSPD	TERM	CMT#	#########
	PCN#	CMT#	ACKW	TERM	CMT#	#########

Including those in Table 1, there are three types of messages and three stages for each of them except for the channel request message which has five stages. The four types of messages include registering messages, channel request messages, communication resume messages, and communication termination messages. Three stages represent the three way handshaking process including request, respond and acknowledgement. For a channel request message, because the PCN needs to respond to both the proactive and passive radios with the allocated channel information, and both of them need to acknowledge receiving the respond message, there are five stages.

3.3 Cell Splitting and Merging

Based on the mobility and dynamics of the entire system, it is often required to split or merge cells to make the system more efficient. Unlike other standard cellular systems, the cell range is changing and the PCN of a cell is able to move around, as are the mobile terminals associating with the PCN. The issues in terms of cell splitting and merging are mainly two: the conditions and schemes of splitting and merging, and handoff processing during and after the splitting and merging. In this section, we will discuss these two aspects separately.

3.3.1 Conditions and Schemes of Splitting and Merging

Cell splitting and merging is made necessary by the dynamic nature of DCCS. The cells serving as base stations can move, and when a base station is overloaded another PPCN can assume the base station function, picking up part of the load and splitting the cell. The situation is different from that in conventional MANETs.

In conventional MANETs, each node can access the available spectrum and are provided stable bandwidth to communicate with each other[15]. The major difficulties in conventional MANET research include dynamically changing topology and the lack of structure; together, theseIP subnetting impossible[18]. The idea of cluster based MANETs has been appeared in [19, 20], and some other literature. Conventional MANETs focus on the use of clusters to reduce the updating overhead during topology changes or in routing process [21, 22]. The head of the clusters in these network serves as a similar role as PCN in DCCS.

In DCCS, all the nodes in the network access the spectrum as secondary users and entire network operates in a DSA scenario. This determines that a PCN has different responsibilities from the head of a cluster in a conventional MANET. A PCN needs to be aware of the spectrum status to avoid interference with the primary users. A PCN needs to be able to receive the messages from its associated CMTs in a DSA scenario to coordinate the network. It also needs to be aware of the spectrum location of its associated CMTs, and able to forward data for them. Because of hardware restrictions, the workload that a PCN can afford is limited. The conventional MANET protocols cannot meet the requirement for DCCS network. For that reason the following model is derived for DCCS.

We set up a model to describe the conditions and schemes. This model is called Cell Splitting and Merging Model (CSM). The objective of this model is to determine whether splitting or merging should be adopted, based on several principles. The model is shown in Figure 10. The yellow circle is the coverage for PCN1, and the blue circle is the coverage for PCN2. The green part is the overlap for these two coverage areas. The principles are listed as follows:

1. If the connections of a PCN with its associated mobile terminals exceed a certain number, splitting is necessary.
2. If the distance between a mobile terminal and a PCN exceeds a certain distance, and the mobile terminal cannot be re-registered with another existing PCN, a split is necessary.
3. A CMT not being covered by a PCN is considered as bad as not being able to be allocated a channel when needed.
4. A parameter called *bad influence factor* is used to determine whether a split or a merge operation should be adopted.
5. Global optimization is not considered in CSM, which means that the influence on other cells by the operation in the system is not considered.

Based on the above principles, the model can be described mathematically as follows.

Parameter Definition:

n_{pcn1}: Number of mobile terminals associate with PCN1

n_{pcn2}: Number of mobile terminals associate with PCN2

n_{max1}: Maximum numbers of mobile terminals associated with PCN1

n_{max2}: Maximum numbers of mobile terminals associated with PCN2

d_{max1}: Maximum distance between a CMT associated with PCN1 and PCN1

d_{max2}: Maximum distance between a CMT associated with PCN2 and PCN2

d_{12}: Distance between node 1 and node 2

$P_{mt}(n_s, n_{pcn1}, n_{pcn2})$: The probability that a new MT cannot be registered before the operation while it can after the operation

$P_{ch}(C_s, n_{pcn1}, n_{pcn2})$: The probability that a CMT can be allocated a channel before the operation while it can't be allocated a channel after the operation

p_r: The probability that a new MT arrives

p_c: The probability that an associated MT requests a channel

c_s: The available number of channels at this moment

Then, the model can be described as:

$$\begin{cases} H_1: f(n_{pcn1}, n_{pcn2}) \geq 0 \\ H_2: f(n_{pcn1}, n_{pcn2}) < 0 \end{cases}$$

$$s.t.: n_{pcn1} < n_{max1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$n_{pcn2} < n_{max2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

$$d_{12} < d_{max1} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

$$f(n_{pcn1}, n_{pcn2}) = p_r * P_{mt}(d_{12}, d_{max1}, d_{max2}) + P_{ch}(p_c, C_s, n_{pcn1}, n_{pcn2})$$

$$P_{mt}(d_{12}, d_{max1}, d_{max2}) = \frac{{d_{max1}}^2}{\left(\frac{\alpha}{2} * d_{max2}^2 - \cos\left(\frac{\alpha}{2}\right) * \sin\left(\frac{\alpha}{2}\right) * d_{max2}^2 - \beta * d_{max1}^2 + \cos(\beta) * \sin(\beta) * d_{max1}^2\right)}$$

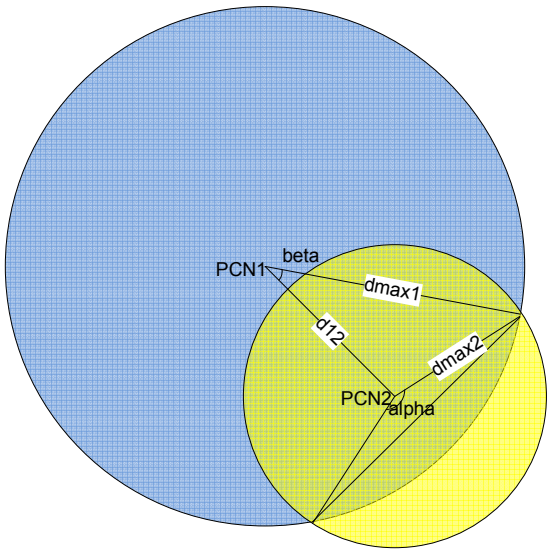

Figure 10: Cell splitting and merging shape comparison

Where, as indicated in Figure 10,

$$\alpha = 2\pi - 2\arg\cos\left(\frac{d_{12}^2 + d_{max2}^2 - d_{max1}^2}{2d_{12}d_{max2}}\right)$$

$$\beta = \arg\cos\left(\frac{d_{12}^2 + d_{max1}^2 - d_{max2}^2}{2d_{12}d_{max1}}\right)$$

$$P_{ch}(p_c, C_s, n_{pcn1}, n_{pcn2})$$

$$= \sum_{k=C_s+1}^{n_{pcn1}+n_{pcn2}} \frac{\left((n_{pcn1} + n_{pcn2}) * p_c\right)^k}{k!} e^{-(n_{pcn1}+n_{pcn2})*p_c} - (1$$

$$- \sum_{k=0}^{C_s \frac{n_{pcn1}}{n_{pcn1}+n_{pcn2}}} \frac{(n_{pcn1} * p_c)^k}{k!} e^{-n_{pcn1}*p_c} \sum_{k=0}^{C_s \frac{n_{pcn2}}{n_{pcn1}+n_{pcn2}}} \frac{(n_{pcn2} * p_c)^k}{k!} e^{-n_{pcn2}*p_c})$$

If H_1 is true, then, Node 1 and Node2 should be independent, which means if they are both PCNs, then they should keep this status and their cells do not merge; and if Node 2 associated with Node 1, their cells should be spitted.

If H_2 is true, then Node 1 and Node 2 should be associated, which means if CMT2 is associated with PCN2, then they should keep the status and their cells should not split; and if they are both PCNs, their cells should merge to be associated.

The above process is only activated when any one of inequalities (1), (2), (3) are violated.

In this model, the distance variables $d_{12}, d_{max1}, d_{max2}$ are not measured as the absolute value, but the relative value. A free space path loss is assumed. The transmit power of PCN1 is in proportion to d_{max1}, the transmit power of PCN2 is in proportion to d_{max2} and the power of PCN1 minus the received power of PCN1 by PCN2 is in proportion to d_{12}.

3.3.2 Handoff for Cell Splitting and Merging

In a DCCS system, handoff doesn't only include ongoing call or data sessions, but it also includes any registered mobile terminals transferring from their current associated PCN to another PCN. The change doesn't include only channel change, but also associated PCN change. In this section, we are going to discuss the handoff process caused by cell splitting and merging. When cell splitting or merging occurs, a handoff that needs to be processed falls into either of two categories: (1) ongoing communication and (2) transferring the information of the CMTs that registered but currently are not actively communicating. For ongoing communications, there are two types of handoff: soft handoff and hard handoff. The advantage of a hard handoff is that at any time, one call only uses one channel, and the advantage of a soft handoff is that the broken of a link happened after the establishing of a new link. Both the current and targeted PCN are aware of the occurrence of the splitting and merging, thus there are sufficient reasons to choose soft handoff.

A special case for ongoing communication handoff processing is for the communication between two MTs within the same cell. Because in this case, the channel is allocated to the two MTs and they are communicating directly. However, after splitting, there is possibility that they are in different cells. And according to the basic channel allocation principle, direct communication

between them needs to be relayed by PCNs. We made an exception in this case; the direct communication between these two will be continued until the termination of the session. The reason is to reduce the frequency of handoffs.

For registered node, the registering information will be transferred to the targeted PCN from the current PCN by message exchange.

3.4 PCN Design

Figure 11 is the block diagram for a PPCN implemented using software-defined-radio-based architecture. “A series of CR nodes functioning as PPCNs begin to sense the surrounding area by sending out a registering request. If the received signal power of respond message from a PCN is lower than a predetermined threshold or no response message is received, it will assume there is no PCN available, and become a PCN itself. Additionally for the CMT, if the PCN which it is registering with is not available because of some unpredictable event, it will switch to a PCN. Once a PCN is initialized within hardware and software, it connects with all other PCNs in the DCCS network and begins updating and broadcasting its routing table.

For PCN intra-cell management, it is PCN’s responsibility to coordinate the transmission within its cell, allocate channel, and serve as gateway for radios that are not cognitive radios. Spectrum information is gathering via energy detection by multiple nodes. A CMT randomly chooses a channel, and uses energy detection to determine the existence of signal in the channel. Based on CSMA, if no signal is present in the channel, it can send out a command message. The PCN listens to all channels sequentially. It processes and responds to the message it receives and collects the channel information meanwhile. Based on the channel information it collects, it allocate the current optimal channel. The method to determine which channel is optimal is shown in Section 4.3. The CMTs are then informed of the results by the PCNs so that the CMTs do not experience intra-cell interference and adjacent-cell interference. Additional functionality for

intra-cell management includes a MAC protocol and gateway and data forwarding to CMTs within the cell. [12]

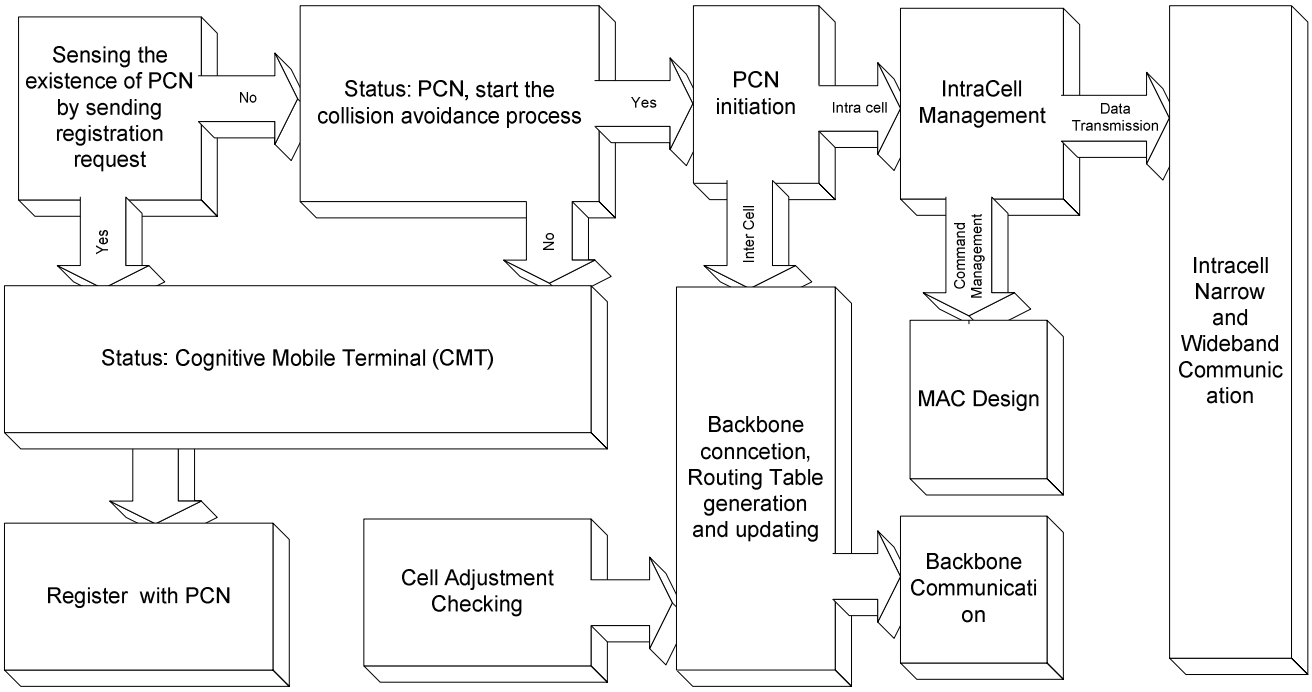

Figure 11: PCN implementation block[12]

3.5 CMT Design

Figure 12 is the block diagram for a CMT. It is necessary for a CMT to follow the protocol described in Fig 3 when the PPCN switches to CMT status. A CMT can also be an individual node which does not satisfy the hardware requirement of a PPCN and can only serve as a CMT. A legacy radio cannot serve as CMT because it is not a cognitive radio, and it cannot reconfigure to adopt the protocols. Many legacy radios can be connected to DCCS system because of the gateway function of a PCN. Based on carrier sensed multiple access (CSMA), a CMT assigns

itself a sub-channel in order to register itself within the cell in accordance with MAC protocols. Once the CMT is registered in a cell, it can be allocated a channel in one of two ways: (i) a CMT is a intended receiver from another CMT; (ii) a CMT itself intends to communicate with another CMT and requests to be allocated an available channel. Once the PCN allocates a channel following the CMT's request, the CMT switches to the allocated channel. A CMT will continue to communicate on the allocated channel until it is interrupted by a number of possible events including but not limited to: (i) a primary user requests the channel, or (ii) another CMT makes a request, or (iii) the CMT travels outside the cell coverage area. When interrupted, the CMT can request to resume communication or inform the PCN it is ready to terminate the connection when finished with its communication.[12]

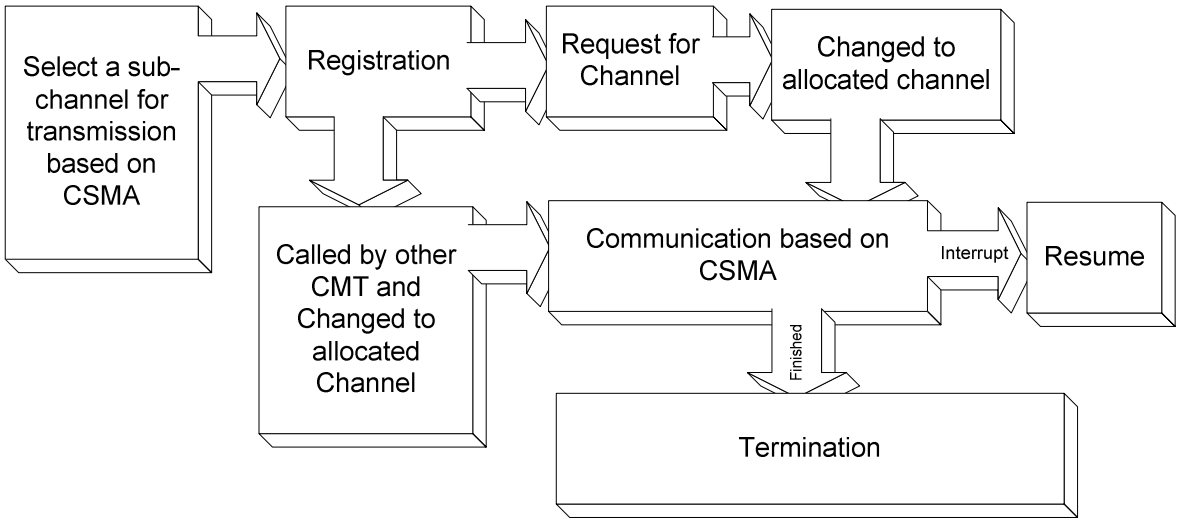

Figure 12: CMT implementation block[12]

Chapter 4 Spectrum Management and Channel Allocation

Spectrum management and channel allocation are important for the DCCS system, because DSA is one of the essential features of DCCS. In DCCS, all the channel access in the system is secondary user based.

In this chapter, we are going to discuss two major problems. One is a multiple-user DSA scheme, which is designed to be the DCCS MAC layer. For DSA, secondary users access a vacant channel when primary users are not using it. For a multiple secondary user coexistence scenario, a well designed MAC layer scheme that can avoid interference to both primary users and other secondary users and an optimal channel allocation scheme are important. The other is the channel allocation algorithm, which is the algorithm to determine the channel which has the least probability that primary users will appear during a certain period based on statistical data about channels collected while sensing the spectrum. By implementing the algorithm, interruptions by primary users are greatly reduced, and the communication quality is higher. These two parts provide better channel resource utilization and user satisfaction in DCCS.

4.1 Spectrum Sensing

Spectrum sensing is a necessary part of DSA. We will first discuss spectrum sensing technology and challenges before describing the solutions to the two problems in DCCS mentioned above. According to [23], The current spectrum sensing technology can be classified into one of the three categories: transmitter detection; cooperative detection; and interference based detection.

Transmitter detection is a non-cooperative detection, and focuses on letting cognitive radio distinguish between used or unused channels based on local observation. Algorithms for

transmitter detection mainly have three types: matched filter, energy detection and cyclostationary feature detection. Matched filters are usually adopted when the primary user signal is known to the secondary user. By using a matched filter, the receiver can maximize the received signal's SNR. Energy detection is more suitable for the condition that the secondary user cannot gather sufficient information about the primary user. It simply determines the existence of a primary user by comparing the received signal power with a defined threshold. Cyclostationary feature detection can be seen as advanced energy detection. The secondary user still doesn't have enough information about the primary user to use a matched filter. Nevertheless, it extracts some information by spectral correlation analysis based on the fact that modulated signals are, in general, coupled with sine wave carriers, pulse trains, hopping sequences, or cyclic prefixes, which result in built-in periodicity[24, 25]. Cyclostationary feature detection usually requires higher complexity and a longer signal capture duration.

There are two major issues about transmitter detection. One is the assumption on which it is based: the primary user is actively transmitting. "If the primary user is in the receiving status, the inference can't be avoided due to the lack of primary receivers information"[25]. The hidden terminal problem is another issue. Thus, cooperation for more spectrum sensing information from other nodes is necessary. This is the motivation for cooperative detection. Cooperative detection structure can be either centralized or distributed. Stochastically, multiple detectors have a smaller probability of making a mistake than a single detector[24]; furthermore, multipath fading and shadow effects are greatly reduced by cooperative detection[26]. However, the increased overhead and traffic are the problem for cooperative detection, and the first issue of transmitter detection is still not fully solved.

The FCC proposes the interference detection method [27]. This method allows a secondary user to transmit under the interference temperature limit for primary users, using the band that has a noise floor lower than the interference temperature limit. This is illustrated in Figure 13. This method avoids interference to the primary users. However, it is difficult for secondary users to detect each other, and this will cause interference among secondary users. Thus, it cannot be used in a multi-secondary-user scenario.

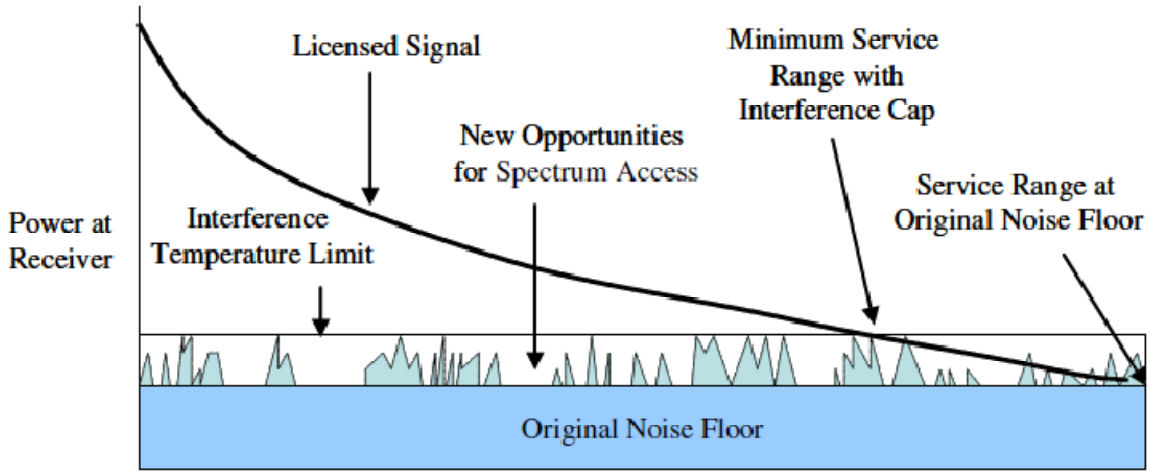

Figure 13: Interference detection model [27]

After all, the motivation or the application of spectrum sensing is opportunistic, because the primary users' behavior is unpredictable. The fundamental limits and practical challenges have been thoroughly discussed in [28]. Three principles should be realized in order to build a spectrum sensing system with acceptable reliability. The first is that "opportunistic use can be high power with long range, as long as the power density is controlled appropriately [28] ". This principle indicates an important concept in spectrum sensing and dynamic spectrum access, power density. Wireless signal propagation has an important feature; fading. If a spectrum sensing result is to be shared, or a secondary user access is to be applied, fading has to be considered in the analysis. The second principle is that "Opportunistic use requires a system, not a device, to deal with fading". The importance of this principle is that it stresses the system level application, instead of a single node behavior. Many systems can measure the spectrum sensitively, even work in really low SNR by an accurate device or a complex algorithm. However, in order to put the system into real wireless communications, the fading and spectrum variation with location has to be considered and thus requires a system level cooperation. The third principle is not only that secondary-to-primary interference needs to be considered, but also secondary-to -econdary interference must be considered. Unlike a primary user to which channel is pre-allocated, secondary users are not aware of each others' existence without system level cooperation. The interference or competition among secondary users doesn't bring any winner, but instead, all are losers because potentially none of them will be able to successfully

communicate. Thus, avoiding secondary-to-secondary interference is just as important as secondary-to-primary interference.

Considering all above issues, we developed our own spectrum accessing scheme in DCCS. It can be categorized as a combination of transmitter detection and cooperative detection, and it deals with primary user detection, fading, and interference at the system level. It doesn't only detect and avoid primary users, but also other secondary users. For other secondary users in the system which obey the protocols, the cooperative information will benefit their communication and for the secondary users outside of DCCS, we treat them as primary users and avoid inference with them. The details appear in Chapter 3.

4.2 Multi-secondary-user DSA Scheme

Each cognitive radio node in DCCS is able to perform transmitter detection. Transmitter detection in DCCS is different, depending on whether the node is a PCN or a CMT. One assumption made is that a PCN is able to perform more complicated tasks although it doesn't have to, and a CMT is assigned relatively simple tasks. The consideration of choosing which transmitter detection method to use includes the complexity, the hardware constraints and the time restrictions. We assume that primary users' modulation information is not known to CMTs or PCNs for general consideration. If in fact the primary user modulation is known to the system, then, a matched filter will be chosen for transmitter detection. In the following discussion, we only discuss the general cases and matched filter detection is not considered.

For a CMT node, energy detection is usually used for transmitter detection. There are two reasons for choosing energy detection. In case primary users frequently access the channel, a fast spectrum sensing algorithm is more important for efficient channel unitization. The strategy is to divide the entire channel in to N sub-channels. If a CMT wants to send a message, it will randomly choose one sub-channel and detect the energy. If energy is detected below the primary user interference temperature limit, the channel is assumed to be vacant, and the message will be sent. After going through three-way handshaking, the message is claimed as successfully sent. If

any step is broken, the message is not successfully sent, and then the same process will be repeated. CMT will randomly choose another sub-channel, and detect the energy and send the message. The interference temperature limit is predefined and can be adjusted during the communication.

There are three reasons for randomly choosing a sub-channel instead of scanning the entire spectrum to find the vacant part before transmitting. The first reason is complexity. The second reason is the time efficiency. The last reason is for multi-user purposes. If we use the method of scanning the entire spectrum and find the vacant part to which to transmit, assume that data transmission from the RF front end to the processor cost time t_1, running time in the process is t_2, RF front end reconfiguration time is t_3, the channel occupancy rate by primary user is λ, which is the primary user occupancy time per second, the number of sub-channel is N, then, the total time to prepare for a secondary user to access a vacant channel is:

$$T_E = t_1 + t_2 + t_3$$

For using the method of randomly choosing a sub-channel, the approximate total time cost is:

$$T_R = \left(\frac{t_1}{N} + \frac{t_2}{N} + t_3\right)(\lambda N) = t_1\lambda + t_2\lambda + t_3\lambda N$$

N is determined by the number of users in a cell. If using GNR radio and USRP as the platform, t_1 is the dominant time compared to t_2 and t_3. In our prototype of DCCS, randomly choosing a sub-channel is more time efficient. Furthermore, the first method also will cause unnecessary competition among secondary users because secondary users who need to send messages at the same time have a large probability of choosing the same vacant part of the channel.

For a PCN node, transmission detection is more complicated. Because of the cell structure, each cell occupies a different range of spectrum and the allocation is negotiated among PCNs. This means that a PCN needs to have a broader overview of the channel conditions. For the channels that are dense, a PCN tends not to include these channels within its cell allocated spectrum. Energy detection itself is not enough for PCN transmission detection. A PCN also needs a transmission detection method that can accurately locate the primary users, especially those that

appear frequently. Thus, besides the energy detection, a direct feature detection algorithm is also used for PCN. Similar to a cyclostationary feature detection algorithm, the motivation for this algorithm is also to extract the features for a primary user, and use this information to identify primary users. As it is introduced explicitly in Chapter 5, the complexity of this algorithm is greatly reduced, and the signal capture duration is also much shorter compared to cyclostationary detection. The information extracted is enough to demodulate primary user, so that locating the primary signal in the spectrum and analysis of the entire channel vacancy condition is feasible. After identifying the entire channel, the spectrum is divided among PCNs for each cell. In each cell, PCN will divide its allocated spectrum into a number of sub-channels. The number of the sub-channels depends on the maximum associated CMTs with this PCN.

A message exchange scheme among CMTs and PCNs is used to share sensing information. The PCN starts with one of the sub-channels and goes through all the sub-channels one by one. In each sub-channel, the PCN will listen to the sub-channel for a small time interval t_{sub}. τ_{sub} can be seen as a time counter. When the start time is reset, it will start to recount the time toward τ_{sub}. During this interval, PCN will determine next step based on the following conditions:

1. If a PCN detects energy, but the signal is not a message sent from a CMT, then, the PCN will move to next sub-channel, and start time of $\tau_{sub}$is reset .
2. If the PCN detects energy, and the signal is the message from a CMT requesting service from the PCN, PCN will leave response message, reset the start time ofτ_{sub}, and wait for the acknowledgement message.
    a. If the acknowledgement message arrives within the time interval τ_{sub}, based on the service requirement, if necessary, the PCN will assign a processor to taking care of the service requested - for example, a forwarding request. Meanwhile, the PCN will move on to the next sub-channel immediately after receiving the acknowledgement message and the start time of τ_{sub} will be reset.
    b. If the acknowledgement message doesn't arrive within the time interval t_{sub}. , the PCN will move to next sub-channel immediately after time interval t_{sub}, and the start time of τ_{sub} is reset. This message exchange has failed and the request needs to be sent again.

3. If PCN doesn't detect any message within the time interval t_{sub}, PCN will move to the next sub-channel immediately after time interval τ_{sub}, and the start time of τ_{sub} is reset.

This process can also be explained by a flow chart, as in Figure 14.

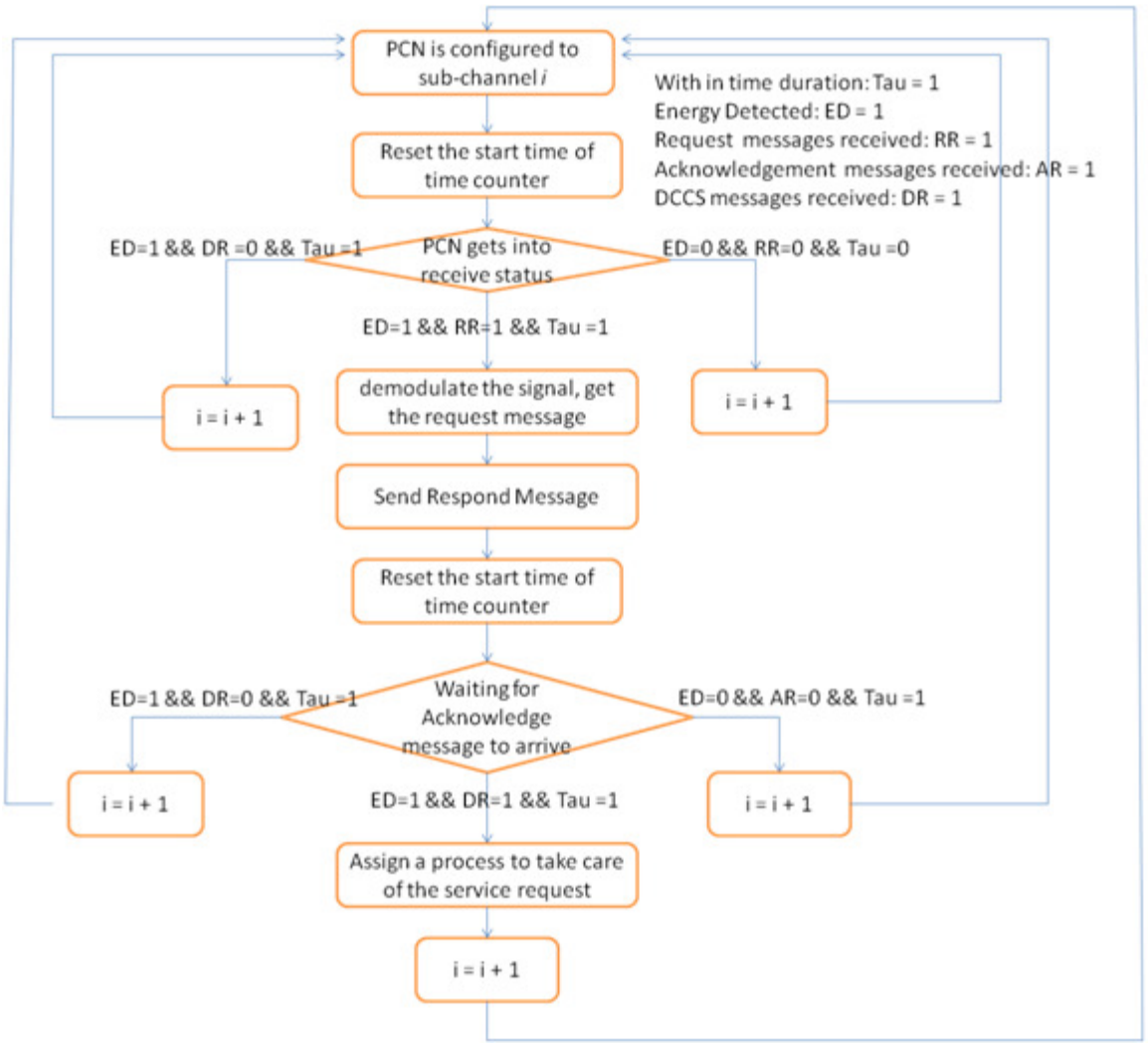

Figure 14: Flow graph for PCN spectrum management

While a PCN cycles through the entire sub channel set, besides the process request, it also records some information about each channel, for example: primary user detected time, channel vacancy duration, etc. Using this information, the PCN is able to predict channel conditions. In item 2, one of the pieces of request information PCN needs to process is allocation of a channel.

PCN can use its predicated channel condition to allocate the channel that has the least probability that it will be interrupted by primary user during the allocated duration. The detailed algorithm and analysis will be shown in Section 4.3.

4.3 Channel Allocation

The purpose of channel allocation is to assign a certain bandwidth to users in an efficient manner while minimizing interference to primary users or other secondary users. “Channels in a wireless communication system typically consist of time slots, frequency bands and/or CDMA pseudo noise sequences, but in an abstract sense, they can represent any generic transmission resource”[29]. Fixed Channel Allocation (FCA), Dynamic Channel Allocation (DCA) and the combination of these two are three main categories for channel allocation strategies.

FCA is defined as permanently assigning a set of channels to each cell, according to the allowed reuse distance. This is the minimum distance necessary to reduce co-channel interference [29, 30]. A call can only be served within the set of channels available to the cell. A call is blocked if no channel is available in the cell when the call arrives. DCA, centrally, allows any system channel to be temporarily assigned to any cell, provided that the constraint on the reuse distance is fulfilled[30]. Using DCA, cells can have different traffic loads and higher capacity. However, the complexity of DCA makes it difficult to implement in a reliable commercial product. As alternatives for DCA, dynamic frequency selection (DFS) is implemented in IEEE802.11h, and spread spectrum is widely used in 3G networks. Spread spectrum has relatively lower spectrum efficiency than DCA, and some researchers consider OFDMA combined with DCA as the 4G channel allocation strategy. Other than spread spectrum, the combination of DCA and FCA, which is also called hybrid channel allocation (HCA), is popular in cellular networks. Channel borrowing is one of the most straightforward hybrid allocation schemes. In this scheme, the basic structure is FCA, in which channels are pre-assigned to cells. “If a cell needs a channel in excess of the channels previously assigned to it, that cell may borrow a channel from one of its neighboring cells given that a channel is available and use of this channel won't violate frequency reuse requirements”[22]. Two methods are usually used to guarantee the frequency

reuse requirements are satisfied, one is to lock the borrowed channel in the affected range[31], and the other is to use power control to avoid locking channels while meeting the frequency reuse requirements[32].

The channel allocation technologies introduced above are designed for primary users in the channels. In DCCS, because the nodes are assumed to be secondary users, the design of the channel allocation strategy is much different, although it adopted a great amount of analysis and theories from the primary user analysis. The channel allocation problem can be divided into two sub-problems, the poll of the channel set for each cell, and the strategy to choose the optimal channel among them. In other words, the first sub-problem is channel allocation among the users and the second sub-problem is to determine the best channel to allocate in a cell when one is requested.

4.3.1 Channel Allocation among Cells

It can be understood that all the channels for intra-cell communication in DCCS are borrowed, without the owner knowing this. Since this is not perceivable by the primary users, there is no need for cooperation with them. Neither power control nor locking can be used to satisfy frequency reuse requirements. This statement also applies to other DSA scenarios. Thus, solving the channel borrowing problem in DCCS is the basis for resolving other channel allocation issues. We define two concepts, receiving range and transmission range, in order to express the following method more clearly. For node A and node B, if A can receive signals from B, A is in the receiving range of B, and B is in the transmission range of A. As we have introduced in Section 3.1.2, an MT sends a request based on CSMA, and a link cannot be created unless a PCN responds. Thus, neither the MT nor the PCN are in the receiving range for any primary user occupying this channel. If it can be guaranteed that neither of them is in the transmission range, the problem is solved. Thus, we just need to make sure that the transmission range of a node in DCCS is smaller than the receiving range of a base station in a standard network. This is true in DCCS because portable devices are used in DCCS and the power transmitted is much less than the standard cellular system or TV signals if used in 700MHz white space. Thus, it is not

necessary to block the channel that is temporarily used in DCCS channel because the utilization is not impacting other nodes in the frequency reuse range. The above solution is made based on an assumption that none of the nodes in the primary user system are only receiving without transmitting. In [33], a method exploiting the local oscillator leakage power emitted by the RF front end of TV receivers to detect the presence of primary receivers is presented. However, the short detection range, long detection time and its specifically designed for TV signal prevent this method to be implemented in DCCS system.

An effective channel allocation scheme should address the following three principles.

1. "Channel allocation schemes must not violate minimum frequency reuse conditions" [29].
2. "Channel allocation schemes should adapt to changing traffic conditions" [29].
3. "Channel allocation schemes should approach the minimum frequency reuse constraints so as to efficiently utilize available transmission resources" [29].

Up till now in this section, we have discussed the first principle. For the third one, because the channels used in DCCS are the ones that primary users are not currently using, any throughput for DCCS can be seen as the gain from the designed primary channel allocation scheme. For the second principle, the following channel allocation among cells would demonstrate its adaption to changing traffic conditions.

There are two options for channel allocation among cells; one is for each cell to use the entire operational spectrum for secondary communication and other method is to divide the entire spectrum into several parts and make each cell only responsible for a part. We are going to analyze both of the options and give the scenarios for which option should be chosen.

The operating band of DCCS is defined as $B_{operation} = (f_1 \sim f_2)$, and it is divided into N_{sub} sub-channels. Each PCN will cycle through each of these sub channels. CMTs are accessing these sub-channels according to the scheme mentioned in Section 3.1.2. The sub-channel for managing message exchange is also the same channel as allocated when requested by CMT. Because of the listening before transmitting principle, the frequency reuse range is automatically formed.

Compared to primary user cellular structure, the secondary user channel allocation strategy is relatively simple, and effective.

There are several advantages for choosing this option.

1. Simplicity. Unlike in a standard cellular system, where any request from a mobile terminal must be on a pre-assigned channel, there is no pre-assigned channel in the DCCS system. Thus, defining the same rule for the frequency accessing method for each of the cell simplifies the entire process. When a CMT comes to an area and requests to join the network, it doesn't have the knowledge of which PCN it can connect to or which cell it currently belongs to. Thus, having any CMT choosing a channel from the same range simplifies the process.
2. Efficiency. For a PCN, the managed message exchange process is also the channel stochastic information recording process. A PCN is always prepared to provide the information which of the sub-channels is optimal when a sub-channel allocation request message is received. The algorithm for determine optimal channel is described in Section 4.3.2
3. Capacity. Because all the sub-channels are available for each cell, the traffic load can be properly allocated. The cell with more traffic will access more sub-channels. This fulfills the second principle for being estimated as a well designed channel allocation protocol.
4. Frequency reuse. Compared to manually defining the range for the waveform propagation, in this system, the frequency reuse ranges are automatically determined. The spot where no signal can be sensed is considered as outside of the propagation range and is useful for access to the spectrum.

However, this strategy works better than option two (to divide the entire spectrum into several parts and each cell is only responsible for a part) in the scenario where the number of sub-channels N_{sub} is small and the channel occupancy by the primary users is high. If the number of the sub-channels is large, it has an obvious drawback. PCN circulation time is increased linearly with N_{sub}. For a software defined radio, although the reconfiguring duration is much less than that required for FPGA on the fly configuration, it is still not fast enough. As we introduced in

Section 3.1.2, $T_E = t_1 + t_2 + t_3$. Thus, the circulation time for going through the entire channel would be at least $T_E N_{sub}$. This circulation time is directly related to the accuracy of the sub-channel evaluation by PCN. The shorter the circulation time is, the more time efficiently a PCN can update the sub-channel status. Thus, for a large number of sub-channels, parallel processing of multiple sub-channels in the PCN is necessary. Currently, the parallel processing is not researched and is left for future work.

After discussing the allocation of sub-channels within a cell, in Section 4.3.2, we will discuss the method to choose the specific optimal sub-channel when a channel request message is received by a PCN.

4.3.2 Optimal Channel Allocation Within a Cell

The best duration sub-channel is defined as the sub-channel that can has the largest probability that it will not be taken back by the primary user during the requested secondary user communication time. To choose the best duration sub-channel when a sub-channel request received by a PCN, we need to be able to predict which channel has the least probability that, if allocated to the secondary users, the communication will be interrupted because of primary users' return. Because the primary users' occupancy of the channel can be described as a type of random variable distribution, the next moment behavior can be predicted to some extent. In this section, we will give a generic format to choose the best duration sub-channel, given the random distribution of the primary user's behavior. We will also give several models of the primary user behavior estimation.

Another important factor that impacts the secondary user's communication is the SNR of a sub-channel. The best quality sub-channel is defined as the sub-channel that has the largest SNR.

Which one to choose - the best duration, or the quality - that is the tradeoff here. Generally speaking, for primary users, two parameters, the maximum interference power level η without affecting the primary users' communication and the maximum probability ζ that the interference

power level perceived by the primary user may exceed η defines the interference constraints. [34] This description can also be used with little modification for secondary users.

In [35], the authors gave a detailed survey of channel allocation strategies. They focus on the method called "Opportunistic Spectrum Access (OSA)" for spectrum opportunity identification, exploitation, and policy regulation. In this paper, the primary users' behavior is modeled like this: " N channels are allocated to the primary users' network. The traffic statistics of the primary system are such that the occupancy of the N channels follows a Markov process with 2^Nstates, where the state is defined as the availability (idle or busy) of each channel." In [36], "the design of optimal sensing strategies has been formulated and addressed within the framework of partially observable Markov decision processes (POMDP)". In DCCS channel allocation we adopted a similar primary users' behavior model with modifications to the above description.

In Fixed Channel Allocation, assuming Poisson call arrivals and exponentially distributed channel holding times, call blocking probability can be derived according to the ERLANG-B formula. In DCCS, we build a model to describe the primary users' behavior and make the best decision based on the currently available description.

Research of channel allocation in a DSA scenario includes individual node detection based [36, 37] and cooperative detection based algorithms [38, 39]. For individual node detection based algorithms, the main focus is on decreasing the error probability when detecting the primary users. Cooperative detection based algorithms attempt to utilize the available spectrum opportunistically and efficiently from the network point of view. In [39], the tradeoff between optimal sensing time and the highest network throughput is discussed. Channel allocation algorithms described in [40, 41] are based on routing decisions or traffic conditions in the network.

In DCCS, because a PCN cycles through all of the channels, it has an overview of all the available channels. The objective of channel allocation algorithm for DCCS is to choose the best

channel among multiple available channels. An algorithm that can evaluate the channel feature and select the optimal channel is important for increasing secondary users' communication quality for CMTs. To realize this objective, an evaluation method of the channels has to be established. Estimated channel throughput is used to evaluate channels in order to be able to transmit more data. The maximum throughput channel allocation algorithm described below aim to build up the model to provide a method to evaluate the channel quality in a DSA scenario.

To make the description easier and more accurate, we define the following variables.

N: Number of sub-channels in DCCS operation range.

$S_{channel} = [S_1, S_2, S_3, \dots, S_N]$: The status (Idle/Occupied) of each sub-channel, where $S_i \in (1 \text{ or } 0)$.

$\Lambda_{channel} = [\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_N]$: The accessing rate for primary users on sub-channels based on observation.

$T_{channel} = [t_1, t_2, t_3, \dots, t_N]$: The time interval between decision making time and time of detection

$P_{channel} = [p_1, p_2, p_3, \dots, p_N]$: The power detected for each sub-channel.

$B_{ch} = [B_1, B_2, B_3, \dots, B_N]$: The bandwidth for each sub-channel. In most cases for DCCS, the bandwidth for each sub-channel is the same. But using different bandwidths for each sub-channel is allowed.

p_{th}: Power threshold for power detection. If $p_i > p_{th}$, then $S_i = 0$, the ith sub-channel is not available for secondary users.

pa: The probability that a CMT can tolerant the link will be interrupted by primary user.

B: The bandwidth of each sub-channel.

$P_{maximum}$: The maximum power allowed for secondary user regulated by FCC or spectrum owner

As we have mentioned earlier, the issue is which measure to choose, the best quality or the best duration? Here, we propose a new objective function to integrate these two factors: providing the channel for the secondary user that has the maximum thoughput. The mathematical formulation is described as follows:

$$Objective: \max(S_i \cdot (1 - F_i) \cdot T_i \cdot C_i)$$

$$s.t.: \quad F_i = 1 - e^{-\lambda_i t_i}$$

$$T_i = -\frac{\ln(1 - pa)}{\lambda_i}$$

$$C_i = B_i log\left(1 + \frac{P_{maximum}}{p_i}\right)$$

$$i = 1, 2, 3, \dots, N$$

Again, in this model, our goal is to maximize the throughput of the sub-channel.

S_i represents for the availability of the ith sub-channel to the secondary user based on the observations. If the detected power of the ith sub-channel is less than desired thresholdp_{th}, S_i equals to 1; otherwise, S_i equals to 0.

F_i represents the probability of a sub-channel not being available even though the detection results show it is available. In other words, it represents the probability that a primary user appears in the interval between last detected time and the decision making time. Thus, $1 - F_i$ represents the probability that the channel is still available at the decision making time given the condition that it was available for the secondary user the last time it is observed.

The expression for T_i gives the probability (1-pa) that a primary user will appear within Ti seconds after the last observation. In other words, $P_{possion}(t < T_i) = 1 - e^{-\lambda T_i} = pa$, where $P_{possion}(\cdot)$ is the Poisson distribution with parameter λ_i, which describes the average rate at which primary users reappear.

C_i is the ith sub-channel capacity given the bandwidth of this sub-channel, the maximum allowed transmission power of a secondary user and the noise floor level of this sub-channel.

The basic principle of this sub-channel selection strategy is based on the assumption that the best channel is the channel that can let the secondary users transmit the maximum amount of information. This model is implemented in the prototype and the analysis is shown in Section 7.2.

Chapter 5 Signal Classification and Synchronization for PCN

In this chapter, we present a systematic description of the signal classification and synchronization system implemented on cognitive radio nodes in DCCS. This system can also be used independently in other applications. This chapter is not organized the same way as other chapters in this dissertation, but instead is an independent chapter. It was published as [42] and represents joint work by Qinqin Chen and the author of this dissertation. Each of us will use it as part of our dissertations. While this procedure might seem a bit unusual, our advisor felt that the work is of such significance it should appear in full, rather than simply being cited as a reference in both documents.

Extracting parameters from a received signal and auto-demodulating based on these parameters without prior information from the transmitter can be beneficial to Dynamic Spectrum Access (DSA), Cognitive Radio, and many other applications. Universal Classifier and Synchronizer (UCS) is conceived as such a self-contained system which can detect, classify, synchronize with a received signal and provide all parameters needed for physical layer demodulation. The accommodated modulations include AM, FM, FSK, MPSK, QAM and OFDM. UCS can be used in different multi-user access schemes. The designed system has been verified by a prototype using GNU Radio in Linux plus a Universal Software Radio Peripheral (USRP), as well as other software defined radio (SDR) platforms. Performance for key components and the entire system has been evaluated by theoretical analysis, Over the Air (OTA) experiments and computer simulations.

5.1 Introduction

Signal classification has many applications in wireless communications for both civilian and military purposes [43]. An M-ary hypothesis testing problem is commonly posed to detect and classify a signal [44]. In our system, we are not only focusing on the existence and type of the

signal, but also extracting its physical layer features for demodulation. A software-based design at IF or quasi-baseband of a cognitive radio makes it able to automatically change configuration settings based on varying channel environments and user requirements. A radio equipped with UCS can respond to this change so that the continuity of communication can be guaranteed. UCS has the ability to demodulate signals without benefit of priori information and forms the heart of a truly cognitive receiver [45].

In the DSA scenario, a pair of cognitive transceivers occupies a certain channel as secondary users until a primary user appears. The secondary users must immediately move to an unoccupied channel. In order for this pair of cognitive radios to stay tuned in to continue communicating, either an out-of-band control channel needs to be used to negotiate setting information, or a pre-defined channel changing protocol needs to be complied with to achieve automatic coordination. The former increases system overhead in control message conveyance and control channel management. The latter also has drawbacks. The primary users' behavior and channel environment conditions are not fully predictable, while under a pre-defined channel changing protocol the secondary users have to alter their settings in a fixed manner, in order to maintain the link. Thus the allowance for a radio to cognitively change its key physical layer parameters according to the channel environment is dramatically limited. Therefore, by sole use of a pre-defined channel changing protocol, optimal resource utilization cannot be achieved. With UCS integrated in the radio, the aforementioned drawbacks can be conquered because the receiver can automatically extract necessary parameters from the detected signal and continue the previous communication. Such a system can be employed in DSA-enabled scenarios to promote throughput, simplify channel management, optimize resource utilization, and provide better robustness under varying conditions, without increasing overhead. For example, if the changed channel is different from the previous one in terms of bandwidth, Signal Noise Ratio (SNR) or other propagation features, the physical layer parameters can be changed for better performance without interrupting the communication. The parameters include carrier frequency, modulation type, symbol rate etc. As shown in Figure 15, during the DSA resource optimization process, unlike the conventional communication scenarios with pre-agreement, the communication waveform is shuffled like a Rubik's Cube, and UCS is in hand to put it back in proper order.

Another example of UCS's applications is that UCS can accommodate multiple devices or support a variety of modulation settings. Whenever some resources become unavailable, the transmitter can switch to use other resources and the receiver with UCS will automatically follow it.

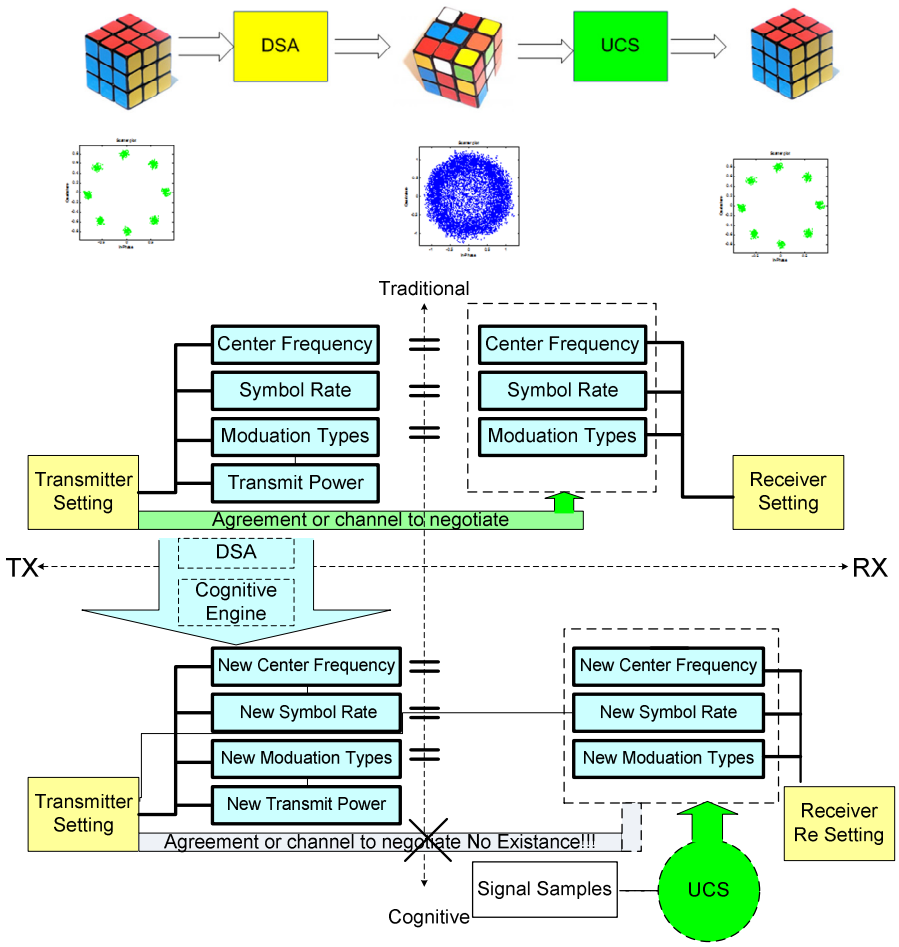

Figure 15: Role of UCS in DSA[46]

The remainder of this paper is organized as follows. In Section 5.2, we present our objectives after introducing the research background and state of the art. Section 5.3 begins with a description of the general cognitive receiver model, followed by the overview of our UCS system. We address the design and implementation details for each module of the UCS system in Section 5.4. Section 5.5 describes the UCS prototype and evaluates the performance for several key components and the entire system by theoretical analysis, OTA experiments and computer simulations. Conclusions are made in Section 5.6.

5.2 Background and State of the Art

As a branch of SIGINT signal classification became an attractive research topic in 1980s. Because of the recently increasing interest in cognitive radio, signal classification is gaining more attention. The methodologies and technologies in this area can be roughly divided into three categories, (a) Maximum Likelihood (ML) based, (b) feature-extraction based, and (c) cyclostationary feature based [47]. Method (a) classifies by comparing the likelihood of candidate signal and modulation types. Reference [48] is a classic article that talks about the optimal classification rules. Reference [49] is about asynchronous classification for MFSK. Method (b) directly extracts phase or amplitude features from the target signal in order to differentiate modulations. Zero crossing and wavelet technology are quite frequently involved in this area[50-52]. Some papers combine (a) and (b) to get better performance. For example, in Reference [53], both ML and extracted features are used for OFDM signal detecting and classification in cognitive radio. Method (c) is attractive for DSA applications because of its ability to detect and classify signals at low signal to noise ratios [54]. The methods mentioned above have excellent performance in certain scenarios. The scenario conditions include channel types, signal types, and equipment. Our objective is to design a universal signal classification and synchronization system which can analyze a signal's physical layer features with minimal prior information and application limits and can demodulate the signal using the acquired information.

5.3 System Overview

UCS has been developed and implemented to identify signals including AM, FM, MPSK, QAM, MFSK and OFDM. The system is constructed to run on Universal Software Radio Peripheral (USRP) with the GNU Radio platform. It is reconfigurable to provide adaptivity in various environments, extendable to accommodate more signal types, and transplantable to other platforms, including Anritsu MS 2781A Signal Analyzer, Lyrtech SFF radio platform etc.

A fully developed UCS system with a reconfigurable demodulator can function as a cognitive receiver to classify, synchronize with and demodulate new signals. The system structure is

shown in Figure 16. A complete cognitive receiving loop starts from signal awareness. The hardware is initially set with a wide frequency span at a center frequency we are interested in, and then zoomed in to the band where a possible signal exists. It is then reset to a reasonable sampling rate, and capture time. The channel estimation and equalization process is launched afterwards. This module is used to classify the channel type and make the necessary compensation. The next step is called suite categorization and is used to determine the signal type. If the signal is analog, then it is sample based, which means we must demodulate it sample by sample, like FM and AM. If the signal is MPSK or QAM, then it is symbol based, which requires symbol timing and synchronization before demodulation. If the signal is wideband, for example OFDM, then it is block based, which means the demodulation is done block by block. If a signal cannot be clearly assigned to one of these three categories (sample, symbol, or block based) with information that has been extracted by the process described above, "knobs" [55, 56] like sampling rate will be "turned" to reconfigure the hardware for data re-collection. Otherwise, the signal is identified as one of the aforementioned three types and then fed to the corresponding reasoning module, where the parameters needed for demodulation are estimated. Three possible results will be returned by the verification process: "reclassify," "ready to configure," and "unknown signal type." If the parameters pass verification, they will be marked as "ready to configure" and formatted in an XML file [55, 56] to configure the demodulator and one complete loop is finished. If they do not pass, but the problem can be fixed by changing the hardware settings, the verification module will send "knobs" to reconfigure the hardware and the loop will be executed again. If the problem cannot be fixed, the result will be "unknown signal type", and the signal will be stored in a database for future classification. The system block diagram shown in Figure 3 includes all the modules implemented in UCS prototype. In the next section, we will describe the details of each module. The entire structure of UCS prototype can be understood as 4 branches and 3 phases. The 4 branches include multi-carrier digital signal, narrowband digital signal, analog signal and standard FSK signal based on the different feature extraction scheme for different types of signals. The 3 phases are briefly concluded as phase 1: classification, phase 2: synchronization and phase 3: demodulation based on different user requirements and scenarios. To implement parts or all three phases depends on the known signal information and user requirements. For example, to detect a FM primary user only needs to

implement phase 1; if a digital signal's center frequency is known, only phase 2 needs to be implemented.

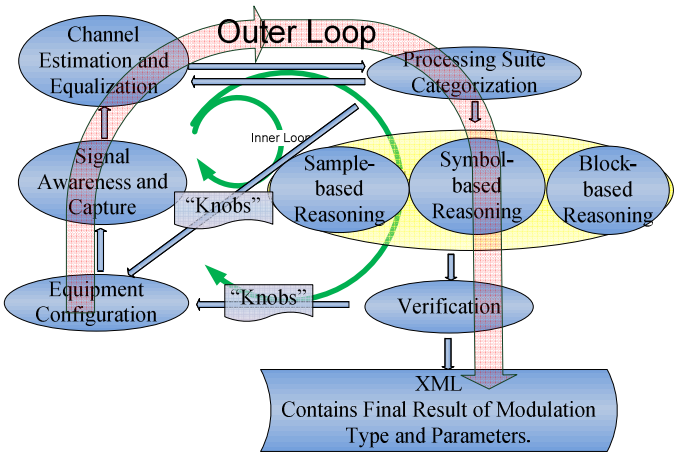

Figure 16: Cognitive receiver system structure

5.4 System Design and Implementation

Before we dive into the details for UCS system, it is necessary to define the important notations that will be used in the rest of this paper. These notations are outlined in Table 5.

Table 5: Notations

Symbols	**Definitions**
$\mathbf{R_s/T_s}$	Sampling rate/sampling interval
$\mathbf{T_c}$	Capture time
R/T	Symbol rate/symbol period
$\mathbf{f_c}$	Carrier frequency at TX
$\mathbf{f_{LO}}$	Local oscillator frequency at RX
$\mathbf{\Delta f = f_c - f_{LO}}$	Frequency offset
$\mathbf{\lceil x \rceil}$	Minimum integer not less than x

$\lfloor x \rfloor$	Maximum integer not larger than x

5.4.1 Spectrum Sensing

Spectrum sensing has two purposes: it provides signal location information in the frequency domain (including center frequency and bandwidth), and also describes the spectrum occupation for the radio environmental map, which is important for DSA. In UCS, after the transceiver RF front-end has been set at the carrier frequency estimated by the spectrum sensing module, the sensed signal will be down-converted to IF or quasi-baseband. Then each sample will be expressed as a complex number for subsequent processing. The amplitude and phase of this complex signal contain important information for classification, synchronization and demodulation.

The spectrum sensing method adopted in UCS is energy detection based on power spectrum density (PSD). From the PSD distribution, in the frequency ranges where the SNR is larger than a certain threshold, we consider that a signal exists. The threshold is defined as the acceptable SNR for UCS to get the correct modulation information extraction.

5.4.2 Signal Capture

Two factors that determine how to capture a signal for analysis are sampling rate and capture duration, which represent bandwidth and resolution respectively in the frequency domain. The captured signal data needs to be detailed enough to guarantee the accuracy of further information abstraction as well as brief enough to simplify the complexity and decrease the effect of fading channels. The settings of these two parameters in the signal collection device influence the entire UCS procedure.

Sampling rate R_s equal to twice the bandwidth Bof the interested signal is enough for waveform recovery. In our case, between waveform recovery and information demodulation, a series of other processes is required in order to acquire symbol rate, center frequency and other physical layer features. Is it necessary to increase sampling rate even higher in order to obtain these

parameters? We will prove in Section 5.4.8 that in order to extract the correct symbol rate, sampling rate R_s has to be larger than $\max(2B, th_{Rs})$, where $th_{Rs}$is the threshold sampling rate required for symbol rate estimation. The value of the threshold is related to the pulse shaping used on the transmitter side.

Capture time T_c is another key factor to be considered. There are several restrictions for capture time. To simplify the calculation, capture time will be chosen as the minimum value which satisfies the following restrictions:

$\frac{1}{T_c} < R_f$: R_f is the frequency resolution required by the signal.

$T_c \times R_s > N_s$: N_s is the minimal number of data samples to correctly extract the features.

$\frac{1}{T_c} < f_m$: $f_m$is the maximum Doppler shift in the channel, which equals to $f_c \times v/c$, where v is the velocity of a mobile radio and c is the speed of light.

$\frac{1}{T_c} < 1/L_R$: $L_R=\sqrt{2\pi}f_m\rho e^{-\rho^2}$, L_R is the level crossing rate and ρ is the tolerable fading level for further classification and synchronization processing. This is only for a Rayleigh channel. If the channel is different, the relationship must be modified accordingly.

5.4.3 Channel Estimation and Equalization

A non-AWGN channel distorts a signal, and causes inaccuracy in UCS results. For example, without any additional processing, a MPSK signal going through a multipath channel might be classified as a QAM signal because of the amplitude distortion. Although the resulting error can be caught by our verification algorithm, an earlier channel analysis will reduce the waste of time and resources. Thus, in this section, we introduce and apply our channel estimation and equalization.

A wireless channel can be described by four aspects: m amplitude attenuations, m delays, m frequency shifts, and noise, where m is the number of propagation paths. As it will be shown in

the system performance analysis, noise below a certain level will not mislead the classification result. Thus, in this section, we focus on eliminating the influence caused by multipath delays and frequency shifts. If we decompose the three aspects, multiple path attenuation generates multiple amplitude deviations of the received signal, multiple delays generate multiple phase deviations, and multiple frequency shifts generate multiple symbol timing deviations. When the radio is not moving fast, frequency shifts are negligible. A convenient method is to use a rake receiver before further signal processing to counter the effects of multipath fading [57]. We adopted this method, but made a slight change:

$$r(t) = \sum_{n=1}^{L_R} s\left(t - \frac{n}{R_s}\right) + n(t)$$

Here L_R is the number of fingers of the rake receiver. Instead of using bandwidth of the channel we use $1/R_s$ as the delay of each finger because we are oversampling. When the radio is moving so fast that the frequency shifts cannot be neglected, instead of using the sum of all the fingers, we only use the maximum one because branches with different symbol timing deviations cannot be simply totaled. Thus,

$$r(t) = \max\left(s\left(t - \frac{n}{R_s}\right) + n(t)\right)$$

The output of the modified rake receiver is a clean signal ready for further processing.

5.4.4 Modern Wireless Communications Modulations and Scenarios

UCS is designed for practical systems. It is necessary to give a summary of the commonly used signal types and multi-user access schemes in modern wireless communications. Listing commonly used signal types helps clarify the research focus of UCS. In this paper, user scenarios analysis mainly focuses on single and multiuser access schemes. User scenarios analysis is beneficial for balancing UCS system implementation phase level and computation complexity. Currently, the most frequently used communication systems include: cellular, Wi-Fi, WiMAX, public safety radios, and, customized cognitive radios. The goal of UCS is to extract physical

layer features. The physical layer modulations for each standard are listed in Table 6, and the multiuser schemes for each standard are listed in Table 7.

Table 6: Modulation Types of Interest

Communication System	Physical Layer Modulation Types Adopted
GSM	GMSK
GPRS	GMSK
3G	CDMA
4G(WiMAX)	OFDM
Wi-Fi	OFDM
Public Safety Radio	FM, C4FM, CQPSK
Customized waveforms for cognitive radio	FM, AM, MPSK, QAM, OFDM

From Table 6, we can see that the waveforms of interest include GMSK, CDMA, OFDM, FM, C4FM (Continuous 4 level FM), MPSK, QAM, etc. Among all these modulation types, we do not include CDMA in the UCS system because processing CDMA requires the spreading code. In standard communication systems, parameters are fixed; and in customized waveforms, the parameters are prone to change with varying environments. Our system can deal with both situations.

Table 7:User Scenarios Description

Communication System	**Multi-user Access Scheme**
GSM	Mixed TDMA and FDMA
GPRS	Unused TDMA channels in GSM
3G	CDMA
4G(WiMAX)	OFDMA

Wi-Fi	CSMA/CA
Public Safety Radio	Random Access or Trunking
Customized waveforms for cognitive radio	DSA

In Table 7, multiuser access schemes fall into two categories: pre-defined multi-user schemes, which include TDMA, FDMA, CSMA and OFDMA, and primary/secondary user scenarios. One of the main purposes of UCS in cognitive radio is to accurately recognize the target of on-going communication. Because of multiuser scenarios, other radios' behavior will impact this recognition process. The basic principle of UCS is to extract physical layer features so that any signal can be demodulated. Technically, if a signal can be demodulated, the identification of the signal is not a problem, and then a multiuser scheme will not prevent UCS application. However, either due to computation complexity or real time requirements, it is not a smart idea to run entire UCS on each of the signals appearing in the spectrum. As illustrated in Figure 17, running necessary phases instead of entire UCS system is more efficient. For example, for a standard primary user, the spectrum location and modulation setting are pre-defined and fixed. This means that for a decided frequency, the primary user features are assured. If it is established that the signal is not a primary user, UCS will be called to determine the signal type, and if the detected signal type might be the type for the target of on-going communication, the next step synchronization and demodulation will be continued to identify the secondary user.

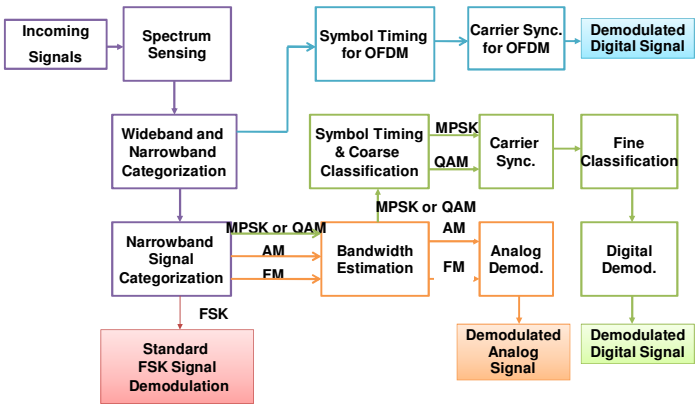

Figure 17:UCS system frame – functional block

5.4.5 Narrowband and Wideband Categorization

As mentioned in system overview, all considered signals in UCS system fall into one of three categories: sample-based, symbol-based or block-based signals. These can also be viewed as analog signals, digital signals and wideband signals. Because these three categories have different processing methodologies, they need to be differentiated before further processing. This section is to differentiate wideband signal and narrowband signal, and the next section focuses on further categorization for narrowband signals. A wideband signal is a signal whose period duration is longer than the maximum delay of the channel, which causes frequency selective fading. Spread spectrum and dividing the entire channel into multiple orthogonal sub-channels are the two commonly used methods to resist this fading. In this paper, we refer to the former as a CDMA signal and to the latter as an OFDM signal. As mentioned in Section 5.4.4, CDMA signal could not be detected by UCS system because its processing requires the spreading code. Thus, an alternative way to achieve the narrowband and wideband categorization is OFDM signal identification.

To identify a signal as OFDM, we correlate the incoming signal with itself [58]. To illustrate this situation, we did OTA experiments for MPSK, analog FM and OFDM signal; the results are shown in Figure 18. The correlated output is different in the case of narrowband modulations and OFDM modulation. This difference is due to the cyclic prefix present in the OFDM signal, which gives us multiple peaks as opposed to a single peak in narrowband modulations.

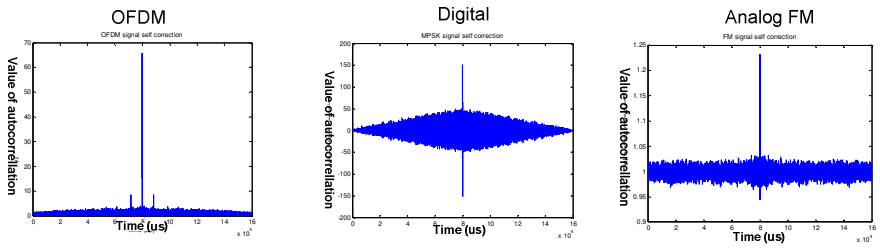

Figure 18: Autocorrelation to detect OFDM

Will the peak be distinct enough to serve as the foundation for differentiation? The following calculation shows the identification accuracy under different SNR's. The duration of an OFDM

symbol is $T = T_d + T_g$, where T_d is the symbol duration without the cyclic prefix and $T_g$is the duration of cyclic prefix.

A continuous-time OFDM signal at baseband can be written as [53]:

$$x(t) = \sum_{k=1}^{K} S_{m(k)} e^{j\frac{2\pi\Gamma(k)t}{T_s}}$$

Here $S_{m(k)}$ is the mth OFDM symbol at the kth subcarrier and Γ denotes the set of K user subcarriers. Channel delay and frequency shift will not influence the ratio between the cyclic prefix peak and noise floor. It can be proved that it is sufficient to distinguish between narrowband and wideband signals. The detailed analysis appears in Section 5.5.

Narrowband signals and wideband signals have different processing methodologies. We introduce narrowband signal classification and synchronization from Section 5.4.6 to Section 5.4.9, and wideband signal classification and synchronization from Section 5.4.10 to Section 5.4.12.

5.4.6 Narrowband Categorization

This section is divided into several sub-sections, each subsequent sub-section building on the information described in the previous sub-section. The first sub-section describes the general formulation of the narrowband waveforms that are currently categorized and provides a general categorization scheme for these waveforms. The second sub-section describes the algorithm for performing the actual categorization of a narrowband signal and provides a general set of metrics for categorization thresholds as affected by noise as well as other irregularities that appear in mobile systems. It is important to understand that it is this coarse signal classification which will make the subsequent processing, e.g. symbol timing, carrier synchronization, and demodulation, much more efficient.

5.4.6.1 Narrowband Generalization

Prior to providing the categorization of current waveforms of interest, it is efficient to describe these waveforms using a general structure. The general form is defined as:

$A(t)\cos[2\pi f_c t + \theta(t)]$

where $A(t)$ and $\theta(t)$ are the amplitude and phase of the waveform, respectively. Note that although the amplitude and phase are described as time-varying for all waveforms, it is possible that for some waveforms these are constants.

5.4.6.2 Analog and Digital Non-linear Modulation

Both FM and continuous-phase FSK (CPFSK) signals may be represented in the time domain by

$$s_{FM,CPFSK}(t) = A_c \cos\left[2\pi f_c t + 2\pi k_f \int_{-\infty}^{t} m(\eta)d\eta\right] \quad (1)$$

where $m(t)$ is the message signal, also called the modulating signal [59], and the constant k_f represents the frequency sensitivity of the modulator [60]. In FM, $m(t)$ is continuous; in CPFSK, $m(t)$ contains discontinuities, but its integral is continuous.

5.4.6.3 Analog and Digital Linear Modulation

The other modulations: AM, PSK, QAM belong to the linear modulation family, which can be expressed in the time domain as follows [59]. $s(t) = \mathrm{Re}[Am(t)\exp(j2\pi f_c t)]$

$$= A\sqrt{m_R^2(t) + m_I^2(t)}\cos\left[2\pi f_c t + \arg\left(m_R(t) + jm_I(t)\right)\right] \quad (2)$$

Here, $\arg(m_R(t) + jm_I(t))$ means the phase component of the modulating signal $m(t) = m_R(t) + jm_I(t)$. For the convenience of the posterior analysis, we represent a full AM signal by:

$$s_{AM}(t) = A_c[1 + k_a m(t)]\cos(2\pi f_c t) \quad (3),$$

where k_a is a constant that determines the percentage modulation [60].

5.4.6.4 General Categorization for Narrowband Modulations

By qualitatively analyzing their features (including instantaneous phase, amplitude, and frequency) embodied in equation (1)-(3), the aforementioned narrowband modulations can be sorted into the categories shown in Table 8.

Table 8: Autocorrelation to detect OFDM

Continuous phase			**Discontinuous phase**	
Constant amplitude		Varying amplitude	Single-value envelope	**Multiple-value envelope**
Continuous freq.	Discrete freq.	AM	MPSK	QAM
FM	CPFSK			

Table 8 gives a general outline of what features to look for when attempting to categorize a received signal. It is also the theoretical basis for the feature-based categorization algorithms, one of which is interpreted by the flow-graph in Figure 19.

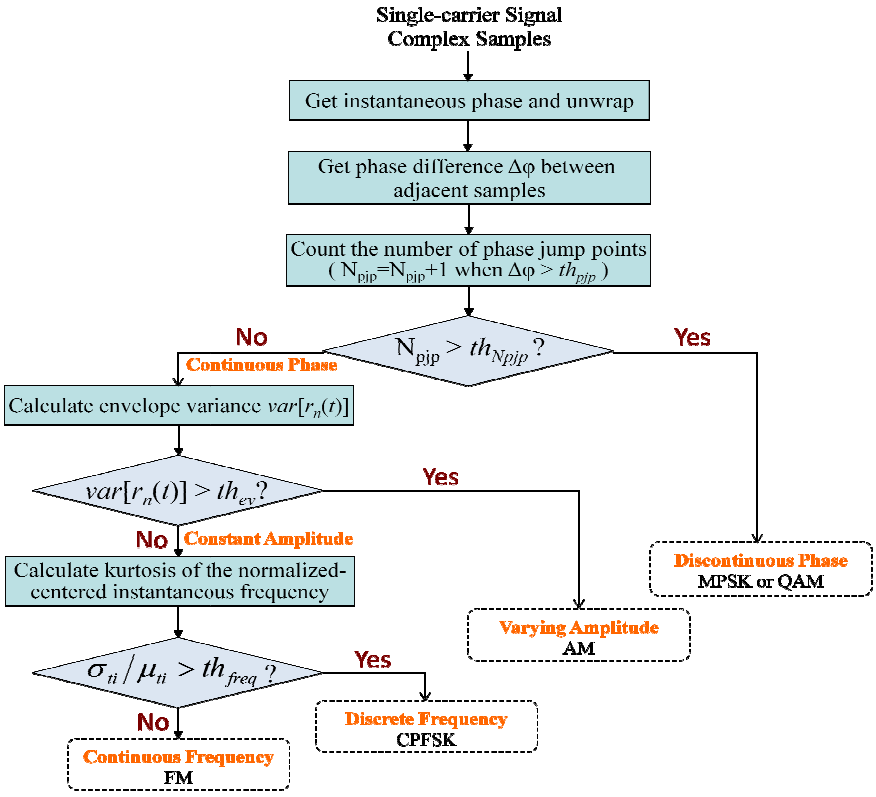

Figure 19: Narrowband categorization flow chart

In the real systems, signals' features may differ from their theoretical counterparts due to distortion caused by noise and imperfections (including Doppler shift, frequency offset between transmitter and receiver etc.). It is important to take into consideration these effects when we derive the quantitative thresholds for the criteria used for the feature-based categorization algorithm, based on theoretical analysis and experimental settings.

5.4.6.5 Effects of Noise to Narrowband Signals

The additive filtered noise $n(t)$ at the receiver's band-pass filter output can be defined by [60]:

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) = r_n(t)\cos[2\pi f_c t + \theta_n(t)]$$

Here the envelope $r_n(t) = [n_I^2(t) + n_Q^2(t)]^{1/2}$ is Rayleigh distributed, and the phase $\theta_n(t) = \tan^{-1}[n_Q(t)/n_I(t)]$ is uniformly distributed over 2π radians. The addition of noise $n(t)$ will change the envelope and (or) phase of the signals as follows:

$x(t) = s(t) + n(t) = r(t)\cos[2\pi f_c t + \varphi(t)]$, where

$$r(t) = \{A^2(t) + r_n^2(t) + 2A(t)r_n(t)\cos[\theta_n(t) - \theta(t)]\}^{\frac{1}{2}},$$

$$\varphi(t) = \theta(t) + \tan^{-1}\left\{\frac{r_n(t)\sin[\theta_n(t) - \theta(t)]}{A(t) + r_n(t)\cos[\theta_n(t) - \theta(t)]}\right\} \quad (4)$$

The envelope $A(t)$ and phase $\theta(t)$ for different modulations are listed as follows:

Modulation Type	**Envelope $A(t)$**	**Phase $\theta(t)$**
AM	$A_c[1 + k_a m(t)]$	0
FM or CPFSK	A_c	$2\pi k_f \int_{-\infty}^{t} m(\eta)d\eta$
MPSK or QAM	$A\sqrt{m_R^2(t) + m_I^2(t)}$	$\arg(m_R(t) + jm_I(t))$

When SNR makes the receiver operate satisfactorily, the signal's amplitude level $|A(t)|$ is large compared with the noise envelope. Then, with reasonable approximations, $\varphi(t)$ can be simplified as:

$$\varphi(t) \approx \theta(t) + \frac{r_n(t)\sin[\theta_n(t) - \theta(t)]}{A(t)}$$

5.4.6.6 Establishing Discontinuous vs. Continuous Phase by Counting Phase Jump Points

Suppose the frequency offset is Δf; the Doppler shift of the channel, if it exists, is $f_d(t)$; the initial phase of the transmitted signal is θ_0, the phase change introduced by noise and the delay of the channel is $\Delta\theta(t)$. Then, the instantaneous phase of the down-converted signal at the receiver side will be $2\pi[\Delta f + f_d(t)]t + \theta_0 + \Delta\theta(t) + \theta(t)$

If we totally captured L_s samples using the even sampling interval $T_s = 1/R_s$, the phase change between the nth and the $(n+1)$th sample is composed of four parts expressed as follows and indicated by circled numbers ①-④:

① $2\pi\Delta f T_s$

After spectrum sensing, the received signal is down-converted to quasi-baseband. In our case, the frequency offset $|\Delta f|$ from DC is around 2kHz. In order to handle signals with bandwidth B up to several hundred kHz, the sampling rate R_s is usually set to be greater than the Nyquist rate 2, i.e. $R_s > 200kHz$. Therefore, the phase change caused by the frequency offset within one sampling period is $|2\pi\Delta f T_s| < 0.02\pi$.

② $2\pi * [f_d(t_0 + nT_s + T_s) * (t_0 + nT_s + T_s) - f_d(t_0 + nT_s) * (t_0 + nT_s)]$

The Doppler shift differs in different application scenarios. In the terrestrial cases, the employed carrier frequencies may vary in a wide range from VHF, UHF to SHF, and the relative velocity between the transmitter and the receiver can be 0~60 mph. For example, when $f_c = 4.9\text{GHz}, v = 60\text{mph}$, the Doppler shift is $f_d \approx 436\text{Hz}$. In most cases, the phase change accumulated by the Doppler shift within one sampling period is much less than $2\pi f_d T_s = 0.00436\pi$.

③ $[\Delta\theta(t_0 + nT_s + T_s) - \Delta\theta(t_0 + nT_s)]$

From equation (4), we can see that noise causes tiny phase changes between two consecutive samples. For narrow band signals, channel delay is much smaller than the symbol interval; the phase changing caused by channel delay is also a small value.

④ $[\theta(t_0 + nT_s + T_s) - \theta(t_0 + nT_s)]$

The phase changes contributed by modulated information depend upon the modulation scheme. From equation (1), we can see that the difference between the instantaneous phases of two adjacent samples for FM or CPFSK is $2\pi k_f \int_{nT_s}^{(n+1)T_s} m(\eta)d\eta$. Considering that the maximum frequency deviation, expressed as $|k_f \max(m(t))|$, is usually 2500Hz for narrowband FM, and 5000Hz for wideband FM, we can get:

$$\left|2\pi k_f \int_{nT_s}^{(n+1)T_s} m(\eta)d\eta\right| < |2\pi k_f T_s \max(m(t))| < 0.05\pi$$

For M-ary CPFSK, $\left|2\pi k_f \int_{nT_s}^{(n+1)T_s} m(\eta)d\eta\right| < 2\pi k_f T_s(M-1)$, which is less than 0.018π for P25 C4FM (its maximum frequency deviation is 1800Hz) if we choose $R_s = 1/T_s > 200kHz$.

For other modulations, the absolute value sets of the phase difference between successive samples are listed as follows.

AM: 0

BPSK: $[0, \pi]$

QPSK: $[0, \pi/2, \pi, 3\pi/2]$

8PSK: $\left[0, \frac{\pi}{4}, \frac{\pi}{2}, \frac{3\pi}{4}, \pi, \frac{5\pi}{4}, \frac{3\pi}{2}, \frac{7\pi}{4}\right]$

16QAM:

$$\begin{bmatrix} 0, & 0.1476, & 0.2048, & 0.2952, & 0.3524, & 0.5, \\ 0.6476, & 0.7048, & 0.7952, & 0.8524, & 1, & 1.1476, \\ 1.2048, & 1.2952, & 1.3524, & 1.5, & 1.6476, & 1.7952 \end{bmatrix}\pi \qquad (5)$$

When the phase change between the two adjacent samples is larger than a certain threshold th_{pjp}, it can be considered that this is a phase discontinuity caused by information present in the signal and we call it the "phase jump point". The value of th_{pjp} should make the afore mentioned signals distinguishable in their phases, which means within the same period of collecting time, a discontinuous-phase signal has a much larger number of phase jump points than a continuous-phase signal does. Therefore, the sensed narrowband signal can be sorted into the discontinuous-phase or continuous-phase group by comparing the number of phase jump points N_{pjp} with a reasonable threshold th_{Npjp}. In Figure 20, the phase of a FM signal is compared with that of a BPSK signal.

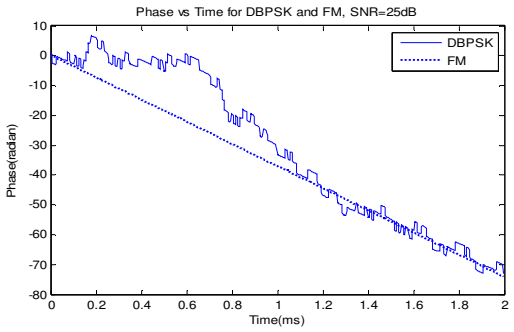

Figure 20: Time varying phase plot comparing FM and DBPSK

The above analysis provides us a theoretical range for the threshold of phase jump points $\mathrm{th_{pjp}}\left(\mathrm{e.g.}\, 0.1476\pi < t\mathrm{h_{pjp}} < 0.2048\pi\right)$, referring to formula (5). Next, we will deduce $\mathrm{th_{Npjp}}$ for the phase-based categorization.

Samples within one symbol contain the same phase information. The jumping occurs when the adjacent samples belong to two different symbols, if these two symbols represent different information. Thus, the number of phase jump points $\mathrm{N_{pjp}}$ out of the number of captured samples $\mathrm{L_s}$ equals to the number of symbol changes within the capture time $\mathrm{T_c} = \mathrm{L_s T_s} = \mathrm{L_s}/\mathrm{R_s}$. Let $\mathrm{T} = 1/\mathrm{R}$ (R is the symbol rate) denotes the symbol duration time. The number of symbols within $\mathrm{T_c}$ is $\mathrm{L} = \lceil \mathrm{L_s T_s}/\mathrm{T} \rceil$.

Because the number of symbol changes within an L-length symbol stream can be any value within a finite set at a certain probability, $\mathrm{N_{pjp}}$ is actually a discrete random variable with finite possible values. Based on the derivation in Appendix A, we get the ratio of $\mathrm{E}\left[\mathrm{N_{pjp}}\right]$ to $\mathrm{L_s}$ for MPSK and M-ary QAM as follows.

$$\frac{\mathrm{E}\left[\mathrm{N_{pjp}}\right]}{\mathrm{L_s}} \approx \left\lceil \frac{\mathrm{T_s}}{\mathrm{T}} \right\rceil \cdot \frac{\mathrm{M}-1}{\mathrm{M}}$$

When $\lceil \mathrm{T_s}/\mathrm{T} \rceil$ is fixed, the smaller M is, the smaller the value of $\mathrm{E}\left[\mathrm{N_{pjp}}\right]/\mathrm{L_s}$ is. Thus, the threshold $\mathrm{th_{Npjp}}$ should be less than $\lceil \mathrm{T_s}/\mathrm{T} \rceil \cdot \mathrm{L_s}/2$. Based on the analysis in the "Signal Capture" section, the sampling rate $\mathrm{R_s}$ is chosen to be several times the roughly estimated bandwidth B, which is output from the spectrum sensing module.

5.4.6.7 Constant vs. Varying Amplitude

The phase-continuous family mainly includes AM, FM, and CPFSK. We distinguish these three modulations by their different envelope characteristics. FM or CPFSK has a constant envelope, but the envelope of an AM signal changes with the modulating signal; the distribution range of an FM or CPFSK envelope is consequently much less than that of an AM. Thus, the envelope variance can be the criterion for the second-stage categorization shown in Figure 5. The corresponding threshold for this criterion will be derived as follows.

At large SNR, $r(t)$ can be simplified as $r(t) \approx A(t) + r_n(t)$. The envelope variance will be:

$$\mathrm{var}[r(t)] = \begin{cases} A_c^2 k_a^2 \mathrm{var}[m(t)] + \mathrm{var}[r_n(t)] & \text{for AM} \\ A^2 \mathrm{var}\left[\sqrt{m_R^2(t) + m_I^2(t)}\right] + \mathrm{var}[r_n(t)] & \text{for QAM} \\ \mathrm{var}[r_n(t)] & \text{for FM, CPFSK or MPSK} \end{cases}$$

Therefore, the envelope variance threshold th_{ev} can be chosen from the range of $\mathrm{var}[r_n(t)] < th_{ev} < var[A(t)] + \mathrm{var}[r_n(t)]$.

Next, we will calculate the instantaneous frequency of down-converted FM or CPFSK signal, which can be expressed by:

$$f(t) = \Delta f + f_d(t) + k_f m(t)\left\{1 - \frac{r_n(t)\cos[\theta_n(t) - \theta(t)]}{A_c}\right\} + \frac{n_Q'(t)\cos\theta(t) - n_I'(t)\sin\theta(t)}{2\pi \cdot A_c}$$

The above equation tells that $f(t)$ will be discontinuous due to the discontinuities of $m(t)$ in CPFSK, but this is not the case for FM. This difference between FM and CPFSK signals can be evaluated in terms of several different metrics:

(1) The ratio of instantaneous frequency's standard deviation (σ_f) to its mean value (μ_f), which is expressed by:

$$\frac{\sigma_f}{\mu_f} = \frac{\sqrt{E[(f(t) - E[f(t)])^2]}}{E[f(t)]} \underset{m}{=} \frac{\sqrt{\frac{1}{L_s - 1}\sum_{i=2}^{L_s}[f(i) - \mu_f]^2}}{\frac{1}{L_s - 1}\sum_{i=2}^{L_s} f(i)}$$

Where $E[\cdot]$ is the expectation value.

(2) The kurtosis of the normalized-center instantaneous frequency ($f_{NC}(t)$), μ_{42}^f, described in reference [43], can be used to measure the compactness of the instantaneous frequency distribution, and it is defined by:

$$\mu_{42}^f = \frac{E[f_{NC}^4(t)]}{(E[f_{NC}^2(t)])^2}$$

The ith sample value of $f_{NC}(t)$ is expressed as:

$$f_{NC}(i) = \left[f(i) - \frac{1}{L_s - 1}\sum_{i=2}^{L_s} f(i) \right] \Big/ R_s$$

FM's instantaneous frequency has higher compactness distribution than CPFSK's.

(3) The ratio of $ti(t)$'s standard deviation (σ_{ti}) to its mean value (μ_{ti}), which is expressed by:

$$\frac{\sigma_{ti}}{\mu_{ti}} = \frac{\sqrt{E[(ti(t) - E[ti(t)])^2]}}{E[ti(t)]} \underset{m}{=} \frac{\sqrt{\frac{1}{N_{ti}}\sum_{k=1}^{N_{ti}}[ti(i) - \mu_{ti}]^2}}{\frac{1}{N_{ti}}\sum_{k=1}^{N_{ti}} ti(i)}$$

In the above formula, $ti(t)$ is the time interval between adjacent zero-crossing points in an FM or CPFSK signal, and N_{ti} equals to the number of zero-crossing points in the captured samples minus 1. It is obvious that $ti(t)$ is inversely proportional to a signal's frequency. Thus, σ_{ti}/μ_{ti} can be used to discriminate between FM and CPFSK signals.

With UCS 1.0 prototype setup which is explained in Section 5.5, we measured the values of the above three metrics under different SNRs for GNU Radio GMSK (different symbol rates), Cobra Walkie-talkie FM, E. F. Johnson P25 C4FM and narrowband FM. Our experimental results tell us that σ_{ti}/μ_{ti} is much more reliable and distinguishable than the other two metrics with SNR

changing. Based on the measurement results, when SNR is greater than 4dB, the σ_{ti}/μ_{ti} value of FM signal varies in the range of (0.01, 0.4), while it is usually larger than 0.7 for GMSK and P25 C4FM signals. Therefore, the threshold value th_{freq}can be set as 0.5 if the SNR is not less than 4dB.

Theoretically, the order of an M-ary CPFSK signal can be obtained by the histogram of $f(t)$. The instantaneous frequency histogram of a C4FM signal is given in Figure 21.

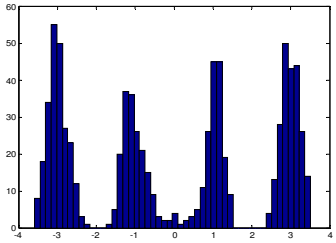

Figure 21: Instantaneous frequency histogram of C4FM

5.4.7 Bandwidth Estimation

Bandwidth estimation can provide a range for symbol rate estimation for a digital signal, as well as filter bandwidth for an analog signal. The accuracy of bandwidth estimation influences the efficiency of the entire system, especially symbol rate estimation time. In our system, we designed a histogram algorithm to replace the traditional -3dB bandwidth estimation method. For an MPSK or a QAM signal, which normally use raised-cosine pulse shaping, if the lower cutoff frequency for bandwidth calculation could start from the starting frequency of the pulse shaping, and the upper cutoff frequency ends at the ending frequency, then $B = (1 + \text{roll\_off})R$. Thus, given the roll off value, we can estimate symbol rate more accurately than by using -3dB bandwidth estimation method.

In Figure 22, we take QPSK as an example to illustrate. The intersection of the PSD plot and the straight line indicates the lower and upper frequency bounds of the desired signal spectrum. The joint point is where the PSD dramatically changes, and can be found by analyzing the histogram of PSD. Figure 23 shows the histogram of the PSD of a DBPSK signal. We use it as an example

to explain how to find SNR and the threshold from a PSD histogram. On the left side, there is a Gaussian like distribution; this is the histogram for noise. The abscissa of local maximum PSD indicates the mean of noise power, and its reciprocal is equal to the current SNR since the received signal is normalized. On the right side, the relatively centralized distribution is the signal. The straight line, where the locally minimal histogram number is, indicates the threshold for bw estimation.

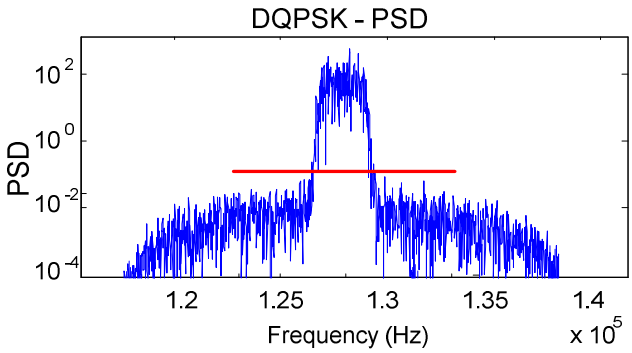

Figure 22: PSD for digital signal

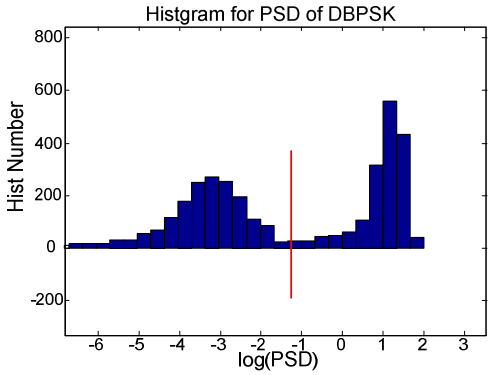

Figure 23: Histogram of DBPSK PSD

5.4.8 Symbol Timing and Coarse Classification

Symbol timing is a key technology in communication systems. It includes both symbol rate searching and symbol synchronization in our system. The accuracy of bandwidth estimation will determine the range of searching space.

Define the estimated bandwidth, which is the output of the Bandwidth Estimation module, as $\mathrm{bw}_{\mathrm{est}}$ (Hz), the sampling rate as $\mathrm{R_s}$ (Hz), the real bandwidth as bw, and true symbol rate as R. Estimated symbol rate is:

$$\mathrm{R}_{\mathrm{est}} = \mathrm{bw}_{\mathrm{est}}/(1 + \mathrm{roll\_off})$$

where roll_off is the roll off value of the root raised cosine filter used at the transmitter. The number of samples per symbol is expressed as:

$$\mathrm{sps} = \left\lfloor \frac{\mathrm{R_s}}{\mathrm{R}_{\mathrm{est}}} \right\rfloor$$

$\frac{\mathrm{R_s}}{\mathrm{R}_{\mathrm{est}}}$ may not be an integer. For the next step's processing convenience, the value of samples per symbol needs to be an integer. Thus, sps equals to the integer part of $\frac{\mathrm{R_s}}{\mathrm{R}_{\mathrm{est}}}$, and we need to resample the sample stream, which was collected at the original sampling rate $\mathrm{R_s}$, according to the redefined sampling rate $\mathrm{R}_{\mathrm{s}_{\mathrm{resample}}}$:

$$\mathrm{R}_{\mathrm{s}_{\mathrm{resample}}} = \mathrm{sps} \times \mathrm{R}_{\mathrm{est}}$$

The value of $\mathrm{R}_{\mathrm{est}}$ is not accurate enough to be used to calculate symbol rate R directly. We therefore need to analyze the accuracy of the symbol rate estimation to set a candidate space S for fine symbol rate estimation and symbol timing. Space S is determined by two factors: the maximum bandwidth estimation error and the tolerated error of the symbol rate when demodulating. Suppose 2l equals to the value of the maximum element in the candidate space S minus the value of the minimum element, and δ equals the difference between the two adjacent elements, then S is defined as:

$$\mathrm{S} = [\mathrm{R}_{\mathrm{est}} - \lfloor \mathrm{l}/\delta \rfloor \delta, \mathrm{R}_{\mathrm{est}} - \lfloor \mathrm{l}/\delta \rfloor \delta + \delta, \dots, \mathrm{R}_{\mathrm{est}}, \dots, \mathrm{R}_{\mathrm{est}} + \lfloor \mathrm{l}/\delta \rfloor \delta - \delta, \mathrm{R}_{\mathrm{est}} + \lfloor \mathrm{l}/\delta \rfloor \delta]$$

Define $\mathrm{R}_{\mathrm{err}}$as the maximum bias error between$\mathrm{R}_{\mathrm{est}}$ and R, i.e.

$$\mathrm{R}_{\mathrm{err}} = |\mathrm{R}_{\mathrm{est}} - \mathrm{R}|_{\max}$$

In order to find the correct symbol rate, we let $\mathrm{l} = \mathrm{R}_{\mathrm{err}}$, which guarantees $\mathrm{R} \in \mathrm{S}$.

Suppose the maximum tolerated symbol rate error for the subsequent synchronization is $\mathrm{err}_{\mathrm{tolerant}}$, then $\delta = 2\ \mathrm{err}_{\mathrm{tolerant}}$, such that

$$\mathrm{Min}(\mathrm{S(i)} - \mathrm{R}) < err_{\mathrm{tolerant}}$$

where S(i) is the ith element of S. So far, the candidate space S is defined. For each element of S, we have a new resampling rate: $\mathrm{Re(i)} = \mathrm{sps(i)} \times \mathrm{S(i)}$, where $\mathrm{sps(i)}=\lfloor \mathrm{R_s}/\mathrm{S(i)} \rfloor$. Re is a vector in which each element represents the resampling rate corresponding to each element in S . Next, we are going to explain how we search for the best symbol rate and finish symbol timing synchronization at the same time.

Figure 24 shows a snapshot of samples for a DBPSK signal at quasi-baseband, which was collected by the Anritsu Signature signal analyzer in the OTA experiment. There are 8 samples per symbol. Dots indicate the sampling points. Red points indicate the correct symbol timing. Black points indicate the incorrect symbol timing. As we can see, only when both symbol rate and the symbol timing moment are correct, the chosen samples have small variance, while the other set has relatively large variance. We need to calculate two parameters from this result: number of samples per symbol and timing position within a symbol.

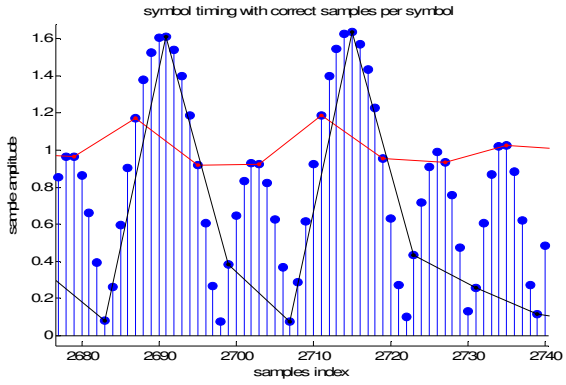

Figure 24: Illustration of symbol timing impact to the received signal

Samples_V is defined as the vector including the quasi-baseband complex samples collected at sampling rate $\mathrm{R_s}$ within capture time $\mathrm{T_c}$.

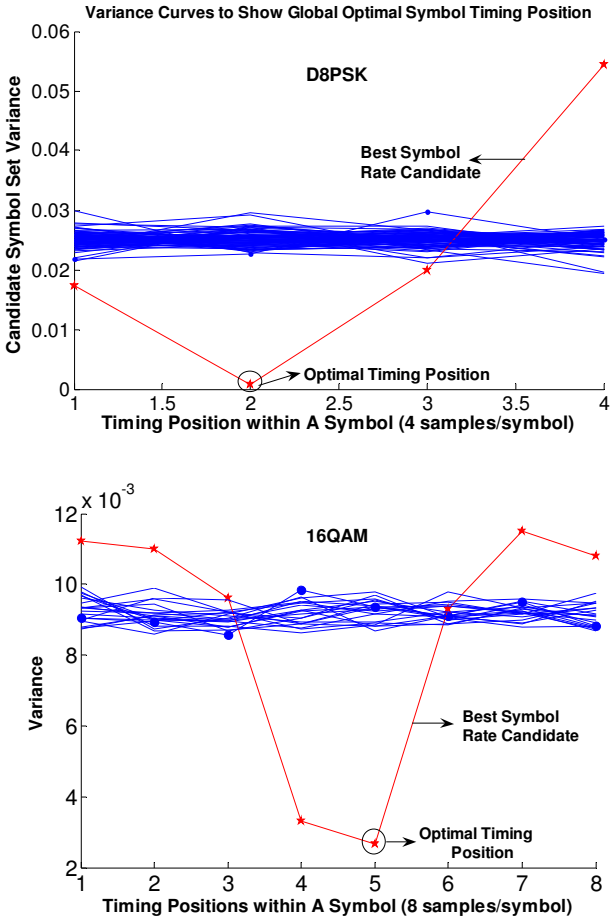

Figure 25: Variance curves implying global optimal symbol timing position

Each element of space S is a candidate for the correct symbol rate. Resamples_V_i is the samples stream after resampling Samples_V by Re(i). SS is defined as the space for candidate symbol set. The number of elements in SS is $\sum_{i=1}^{2\lfloor l/\delta \rfloor+1} \text{sps(i)}$.

Each element of SS is a vector, which is expressed by

$$\text{SS}\left(\sum_{k=1}^{i-1} \text{sps(k)} + j\right) = \left\{\text{Samples\_V(z)} \middle| \left(z \bmod \text{sps(i)}\right) = j\right\}$$

where $i = 1,2,\ldots,2\lfloor l/\delta \rfloor + 1$, $j = 1,2,\ldots,\text{sps(i)}$.

SS is the candidate space for optimal symbol timing. Each element of SS is potentially a correctly sampled symbol set. Our target is the optimum one. Before searching for the optimal symbol set, we need to first distinguish QAM from MPSK.

QAM and MPSK can be differentiated by analyzing their envelopes. The desired envelope of MPSK symbols is a single constant value, and the desired envelope of QAM is a set of several constant values. In other word, if we cluster samples' envelope in each element of SS, and the clustered result is centralized around one constant value, then it is MPSK. If not, then according to the number of centralized values, we can classify the samples as 16QAM (3 values), 64 QAM (9 values), etc. The number of the centralized values is called the envelope order. Because only one element of space SS is the correctly sampled symbol set, if the received signal is MPSK, for example, other elements' envelope order may be equivalent to or greater than 1 due to the incorrect sampling. Each element's envelope order is calculated and saved in a vector called $\mathrm{Envelope}_{\mathrm{order}}$. Based on the clustering result, we calculate the variance for each of the symbol clusters and add them to get the total variance. The variance is saved in vector $\mathrm{Var}_{\mathrm{SS}}$. The minimal value of this vector is marked as $\min\left(\mathrm{Var}_{\mathrm{ss}}\right)$

Sampling at the right position will guarantee the highest SINR (Signal to Interference plus Noise Ratio). We therefore consider that $\mathrm{SS}(\min\left(\mathrm{Var}_{\mathrm{ss}}\right))$ has the best symbol timing. The corresponding symbol rate R is found at the same time. This is shown in Figure 25.

$\mathrm{SS}(\min\left(\mathrm{Var}_{\mathrm{ss}}\right))$ and R will be fed to the next module, as well as the $\mathrm{Envelope}_{\mathrm{order}}(\min\left(\mathrm{Var}_{\mathrm{ss}}\right))$, which is important for information removal in carrier synchronization.

In Section 5.4.2, we mentioned that the sampling rate is related to symbol timing. We will explain how to select the sampling rate. A digital signal can be expressed as:

$$\mathrm{v(t)} = \sum_{\mathrm{n}} \mathrm{I_n g(t - nT)}$$

Here $\mathrm{g(t)}$ is a pulse shaping function. Thus,

$$\mathrm{g(t - nT)} = \begin{cases} \mathrm{g(t)} & (\mathrm{nT} < t < (n+1)T) \\ 0 & \text{otherwise} \end{cases}$$

With a sampling interval $T_s = 1/R_s$, sampling instants are $t_0 + kT_s, k = 1,2, \ldots$, and the sampling values are:

$$v(t_0 + kT_s) = \sum_n I_n g(t_0 + kT_s - nT)$$

Let $a = T/T_s$, then we have

$$v\left(t_0 + \frac{kT}{a}\right) = \sum_n I_n g(t_0 + (\frac{k}{a} - n)T)$$

Because of the re-sampling strategy we applied, the kth resampling value is:

$$v_{resample}\left(t_0 + \frac{kT}{m}\right) = \sum_n I_n g(t_0 + (\frac{k}{m} - n)T)$$

Here m is an integer and $m = \lfloor T/T_s \rfloor / T \times T = \lfloor a \rfloor$.

For the pth element of SS, the variance is:

$$var(p) = E\left[\left(\sum_n I_n g(t_0 + (\frac{q}{m} - n)T)\right)^2\right] - E^2\left[\sum_n I_n g(t_0 + (\frac{q}{m} - n)T)\right]$$

Here $q \bmod sps(i) = j$ and $\sum_{k=1}^{i-1} sps(k) + j = p$.

After calculation, we have:

$$var(p) = \begin{cases} g^2(t_0 + \frac{\Delta}{m}T) & q = lm + \Delta, l = 1, 2, \ldots \text{ and } \Delta \text{ is an integer} \\ var[g^2(t)] & \text{otherwise} \end{cases} \quad (6)$$

To find out the correct symbol timing, the following inequality has to be met:

$$\max (g^2\left(t_0 + \frac{\Delta}{m}T\right) > var[g^2(t)] \quad (7)$$

To satisfy the above condition, m needs to be greater than a certain threshold. Take a raised cosine pulse shaping function for example as in Figure 26.

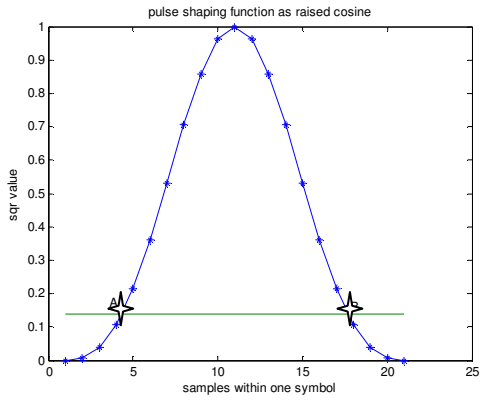

Figure 26: Pulse shaping of raised cosine function

So long as one of the samples in a symbol is located between A and B, then, formula (7) can be guaranteed. Thus, for raised cosine pulse shaping function, 2 samples per symbol is enough for symbol timing recovery. Considering both the condition in formula (7) and the accuracy of bandwidth estimation, we usually select a sampling rate 3 or 4 times of estimated bandwidth.

5.4.9 Carrier Synchronization and Fine Classification

When the modulation types and parameters are known, a Phase Lock Loop (PLL) can be used to accomplish carrier synchronization. This is necessary to overcome equipment limitations. For example, typical of low cost SDR platforms, the USRP has an inaccurate oscillator which will generate a frequency offset – in this case less than 2 kHz. Phase offset is caused by noise, channel delay, and different phase references in the transmitter and receiver.

In the UCS scenario, both the order of MPSK and the center frequency are unknown. As a result, the methods described above will not produce accurate results. In [61], we designed the Universal Synchronization Algorithm, which doesn't need to know the order of MPSK. The key idea is that we remove the information before the symbols get into the PLL. The carrier synchronization architecture is composed of five parts: information removal, frequency estimation, frequency rectification, phase estimation, and phase rectification.

In M-ary PSK modulation, the amplitude of the transmitted signal is constrained to remain constant, thereby yielding a circular constellation [59]. In addition, the M constellation points of an MPSK signal uniformly distribute on a single circle. This means the phase difference, contributed by the information-bearing elements, between any two adjacent symbols of a MPSK signal, can be represented as $n \cdot 2\pi/M$, where n is an integer and M is the order of a MPSK signal. Therefore, it is convenient to use a phase-based method to remove information for MPSK. However, in QAM the amplitude also varies along with the phase. The constellation points of a QAM signal uniformly distribute on squares. Therefore, it seems that a QAM signal does not possess the same phase features as MPSK, and we'll need a square-slicer to cluster the symbols. We observe that the constellation points of a QAM signal also lie on circles. For example, in ideal conditions, the constellation points of a 16QAM signal distribute on three circles, while 64QAM has nine circles. For 16QAM, the phases of points on the inner (the 1st) and outer (the 3rd) circles have the same distribution as a QPSK signal. Therefore, we can select the symbols on the 1st and 3rd circles of 16QAM, or the 1st, 3rd, and 9th circles of 64QAM for phase-based information removal. When the SNR is sufficient, the symbol timing and coarse classification module will provide an accurate envelope order, and a properly designed envelope-slicer will pick out the desired points for information removal.

The loop gain of the PLL is a critical parameter for which we need to account. The PLL will not converge if the loop gain is less than the actual frequency offset Δf. When the value of loop gain is less than $3\Delta f$, the PLL will converge on an accurate result and output a frequency offset estimation. Therefore, in UCS, we use multiple iterations to achieve better estimation precision and adaptive loop gain to ensure the convergence of our algorithm, called "Multiple-iteration frequency tracking algorithm with loop gain adaptation." This algorithm is implemented by a while loop. At each iteration of the while loop, the loop gain is updated on the basis of the just estimated $\mathrm{delta\_freq}$. This adaptive scheme improves the robustness and reliability of our frequency tracking algorithm. For the while loop, we chose the maximum iterations $\mathrm{max\_iteration}$ and, variance of the phases of the information-removed points $\mathrm{var\_inphi}$ as the criteria. With the increase in the number of iterations, the residual frequency offset should become smaller and smaller. Thereby, where $\mathrm{var\_inphi}$ develops a trend of reduction and its

value becomes less than a threshold th_{var_inphi}, we can attain clear constellation points. The setting of th_{var_inphi} is based on the result presented in[62]. th_{var_inphi} is related to the symbol rate and SNR.

The final result for frequency offset estimate is the ultimate value of freq_corr,which has been updated by freq_corr = freq_corr + delta_freq(j) at jth iteration of the while loop. Applying the above schemes to our carrier frequency offset estimation algorithm, the error between freq_corr and the actual frequency offset is within 1 Hz. Finally, we can get a clear constellation diagram.

Figure 27 shows the constellation diagrams for 8PSK and 16QAM at different processing stages: the first column figures display the snapshots of constellation points (samples) before symbol timing. In the middle, the symbol stream output from the symbol timing module is plotted. The third column contains the constellation diagrams for synchronized samples.

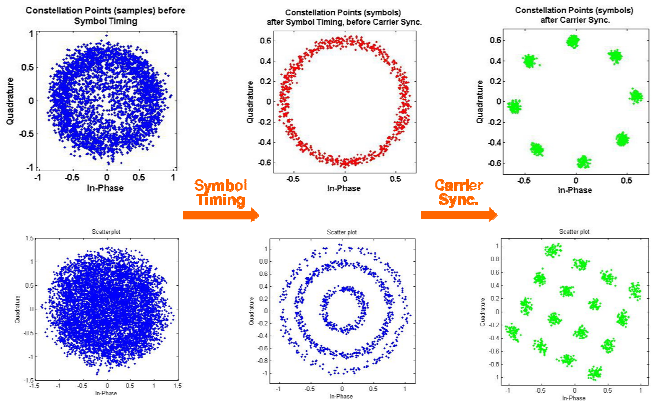

Figure 27: Result of UCS (a|b|c)

From the instantaneous phase distribution histogram of a carrier-synchronized signal, we can easily estimate the bits per symbol of MPSK or M-QAM, as shown in Figure 28. Thus, the fine classification after carrier synchronization has been achieved.

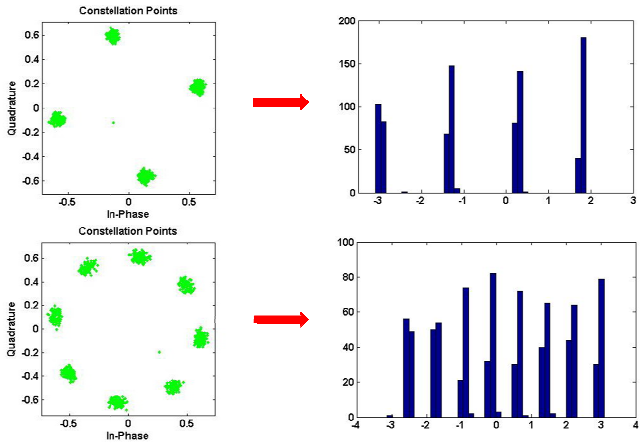

Figure 28: Fine classification based on instantaneous phase distribution histogram

5.4.10 OFDM Signal Scenario and Application

From this section, we are going to discuss the other branch: wideband signals. As we mentioned in Section 5.4.5, in our current stage, we only include OFDM signals in our system. More detailed information about OFDM signal classification and synchronization can be found in [58]. From Section 5.4.10 to Section 5.4.10.2, extracting parameters from OFDM signal for demodulation is illustrated.

OFDM's flexibility in bandwidth is particularly suited to the DSA scenario. Two schemes can be used to change an OFDM signal's bandwidth[63]. One method is to turn off certain subcarriers, which is the scheme applied in Orthogonal Frequency-Division Multiple Access (OFDMA). The other method reduces the subcarrier width and inter-subcarrier spacing, allowing the signal to adapt to dynamically available bandwidth while maintaining a constant number of subcarriers. The symbol rate adaption is controlled by the bandwidth of one subcarrier. Both methods have the same effect on bandwidth and data throughput as shown in Figure 29. We assume that OFDM signals that can be identified by UCS use the second method. By keeping the number of subcarriers constant, it allows us to reduce computational complexity. For standard applications including 802.11a, 802.11g, 802.11n of Wi-Fi family and WiMAX, which use OFDM or OFDMA, a matched filter can be used for identification.

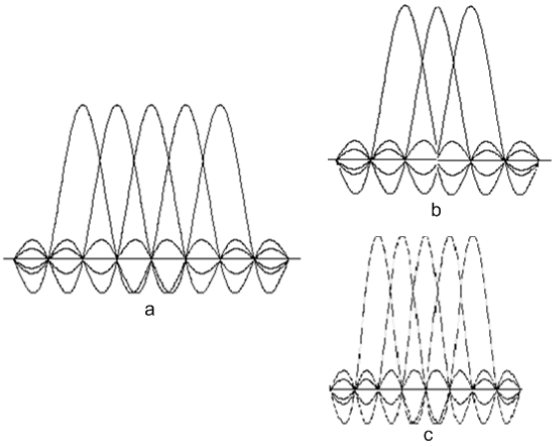

Figure 29 : Original OFDM signal and two schemes for changing the bandwidth of an OFDM signal

Figure 30 shows an overview of OFDM branch after wideband/narrowband classification. The process detects the start and end of a single OFDM symbol, measures the length of the CP, compensates the frequency offset, adjusts the symbol timing, and analyzes the subcarrier modulation type and settings.

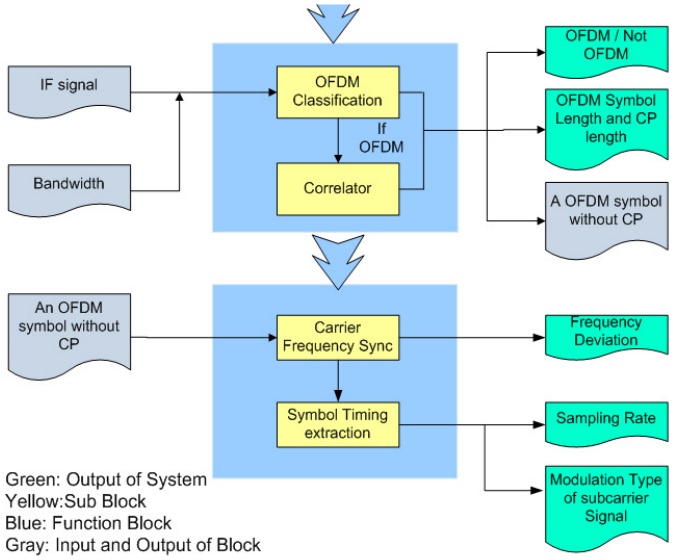

Figure 30: Overview of OFDM synchronization and parameter extraction

5.4.10.1 Estimation of Symbol Length and CP Length

By correlating the incoming signal with itself, we are able to separate an OFDM symbol into a data part and a cyclic prefix part. From Figure 4, we observe that the OFDM plot has three distinct peaks. The two smaller peaks are due to the presence of the cyclic prefix. The length of

the data part of a symbol is the distance between the highest peak and the smaller peak. The number of samples between these two peaks is defined as n_{rx} ; then the data part duration at the receiver side is n_{rx}/R_s, where R_s is sampling rate. For finding the CP length, we convolve the data part of an OFDM symbol with the rest of the OFDM symbol as shown in Figure 31. The CP' copy in data part will overlap with CP. This will result in a peak as shown in Figure 32. The position of the peak determines the length of the CP.

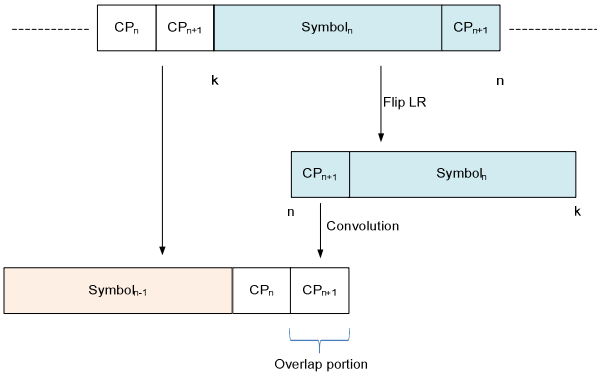

Figure 31: Estimation of CP length

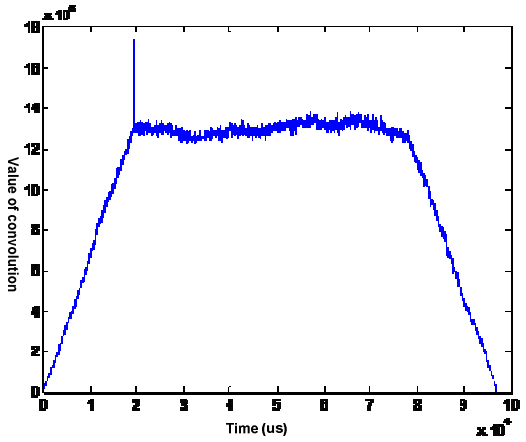

Figure 32: Convolution of symbol and cyclic prefix

Meanwhile, we also get the data part of an OFDM symbol. This part is called V_{rx}. When we use the data part of an OFDM symbol for carrier synchronization and per-subcarrier symbol timing, it is necessary to have an integer numbers of samples per MPSK signal symbol which is acquired after FFT and parallel-to-serial conversion. To meet the requirement of integer number of samples per symbol, we need to resample V_{rx}.

Figure 33 shows the serial to parallel (S/P) processing at the transmitter side. t_{tx} is symbol duration before S/P,$t_{tx} = 1/R_t$, and R_t is the symbol rate at the transmitter side. F_s is the number of subcarriers. Thus, the data part duration is F_s/R_t.

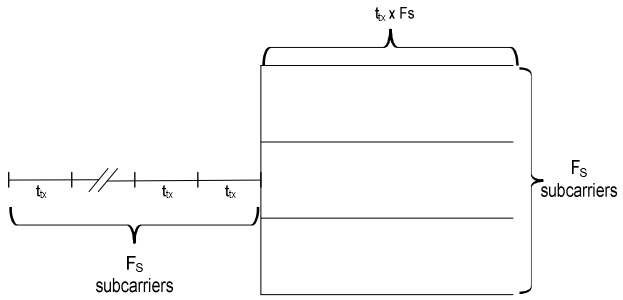

Figure 33: Serials to parallel of OFDM in transmitter side

The data part duration at the transmitter side is the same as that at the receiver side. Thus, we have:

$$\frac{n_{rx}}{R_x} = \frac{F_s}{R_t}$$

$\frac{n_{rx}}{R_x}$ represents average samples per symbol and is determined by the value of R_x and R_t , and cannot be guaranteed an integer. Thus, we resample vector V_{rx} and the number of elements of V_{rx} becomes $\text{round}(n_{rx}/F_s)F_s$ after the resampling. New number of samples per symbol is $round(\frac{n_{rx}}{F_s})$. The new vector after resampling is V_{resample}.

5.4.10.2 Carrier Synchronization for OFDM Signals

For an OFDM signal, carrier frequency offset has a completely different influence on the symbol constellation as compared to a narrow band signal. Figure 34 shows the effects of frequency offset on an OFDM signal and on an MPSK signal.

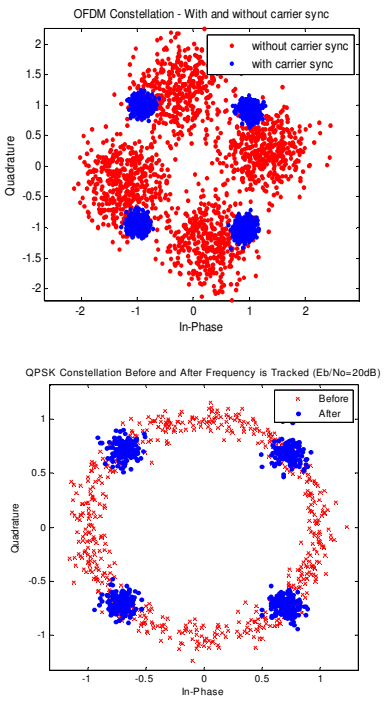

Figure 34 : Comparison between the effect of frequency offset on OFDM signal and QPSK signal

In Figure 34, the comparison is based on the assumption that the symbol rate is correct. Frequency offset causes an OFDM constellation to spread with errors in both amplitude and phase. The frequency offset introduces only phase errors in MPSK signals. The reason for this difference is that frequency offset becomes a timing delay after FFT at the OFDM receiver. This delay will lead to incorrect symbol timing, which will spread the signal constellation.

The frequency offset estimation algorithm is designed as follows. The step size of the frequency offset is defined as Δf, and the range of the frequency offset estimated is defined as f_{range}. We search from $-\mathrm{f}_{\mathrm{range}}$ to f_{range} using step size Δf. As we compensate for frequency offset, the variance of the amplitude of the symbols changes. The minimal variance corresponds to the carrier frequency offset.

5.4.11 Verification Schemes for UCS System

Low SNR decreases the accuracy of both the bandwidth estimation and of the phase values of the received samples. The former will increase the range of searching space in symbol timing step, and the latter will decrease the synchronization accuracy. Both of them will increase the average time consumed for one iteration of correct signal recognition. In our system, which is distinguished from other communication systems operating with bit errors, the final result can only be either right or wrong. Thus, a verification scheme is necessary for our system. There are three parts of verification embedded in our system: noise versus signal verification, symbol timing verification (bandwidth estimation verification), and carrier synchronization verification. If a result does not pass the aforementioned verification, it will be treated as incorrect, and the system will automatically discard the results and recollect data for recalculation. Therefore, the introduction of verification will in fact increase the average time consumed per correct calculation, although it can provide a high correct rate. Accordingly, we can convert the tradeoff of SNR versus correct rate into the tradeoff of SNR versus average iteration time by always keeping the correct rate high. This section involves two parts; the first is to explain the verification algorithm and the other to discuss the relationship between SNR and average consuming time.

Noise versus signal verification is located right after channel estimation. The objective of this function is a pre-decision on whether the captured data will be successfully processed by the system. The failure case occurs when only noise has been captured or the collected signal has a SNR lower than system tolerance. In Figure 23, the ration between the right peak position value and the left peak position is considered as Expected SNR (ESNR). Only a signal with ESNR larger than a certain threshold will be considered. We use ESNR instead of SNR because ESNR is not influenced by the oversampling rate, while SNR is. The threshold is determined by system running environment and user requirement.

The second verification is symbol timing verification. In the symbol timing section, we defined the range for searching. And if it has reached this range, but no satisfactory solution has been found, we drew the conclusion that the data should be discarded and new data should be

collected. This part of verification only existed in narrowband signal branches, and not for OFDM.

The third verification is for carrier synchronization. We define a threshold for loop iterations. If the loop iterations have reached the threshold and not satisfied the criterion, then we concluded that the data should be discarded and new data should be collected.

5.5 UCS Prototypes and Performance Evaluation

UCS has been implemented and tested over the air. The UCS 1.0 implemented in the Anritsu MS2781A Signature Signal Analyzer focuses on demonstrating the classification and synchronization functions of UCS. The UCS 2.0 transplants the system to GNU radio and USRP platform, which dramatically decrease the equipment expense. In UCS 3.0, now under development, the system will be transplanted to a Lyrtech SFF radio platform, which integrates DSP and FPGA design and can greatly speed the entire UCS process. The OTA demonstration setup for UCS 1.0 and UCS 2.0 is shown in Figure 35 and Figure 36. In this section, we focus on UCS 2.0.

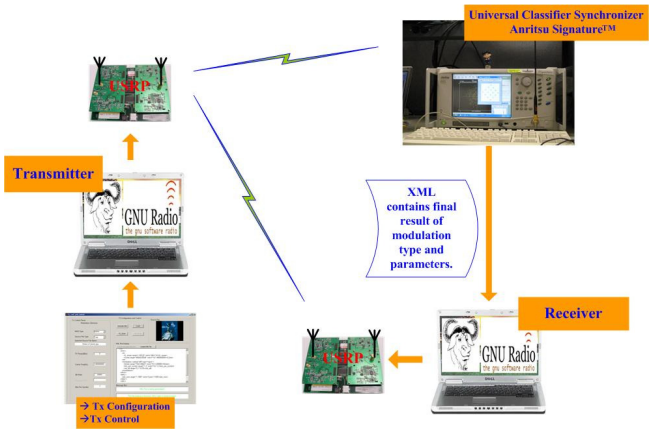

Figure 35 : OTA demo setup for UCS 1.0

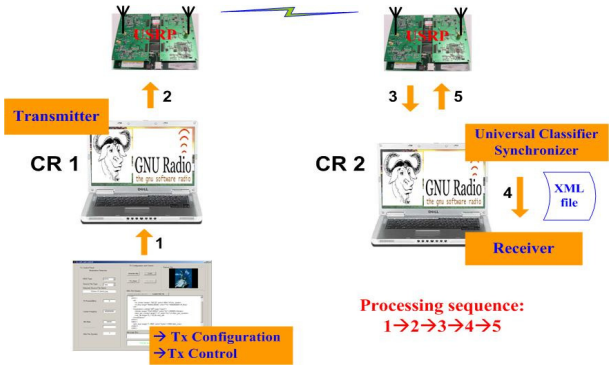

Figure 36: OTA demo setup for UCS 2.0

CR1 and CR 2 are two cognitive radio nodes based on GNU Radio with USRP as their hardware platform. Narrowband signals, including FM, AM, MPSK (M=2,4,8), QAM(16) and OFDM are transmitted by CR1. CR2 will receive the signal over the air from CR1 via the antenna. All the information required for the CR2 to correctly demodulate the signal is extracted by the UCS algorithm and stored in XML format. The updated XML file will be used to (re)configure the radio framework of CR2. Then the connection between CR1 and CR2 can be created, and they commence communications.

How to assess a system is at least as important as actually building it. We evaluate our system from three aspects: accuracy, SNR requirement and time consumption. The decision making in our system is step by step, thus, our analysis and evaluation are also step based. Next, we will give the performance curves for the four key steps in UCS system: wideband/narrowband, narrowband categorization, symbol timing, and carrier synchronization.

Figure 37 shows the error probability of wideband/narrowband signal differentiation under different SNR with different product values of sampling rate and guard interval. As we can see, it is related to the number of samples within a cyclic prefix period. The larger the product of sampling rate and guard interval is, the lower the error rate is under the same SNR. Thus, we can conclude that normally, the SNR requirements for correctly differentiating between narrowband and wideband signals with high probability are not tight, and thus easy to meet.

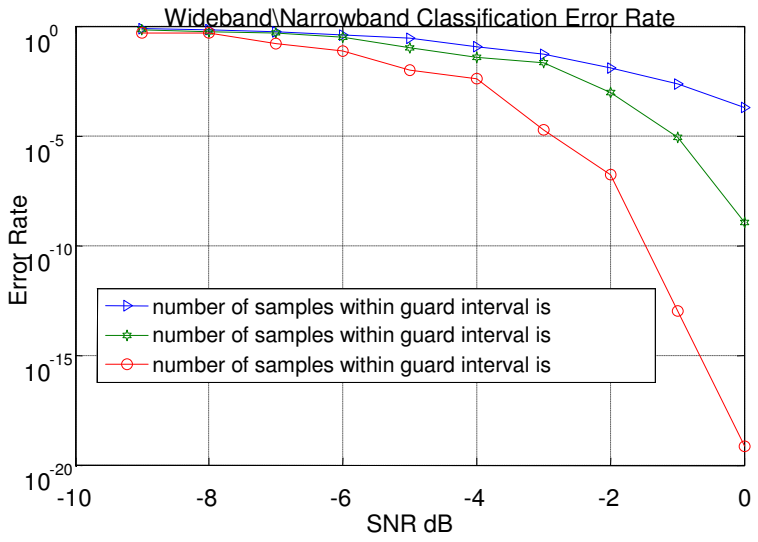

Figure 37: Wideband/Narrowband error detection probability

As seen in Table 6, narrowband signal categorization includes several steps. Here, we mainly analyze the performance of phase-based grouping at different SNR. The probability of mistaken continuous-phase/discontinuous-phase differentiation P is:

For mistaken continuous-phase to discontinuous-phase:

$$P = \sum_{k=th_{Npjp}}^{L_s} \binom{L_s}{k} p1^k (1-p1)^{L_s-k}$$

For mistaken discontinuous-phase to continuous-phase:

$$P = \sum_{k=aL_s-th_{Npjp}}^{aL_s} \left\{ \binom{aL_s}{k} {p_2}^k (1-p_2)^{aL_s-k} \left[\sum_{j=0}^{th_{Npjp}+k-aL_s} \binom{(1-a)L_s}{j} {p_1}^j (1-p_1)^{(1-a)L_s-j} \right] \right\}$$

Here, aL_s is the average number of jump points for each modulation type, p_1 is the probability that a non-jump point is mistaken to be a jump point, p_2 is the probability that a jump point is mistaken to be non-jump point. As explained in Figure 24, p_1 and p_2 can be expressed as:

$$\begin{cases} p_1 = f(th_{pjp}) \\ p2 = f(\beta - th_{pjp}) \end{cases}$$

Where β is the angle for a jump point and $f(\alpha)$ is a function that

$$f(\alpha) = \begin{cases} 1 - (P_r(\tan\alpha) + \int_{\tan\alpha}^{\infty} f_r(x) \dfrac{\arccos\left(\frac{\sin\alpha}{x}\right)}{\pi} dx) & \alpha < \dfrac{\pi}{2} \\ \int_{1}^{\infty} f_r(x) \dfrac{\arccos\left(\frac{\sin\alpha}{x}\right) + \frac{\pi}{2} - \alpha}{\pi} dx & \pi > \alpha \geq \dfrac{\pi}{2} \end{cases}$$

In Figure 38, we only provide the condition of $< \frac{\pi}{2}$; other conditions are similar to it. $P_r$is the cumulative distribution function of a Rayleigh distribution with $\sigma^2 = \mathrm{SNR}$, and f_r is the probability density function.

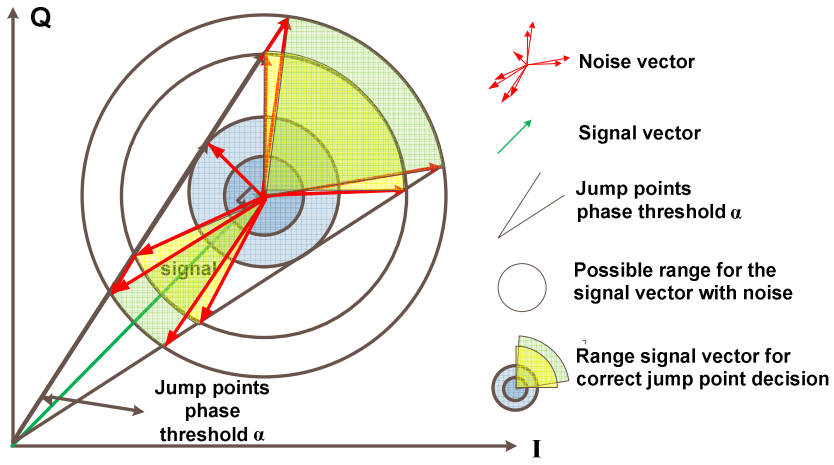

Figure 38: Illustration for probability of mistaking jump/non-jump points decision caused by noise

In our system, $\mathrm{th}_{\mathrm{pjp}} = 0.2\pi$ and $\mathrm{th}_{\mathrm{Npjp}} = L_s/8$. Figure 39 shows the distribution of p_1 and p_2 under different SNR and Figure 40 shows the performance of the continuous-phase/discontinuous-phase classifier under different SNR. We can see that for digital signals, the SNR requirement for this step is low; however, for analog, an error rate less 10^{-3} than requires a SNR larger than 8 dB.

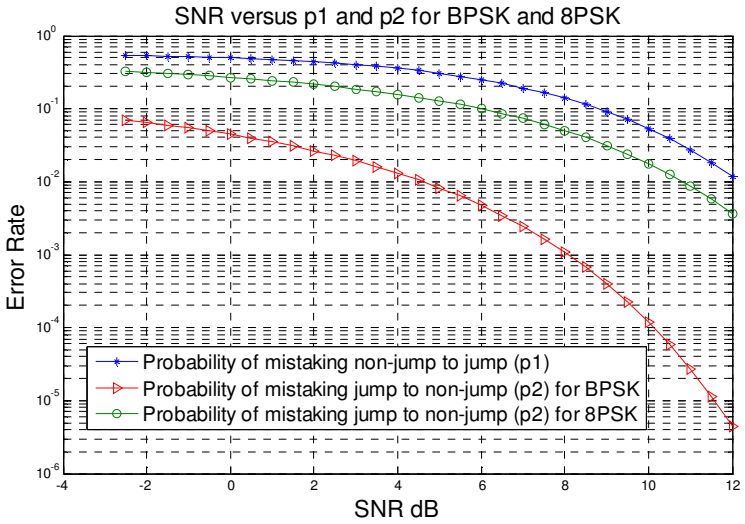

Figure 39: Probability of mistaking jump/non-jump points

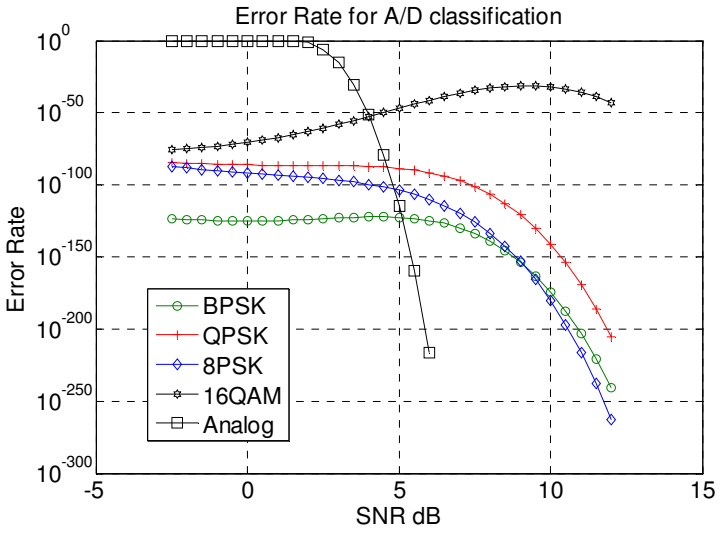

Figure 40: Probability of mistaken continuous-phase/discontinuous-phase differentiation

Symbol timing accuracy depends on sampling rate and pulse shaping. We take a raised cosine for an example. Figure 41 shows the symbol timing correctness under different SNR conditions for a narrowband signal. For OFDM signals, because the symbol timing determination is combined with the step of narrowband/wideband differentiation, no additional error rate will be added in this step.

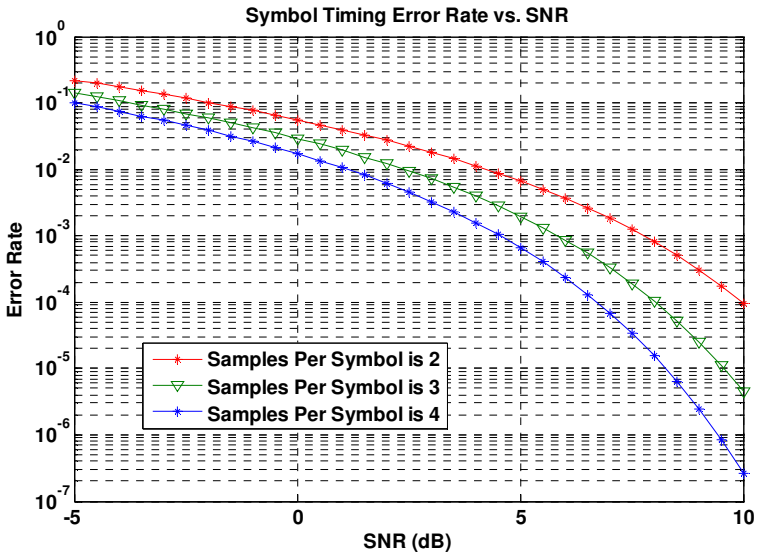

Figure 41: Symbol timing error rate for narrow band signal

To explain the narrowband case, we need to recall the analysis in Section 5.4.8. We state that if $\max\left[\, g^2\left(t_0 + \frac{\Delta}{m}T\right)\right] > \beta \cdot Est(var[g^2(t)])$, the symbol rate is acceptable. Est(var$[g^2(t)]$) is the estimation of var$[g^2(t)]$. In real system, we use the variance of all the samples as the estimation of var$[g^2(t)]$. Because of the inaccuracy of the estimation, a tolerant parameter β ($\beta > 1$) is used in order to guarantee the probability of $\beta \cdot \text{Est}(\text{var}[g^2(t)]) < var[g^2(t)]$ is small.

Considering an AWGN channel with a distribution of$N(0, \sigma^2)$, as seen in formula (6), when correct symbol timing, $\text{var}(p)/\sigma^2$ is a non-central chi square distribution.

Define $y = \max\left[\, g^2\left(t_0 + \frac{\Delta}{m}T\right)\right]$ and $x = t_0 + \frac{\Delta_1}{m}T$, where $g^2(x) = \max\left(\, g^2\left(t_0 + \frac{\Delta}{m}T\right)\right)$. We can see that, x is uniform distribution in the range of $[0, \frac{T}{m}]$. Then, given x, the conditional probability that when the estimated symbol rate is correct and $\frac{\max\left(g^2\left(t_0+\frac{\Delta}{m}T\right)\right)}{\sigma^2} < \frac{\beta \cdot \text{var}[g^2(t)]}{\sigma^2}$, which means the conditional error rate for symbol timing is

$$P_{condition}(x) = F_{ncx2}(\beta * \text{var}[g^2(t)]/\sigma^2)$$

$F_{ncx2}(\cdot)$ is a non-central chi square distribution cumulative distribution function with parameters $k = 2$ and $\lambda = g^2(x)/\sigma^2$.

Thus, the probability for incorrect symbol rate estimation is

$$P = \frac{m}{T} \int_{x=0}^{x=\frac{T}{m}} P_{condition}(x)dx$$

Here, SNR represents the ratio between signal power and noise power, expressed as a numerical ratio and not in dB. In Figure 27, $\beta = 1.4$ and a raised cosine filter with roll off $= 0.35$ is used, which are also the settings we adopted in our implemented system.

As we can see in Figure 27, the number of samples per symbol does influence the performance. In our implemented system, we use samples per symbol between 3 and 4. For QAM signal, the calculation for variance is based on cluster. SNR requirement will be larger than 5dB in order to reach error rate less than 10^{-3}. Figure 27 doesn't include verification, and our implemented system with verification has lower SNR requirement, around 3dB to 4 dB.

As we described in Section 5.4.8, MPSK and M-ary QAM signals are distinguished by analyzing the envelope values of the symbol stream that is output from the Symbol Timing module. For 16QAM, its constellation points lie on three circles. From the inner circle to the outer circle, we use the 1st, 2nd, 3rd to denote the corresponding circle, respectively. Therefore, sorting the 16QAM symbol points (complex) into circles is a three-hypothesis testing problem. After derivation, we get the performance curves plotted under different SNR with different threshold values ($[0.50L, 0.52L, 0.54L, 0.56L, 0.58L, 0.60L]$), as shown in Figure 28. L is the number of symbols. Without knowing the modulation type, these symbols are clustered into three groups no matter the incoming signal is MPSK or 16QAM. For symbol clustering, the normalized complex symbols will first be scaled to match the three circles' radius of an ideal 16QAM signal. The clustering principles are related to the symbol envelope r:

$$\begin{cases} \text{sort into the 1st circle,} & n1 = n1 + 1 \quad \text{when } r \le \sqrt{6} \\ \text{sort into the 2nd circle,} & n2 = n2 + 1 \quad \text{when} \sqrt{6} < r \le \sqrt{14} \\ \text{sort into the 3rd circle,} & n3 = n3 + 1 \quad \text{when } r > \sqrt{14} \end{cases}$$

$n_i (i = 1,2,3)$ denotes the number of symbols that are sorted into the ith circle range. Thus, if $n2 \geq th$, the signal is judged as MPSK, otherwise, it will be 16QAM. Figure 42 tells that the threshold th should be changed with the varying of SNR to meet corresponding requirement to error probability.

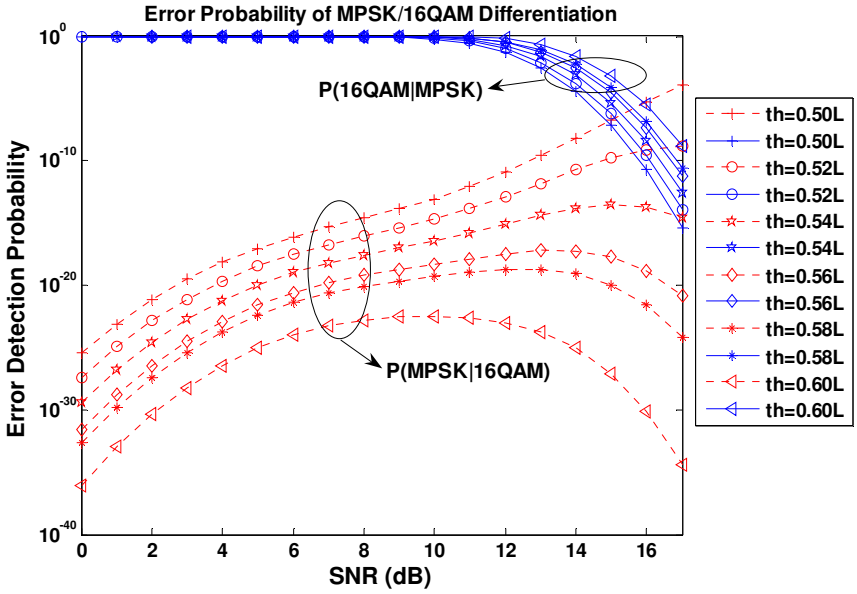

Figure 42: Probability of mistaken MPSK/16QAM differentiation

Carrier synchronization for narrow band signals is required to compensate the frequency offset so that the symbol stream has the least phase variance. It has the same performance as a standard phase lock loop [57, 64, 65]. For OFDM signal, it is to find the least amplitude variance. The searching step of a frequency offset matters to the accuracy of the frequency offset, as well as to subcarrier synchronization. In our implemented system, we use a frequency offset step of 100 Hz .

Before we proceed to entire system evaluation, we would like to discuss hardware issues. Compared to expensive radio devices (for example, Anritsu signal analyzer we used for our first generation system), the RF filter, low noise amplifier, IF filters, local oscillator and analog to digital converter found on the USRP board are inexpensive components. The USRP board is influenced by temperature or other factors caused by running for a long time. Using improved devices will achieve better performance. However, the verification scheme would compensate for the degradations, with a sacrifice in running time.

What can we conclude the performance of the entire system? It has to be noted that when we evaluated above performance step by step, we did not count in verification scheme. In our implemented system, as we mentioned, because of the unstable hardware performance, verification is an important factor. In our OTA experiment, we do include verification scheme. The verification scheme generates a rate of improper classification from 10^{-4} to 10^{-3} when SNR is larger than 8dB and is within acceptable average time per iteration. An error is defined as that any parameters are not correctly extracted and cause a failure for demodulation. According to our experiment, SNR is less related to error rate, but more related to average running time. And in the real scenario, average running time matters significantly to the users. Thus, instead of giving the error rate, we provide the experiment result of SNR versus average running time. The details are shown in Figure 29.

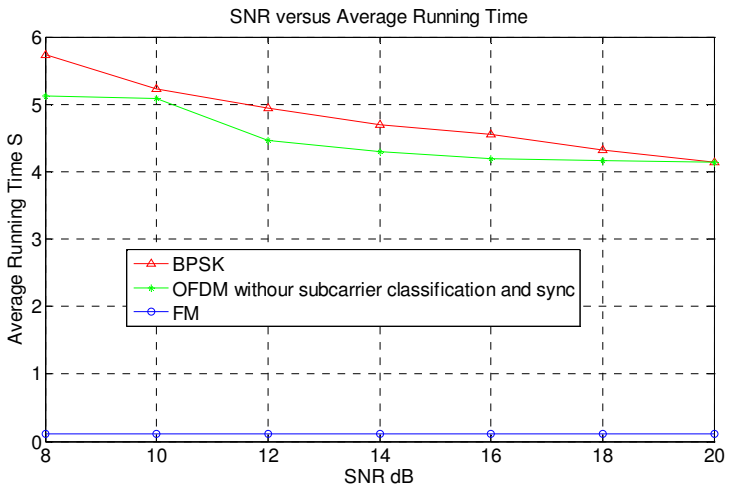

Figure 43: Average running timing under different SNR conditions

In Figure 43, we give the approximate average time for three types of signals. OFDM running time doesn't include single-carrier classification and synchronization time and it is based on simulation. The digital and analog curves are based on OTA experiment.

5.6 Conclusion

In this paper, we have discussed the design and implementation of a universal classifier and synchronizer, and evaluated its performance in terms of accuracy, SNR, and time consumption.

It has been verified in both theoretical analysis and practical experiments that the system works with high accuracy. The UCS prototype developed over an inexpensive SDR platform (USRP plus GNU Radio) can observe the environment, make decisions and autonomously act on those decisions. With the aforementioned advantages, a UCS can serve as an accurate sensor in a complex, distributed cognitive radio or network, or act as an independent physical layer cognitive receiver. UCS can be easily imported to other SDR platforms (such as Lyrtech SFF) to achieve faster processing speed.

Appendix A

Before starting the derivation, it is reasonable to make two assumptions:

(1) For MPSK and M-ary QAM, each of the M possible symbol states occurs with the same probability 1/M;

(2) Within a symbol stream, the occurrence of each symbol is independent of other symbols.

Define a sequence of discrete random variables $N_1, N_2, \cdots, N_i, \cdots, N_L$, where X_i ($i \in [1, L]$) means the number of phase jump points that the occurrence of the ith symbol in the L-length symbol stream will contribute to N_{pjp}. Thus, the discrete random variable N_{pjp} can be represented by:

$$N_{pjp} = N_1 + N_2 + \cdots + N_i + \cdots + N_L$$

The assumptions in (1) and (2) imply that $N_1, N_2, \cdots, N_i, \cdots, N_L$ are independent, identically distributed (i.i.d.) discrete random variables with the common probability mass function as follows:

$P[N_i = 0] = 1$, when $i = 1$;

$$\begin{cases} P[N_i = 0] = 1/M \\ P[N_i = 1] = 1 - 1/M \end{cases}, \text{ when } i \in [2, L].$$

Therefore, the mean value of N_{pjp} will be:

$$E[N_{pjp}] = E\left[\sum_{i=1}^{L} N_i\right] = \sum_{i=1}^{L} E[N_i] = (L-1) \cdot \frac{M-1}{M}$$

Since $L = \left\lceil \frac{L_s T_s}{T} \right\rceil$, the ratio of $E[N_{pjp}]$ to L_s will be:

$$\frac{E[N_{pjp}]}{L_s} = \frac{\left\lceil \frac{L_s T_s}{T} \right\rceil - 1}{L_s} \cdot \frac{M-1}{M}$$

In the real system, we usually choose an oversampling rate in the range of [3, 8], and capture several hundred or thousand of samples. Thus, it is reasonable to conclude $\lceil T_s/T \rceil \gg 1/L_s$, and the above equation can be approximately simplified as:

$$\frac{E[N_{pjp}]}{L_s} \approx \left\lceil \frac{T_s}{T} \right\rceil \cdot \frac{M-1}{M} \qquad (B-1)$$

It is worth mention that the above analysis ignores the cases where long strings of the same symbol are contained in the captured symbol stream. However, the value of $E[N_{pjp}]/L_s$, derived under the restriction that the maximum allowed length of consecutively identical symbols is 3, almost equals to the result in equation (B-1).

Chapter 6 Inter-cell Communication

Inter-cell communications among PCNs in DCCS resembles both a mobile ad hoc network (MANET) and as a wireless mesh network's backbone connection. The former stresses the self-organization features and the latter focuses on its overall connection: each node is connected to another node directly or through one or more intermediate nodes. Although each PCN performs a similar role to a base station in a standard cellular network, the connections among them are quite different. In cellular networks, the connection between base stations is wired, and interference or capacity are not major issues. However, in DCCS, connections between PCNs are wireless, and neither access to the spectrum nor the required capacity are guaranteed. This means that in DCCS, we have more to consider than the traditional backbone connection in standard cellular telephone networks.

In contrast to intra-cell communication, inter-cell communications is more of a network layer problem, instead of a lower layer problem. There are many advanced MANET protocols and systems that we may consider. In [66], a well developed system "Dynamic MANET On-demand Routing Protocol (DYMO)" is proposed. They mentioned three principles of design guidance: extendible via new element types, known behavior for unsupported elements, and simple implementation[66].

The structure of this chapter is shown in Figure 44. It includes mainly three problems. The first is spectrum utilization. Two solutions are discussed: temporary leasing and OFDMA. In the temporary leasing, we also discuss a special case: 700MHz public safety application. The second problem is about the control message communication, and it includes neighborhood discovery protocols and the format of control messages. At last, a summary of inter-cell communication is given.

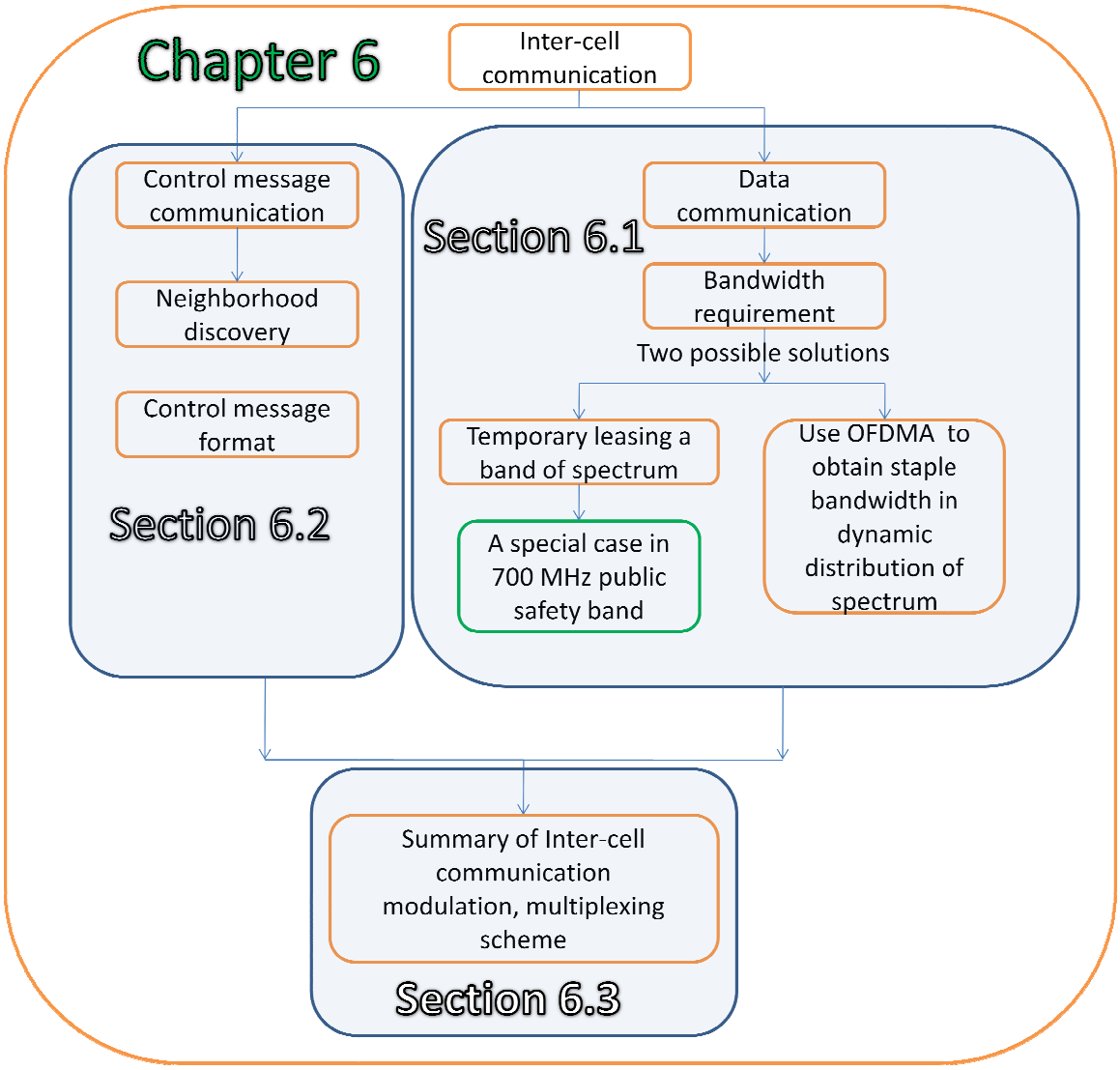

Figure 44: Structure of chapter 6

6.1 Spectrum Utilization

The spectra for inter-cell communication and intra-cell communication in DCCS are non-overlapping. Inter-cell communication has a difficult requirement for the spectrum than intra-cell communication. It usually requires more bandwidth and more stable access to the spectrum. If PCNs can access certain spectrum as primary users when performing inter-cell communication, the problem is solved. So, the first solution is to find possible spectrum that could be temporarily allocated or licensed to DCCS for inter-cell communication. However, in most of the cases, DCCS cannot access the spectrum as a primary user. This might be possible in public safety

applications but not in TV white space. In order to maintain the quality of communications, a robust and flexible channel utilization scheme should be use to obtain a stable bandwidth in a dynamic distribution of available spectrum. This robust and flexible channel utilization scheme is OFDMA. We will explain why and how to use OFDMA for inter-cell communication later in this section.

As we discussed in Section 1.5, to accurately detect the availability of white space, GPS technology is usually combined with spectrum sensing to make the correct decision. The IEEE, together with the FCC, is pursuing a centralized approach to available spectrum discovery [67]. Each base station would be armed with a GPS receiver which would allow its position to be reported. This information would be sent back to centralized servers, which would respond with the information about available free TV channels and guard bands in the area[68]. IEEE 802.22 is a standard for Wireless Regional Area Network (WRAN) using white spaces in the TV frequency spectrum. The development of the IEEE 802.22 WRAN standard is aimed at using cognitive radio techniques to allow sharing of geographically unused spectrum allocated to the Television Broadcast Service, on a non-interfering basis, to bring broadband access to hard-to-reach, low population density areas, typical of rural environments"[68].

The combination of database and spectrum sensing policy methods can open the possibility of leasing the spectrum temporarily according to some policy agreement. Note that this is leasing exclusive secondary user rights. Assuming that the policy allows, channels that are locally available can be temporarily leased to a certain group of users, and a public database for non-leasing parties to download will show who is allowed and who is forbidden to transmit in the leased band. See Figure 45. The leasing party still needs to sense the spectrum in the leasing band before transmitting in case the primary users appear in the spectrum, but they are the only secondary users allowed to use the spectrum. The probability of primary users showing up is low. In this way, the inter-cell communication will have a relatively stable spectrum to access because for both TV white space and public safety, the frequency range is from 54-862MHz and the coverage area is relatively large.

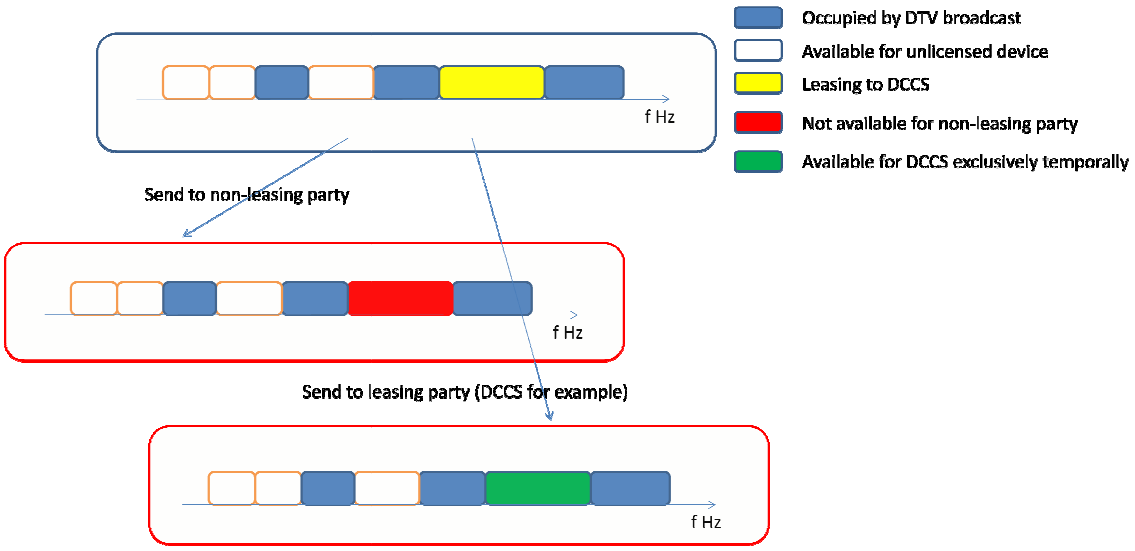

Figure 45: Temporary leasing proposal

A special case for the temporary leasing method is 700MHz spectrum for public safety broadband. In Figure 46, the public safety band in the now reallocated UHF TV band is shown.

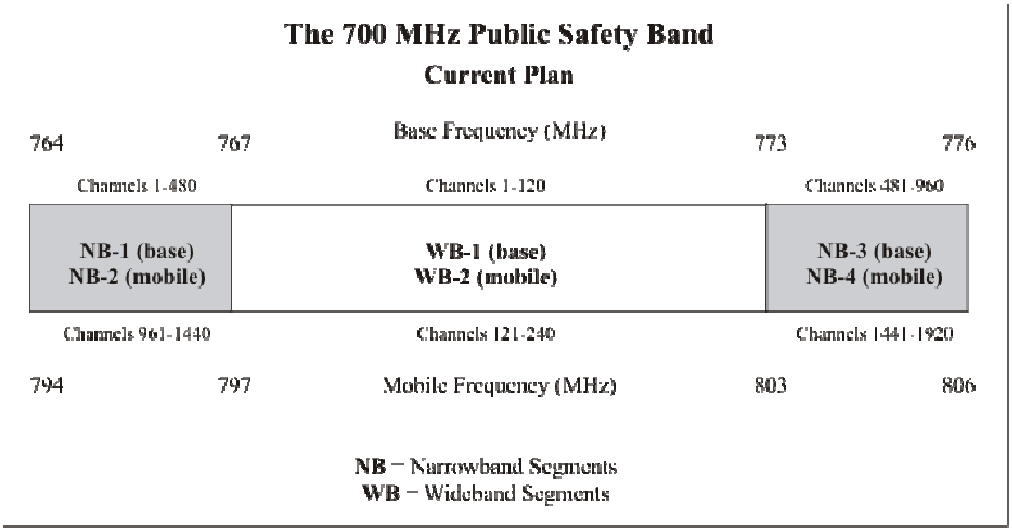

Figure 46: 700MHz spectrum for public safety broadband [69]

FCC has decided to modify the spectrum allocation to accommodate broadband communications by changing the legacy WB to accommodate broadband communication for the purpose of transmitting video, high resolution pictures, etc. The national public safety broadband licensee

will operate a broadband network, which has full access to 767MHz-773MHz and 797MHz-803MHz and has secondary access to bands 764MHz-767MHz, 773MHz-776MHz, 794MHz-797MHz, and 803MHz-806MHz. The licensee has the authority to provide capacity to commercial service providers on a secondary user basis[69]. Full access to the wideband segments can be considered as a special type of leasing the spectrum to certain parties.

When applying DCCS in the 700MHz public safety for the D-block auction, the full access bands are used by inter-cell communication and the secondary access band is for intra-cell communications.

However, the spectrum leasing method requires an FCC policy update or modification, which may take a long time. An alternative method is to use OFDMA to obtain stable bandwidth of spectrum in a dynamic distribution of available channels. OFDMA is a multi-user version of the OFDM digital modulation scheme. Multiple access is achieved in OFDMA by assigning subsets of subcarriers to individual users. OFDMA can also be understood as a spectrum combination method for secondary users. Secondary users collect information about the available channel distribution and combine the subcarriers together to transmit. This is shown in Figure 47.

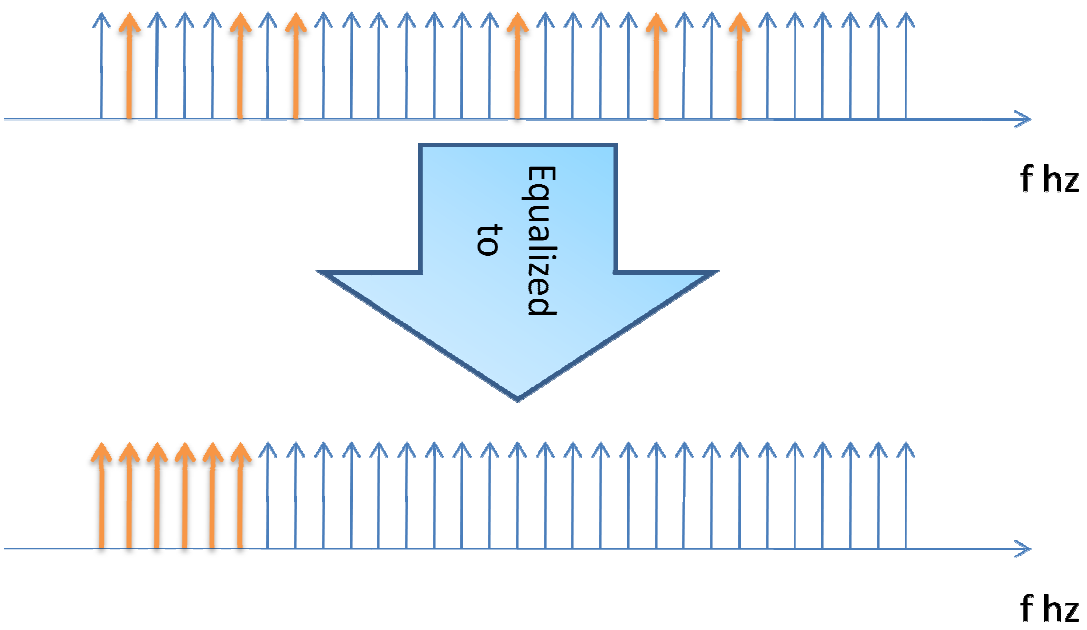

Figure 47: OFDMA scheme of collecting randomly distributed channels to be used by one user

OFDMA is used in WiMAX for multi-user multiplexing. The pattern of the subcarriers distribution for each user is static in WiMAX. In DCCS, the distribution of the available channel varies by both time and space. Coordination among the nodes is thus more complicated than it is in WiMAX. A simple pre-allocation scheme is designed for DCCS. The rest of this section is the detailed OFDMA scheme applied for DCCS.

Pre-allocation is for multiplexing the PCNs inter-cell communications. All the channels in the band are divided into $N+1$ groups. The ith channel belongs to the jth group, where $j = mod(i, N+1)$, then $0 \leq j \leq N$. For example, PCN A will choose one group of channels from group 1 to group N exclusively. Group 0 is used for message exchange among PCNs for inter-cell communications. If another PCN B is in the propagation range of PCN A, it will choose another group of channels. If PCN C is out of the propagation range of PCN A, it can choose the same group of channels of PCN A because they will not interfere with each other. This is shown in Figure 48. The arrow represents the channel. The channel in the same group is shown in the same color. Group 1 to group N are for data communication, and group 0 is for message transmission. Each time a PCN chooses a group of channels, it will inform other PCNs through channels in group 0. The detailed physical layer settings for message transmission are described in Section 6.2.

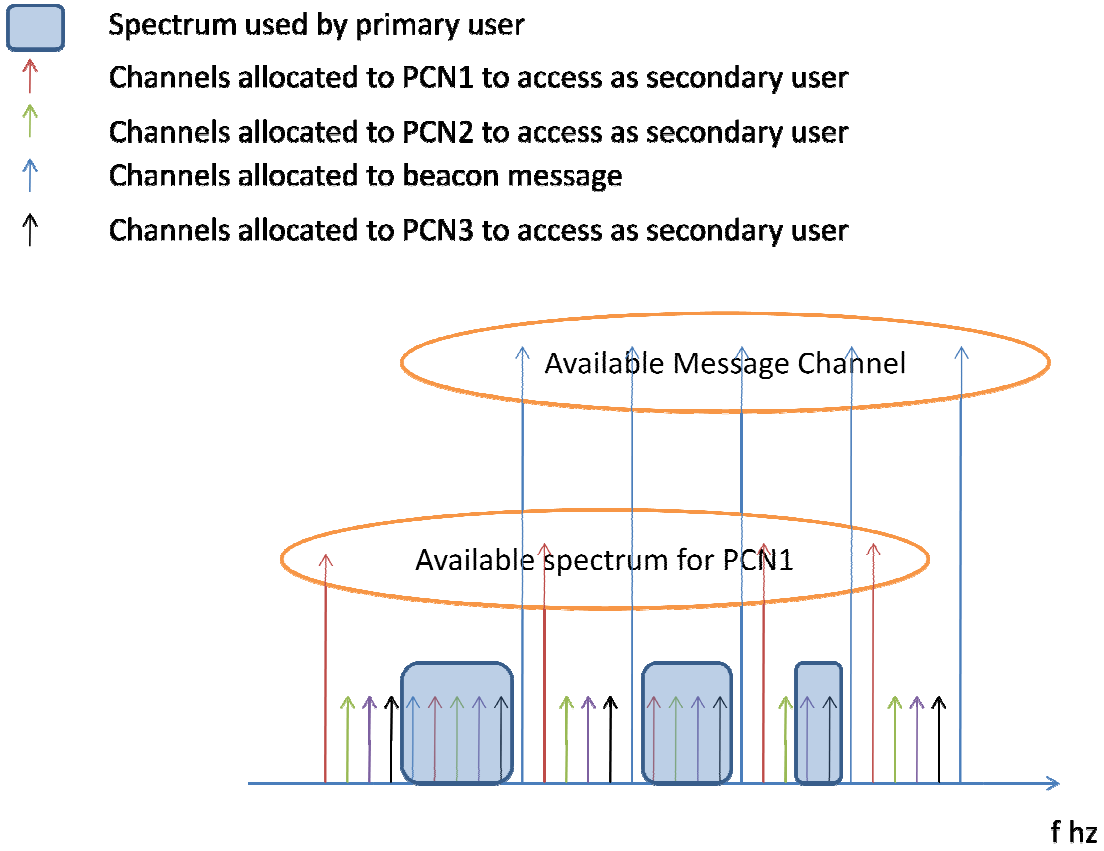

Figure 48: Pre-allocation of OFDMA scheme for inter-cell communication

For example, PCN 1 choose the channels in the red group. Suppose it is called group k. Then, the vacant channels in group k are the channels for OFDMA based data transmission of PCN1. Vacant channels in group 0 are the channels for message transmission. To correctly receive OFDMA signals, it is necessary to know which channels are used. This information is broadcast by PCN 1.

6.2 Control Message and Neighbor Discovery Protocol

The message exchange scheme plays an important role in inter-cell communication. The agreements and negotiations among PCNs are achieved by message exchange. In this section, three aspects of message exchange are discussed. The first aspect is the physical layer settings of the message transmission and receiving scheme. This aspect includes the channels to use, modulation type, symbol rate etc. for transmitting messages. The second aspect is about different types of message, including the functions of the messages, contents of the message, and format

of the messages. One type of the messages is for neighbor discovery protocol. The third aspect is the detailed description of neighbor discovery protocol.

As mentioned in Section 6.1, messages for inter-cell communication are transmitted through the message channel, which is called as group 0. The same message is transmitted through every vacant channel in group 0 instead of just sending it in one of the vacant channels. Because the message is short, and in a dynamic spectrum environment, it is easy to miss some of the messages. And the message is important for correctly receiving the data packages in the data communication channels. For intra-cell communication, a hand shaking scheme is used to guarantee that a message is successfully delivered. In inter-cell communication, messages are broadcast to multiple PCNs and handshaking is complicated to implement in this situation. Transmitting the same message through different channels brings frequency diversity, and thus enhances the successful rate for message exchange. The intended receivers will observe all the sub-channels in group 0. Error detection code in the data package is available in GNU radio. The corrupted message will be discarded. Among the successfully received messages, the same messages will be only processed once.

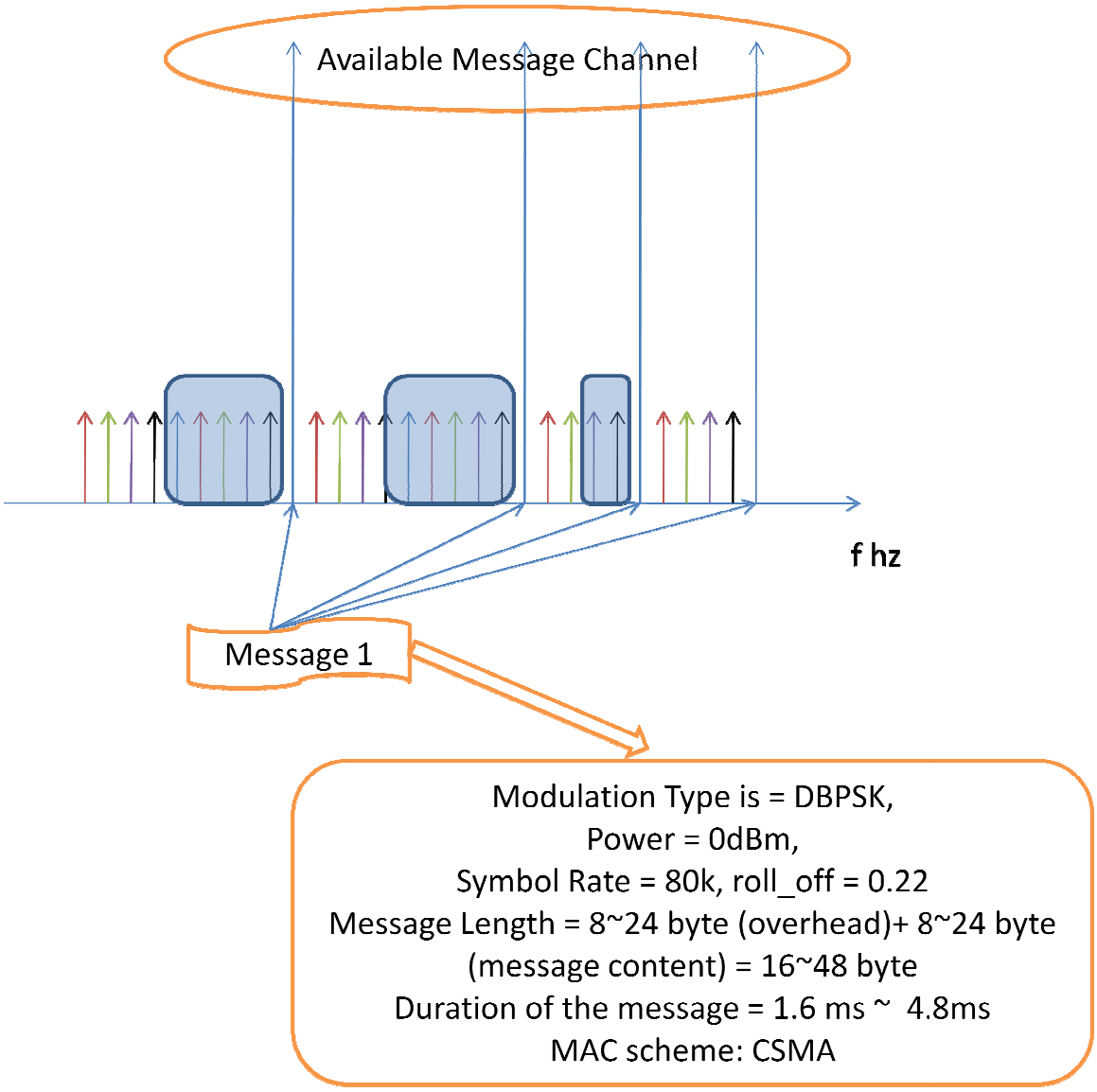

Figure 49: Message exchange for intra-cell communications showing parameter values used in my demonstration.

In Figure 49, the physical layer setting of the message is shown. DBPSK is chosen as the modulation type because it is the simplest digital modulation and has a low BER for a given SNR scenario. The symbol rate is 80k; because the channel bandwidth is 100 kHz and the raised cosine pulse shaping filters roll_off factor is 0.22. The overhead of a package in GNU radio is about 24bytes and it is adjustable. The content of each message is less than 8 bytes. CSMA is used as the multiple users sharing scheme. When transmitting one message through multiple channels, multiple randomly back off times are chosen for different channels before transmitting.

Based on the functions, there are three types of messages. Message type I is for routing table updates. Message type II is for OFDMA active subcarrier information, which means it lists the channels used for the OFDMA packages. Message type III is for neighbor discovery, which will be described later in this section.

Message type I contains information about currently registered CMT users. A PCN sends out message type I on a regular time period basis and also when a new CMT has registered or a CMT has left the cell. There are two conditions by which a PCN knows that a CMT has left the cell. The first condition is that a CMT has been added to the other PCNs registration list. The second condition is when a CMT is requested, and the PCN cannot deliver a message to the CMT. All the other PCNs which received this message will update their databases accordingly. The format of this message type I is shown in Table 9.

Table 9: Message type I

Message type (1 bit)	Message Generated Time(1 byte)	Source PCN ID or IP (15 bit)	Current registered CMT ID(1 byte)	Current registered CMT ID(1 byte)	Current registered CMT ID(1 byte)	Current registered CMT ID(1 byte)	Current registered CMT ID(1 byte)
I	Time	ID or IP	CMT#	CMT#	CMT#	CMT#	CMT#

When another PCN receives the message, the PCN uses it to update its data base, it and uses a routing algorithm to update its routing table. The routing algorithm is not included in this dissertation. Further research can be done on designing specific routing algorithm for DCCS.

In DCCS, it is specified that each PCN needs to have an IP address, but a MT does not necessarily have an IP address. If a MT doesn't have an IP address, its associated PCN is able to allocate a temporary ID to it in DCCS. On one hand, each PCN's having an IP address enables the implementation of some TCP/IP protocols in DCCS. It also provides DCCS with convenient interfaces to some IP architecture networks, for example, Wi-Fi or WiMAX. On the other hand, the allowance of MTs with/without IP address can accommodate more devices and conditions. Some devices are not designed for IP architecture; the requirement that each node have an IP

address would restrict the application for these devices in DCCS. Because of the centralized structure within a cell, being without an IP address doesn't prevent effective communications.

Message type II contains the OFDMA active subcarrier information. The information includes the group number and the current OFDMA active channels. PCNs access the channel as secondary users, and the channel conditions vary with time. Having such type of messages to inform other PCNs and keep receivers synchronized is important. The format of message type II is shown in Table 10

Table 10: Message type II

Message type(1 bit)	Message Generated Time (1 byte)	Source PCN ID Or IP (15 bit)	Group Number (1 byte)	Active channel number (4 byte)
II	Time	ID or IP	Group #	Active channels #

Message type III is for neighbor discovery. Before defining the format this type of message, we will describe the neighbor discovery protocol.

The architecture of inter-cell communication among PCNs is distributed. To avoid being limited by the number of nodes, the topology needs to be as flat as possible[70]. What we are looking for is a protocol that determines, in a distributed, mobile, and scalable manner, what nodes are neighbors in a wireless network. Because PCNs in DCCS are portable devices, there are strict power restrictions, especially for accessing spectrum in TV white space[70, 71] and battery life is a consideration. Thus, the energy cost of network-wide synchronized protocols is not acceptable for DCCS, and the designed neighborhood discovery protocol must be asynchronous. The problem of neighbor discovery can be stated as follows:

General assumptions:

1. Before becoming neighbors, nodes are asynchronous, and they can be synchronized after becoming neighbors.

2. Neighborhood relationship is bi-directional. If node A is node B's neighbor, messages transmitted by A can be properly received by B, and messages transmitted by B can be properly received by A.
3. All PCN in DCCS are cooperating, and every PCN can be trusted.
4. The neighborhood list for a PCN can be incomplete.
5. A node that is a neighbor to any node in a group has a neighbor relationship with the group.
6. Potential neighbors of a node A are the nodes that are not yet claimed as the neighbors of A but can receive A's messages.

A simple protocol can be divided into three steps:

1. Every node broadcasts a request containing neighbor discovery information
2. Every node that receives the request responds with a neighbor discovery reply
3. By the identification information in the message, each node finds its neighborhood

However, a well performing neighborhood discovery protocol needs a lot more consideration. A great amount of research has been done in this area, for example, IPV6 Neighborhood Discovery Protocol (NDP), distributed neighborhood discovery using directional antennas[72], asynchronous neighborhood discovery algorithm for wireless sensor networks [73], or a self-organization algorithm using super frame and boot up time [70] and some other methods and protocols. For most of the protocols, "if the network topology is expected to change during network operation, the neighbor discovery algorithm could be re-run"[73].

For the purpose of fitting into DCCS inter-cell communication, we designed a relatively simple model for neighbor discovery. Based on general assumption 1, nodes which are neighbors can perform in a synchronous manner. Three neighborhood discovery modes are defined as described in Figure 50. They are:

A. One node requesting to join a network
B. Two networks requesting to join together
C. One node requesting to be a neighbor of another node

In Figure 50, the same color nodes in each block are able to perform in a synchronous manner. In case A, we call the blue dots group ζ, and the single yellow dot α. α first listens to the message channels. Based on general assumption 3, the neighborhood relationship is bi-directional. Thus, α is able to receive messages transmitted by potential neighbors in ζ. α then transmits a joining request. The PCNs who received this message will send A the most updated routing table and add A as a neighbor. A will confirm the neighbor relationship and officially join the network.

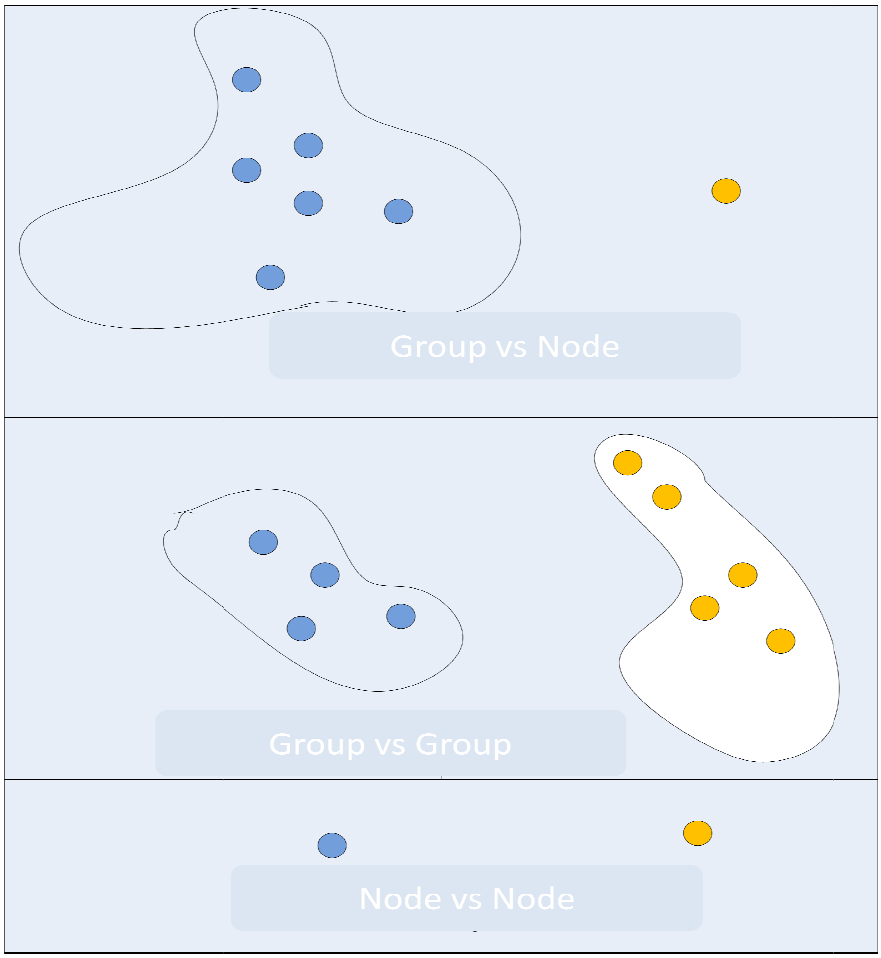

Figure 50: Neighborhood modes explanation

For case B, when two groups meet, a similar process will be performed with some modification. Should any node in group ζ receive a message other than from its neighbor, it will exchange the routing table with the transmitter. Routing tables from the two groups will be merged, and a larger network is formed.

For case C, when two nodes become the only neighbor of each other, three-way handshaking is performed

Different from message type I and message type II, the neighbor discovery message, message type III, is three-way handshaking. Routing table information is exchanged if applicable. The format of message type III is shown in Table 11.

Table 11: Message type III

Message type(1 byte)	Message Generated Time (1 byte)	Source PCN ID Or IP (15 bit)	Group Number (1 byte)	Routing Table
III RQST or III RESD or III ACKW	time	ID or IP	Group #	

6.3 Summary of Physical Layer Modulations and MAC Layer Protocols for inter-cell Communication

In this section, we will give a summary of the physical layer modulations and MAC layer protocols for inter-cell communication.

In the physical layer, two types of communications are required for inter-cell communication: data communication and message communication. Channels are divided into groups, and each PCN uses one group to transmit. OFDMA is used for data communication. An available channel in the group is a subcarrier of the OFDMA signal. Because of the varying distribution of available channels, each PCN inform its neighbors about the active OFDMA subcarriers.

Messages are transmitted through channels in group 0. Their modulation setting is shown in Figure 49. The same message is transmitted through every available channel in group 0, gaining frequency diversity in the price an increasing overhead.

Message types are defined based on their functions. Type I messages transmit CMT members' information, so that the receivers can update their databases and the routing tables accordingly. Type II messages transmit OFDMA active channel information to assist OFDMA demodulation. Type III messages transmit routing table and neighbor information so that new PCN is able to

join the network and share the routing information with others, and thus small DCCS networks can join together.

Chapter 7 DCCS Prototype and Demonstration

With all of the functional system designed, in this chapter, we will present the prototype of a DCCS system. In this chapter, we will first give an overview of the current development progress of DCCS. DCCS is a multi-function system, and there is still a great amount of development and implementation work to do in the future. Then, we will introduce the prototype for a PCN, and then the prototype for a CMT. After that, we will describe the demonstration settings for DCCS system prototype. At last, the test case for each function integrated in the prototype of DCCS is given as well as the testing results.

7.1 Current Development Progress of DCCS

The current development progress of DCCS is shown in Figure 52. DCCS was first proposed in November 2007 for public safety purposes. The assumed scenario is in natural disaster area where infrastructure is damaged, for example, the earthquake in China and the area affected by Katrina. First responders from different departments come to this area. Different types of communication are needed and used. Spectrum is accessible partly on a secondary user basis and partly on a primary user basis. This scenario fits into the U.S. 700MHz D Block auction for a public safety nationwide broadband network. The early development of DCCS followed this path. PCN/CMT switching, interoperability, dynamic spectrum access, signal classification and synchronizations are the main focus of DCCS. The basic DCCS functions were developed during this period. These functions include cell formation, PCN/CMT switching, UCS, DSA, message exchange for intra-cell communications.

UCS was developed as an independent system beginning in December 2007, by Qinqin Chen and I. It can be fully integrated on PCN, and it is an important piece of a PCN.

As I became more deeply involved in DCCS research, I found out that a more widely usable application of UCS would be TV white space. DCCS can serve as a white space device network.

Such a network is reliable in the avoidance of interference to primary users – TV receivers. Interference to primary users is so far the largest obstacle for the huge market of white space device application. This network can also provide a higher throughput for the limited spectrum resources by channel allocation. Having high spectral efficiency is important for white space networks. In this stage, channel allocation, a more flexible and reliable DSA scheme, and more rational network architecture are my research focus. Channel allocation, cell splitting and merging, inter-cell data and message communication were researched during this period.

DySPAN 2008 provided an excellent opportunity to demonstrate, and more importantly, to test the DCCS system in a harsh RF environment. More than ten teams participated in the demonstration session, and all the teams shared the same spectrum in a small room. To obtain the quality communication in this spectrum environment requires a robust and flexible network. Many graduate students in our group contributed in preparing the DCCS demonstration. My programming work mainly focused on the PCN node. DCCS was developed more robust and more functions are added during this period. GUIs for PCN and CMT were built. The handshaking schemes were fully integrated and tested. Using Python, the channel allocation scheme was programmed, integrated into the PCN code and tested. Hand-off processing between CMT and PCN nodes was implemented. Interfaces to Wi-Fi network and gateway to Walkie-Talkie were integrated to DCCS. The application layer was integrated with the system, enabling the system to interface to different user defined applications, for example, video transmission. The one-hop forwarding function was developed and integrated. More than ten teams accessing the spectrum simultaneously provided a unique testing environment. Channel changing occured frequently and the noise floor was also higher than usual because of the power leaking from different devices. DCCS was successfully demonstrated in DySPAN 2008. In Figure 51, some pictures from that meeting are shown.

Figure 51: Pictures from DySPAN 2008

After DySPAN 2008, more functions were developed in DCCS, including inter-cell OFDMA communications, inter-cell message exchanging scheme, and power control scheme of DCCS. Evaluation of the functions and algorithms in DCCS is also researched. In Figure 52, different block colors show the current development stage of DCCS. The green blocks are the functions that have been fully implemented and demonstrated. Given better hardware and more time, functions shown in other colors can be implemented in the future.

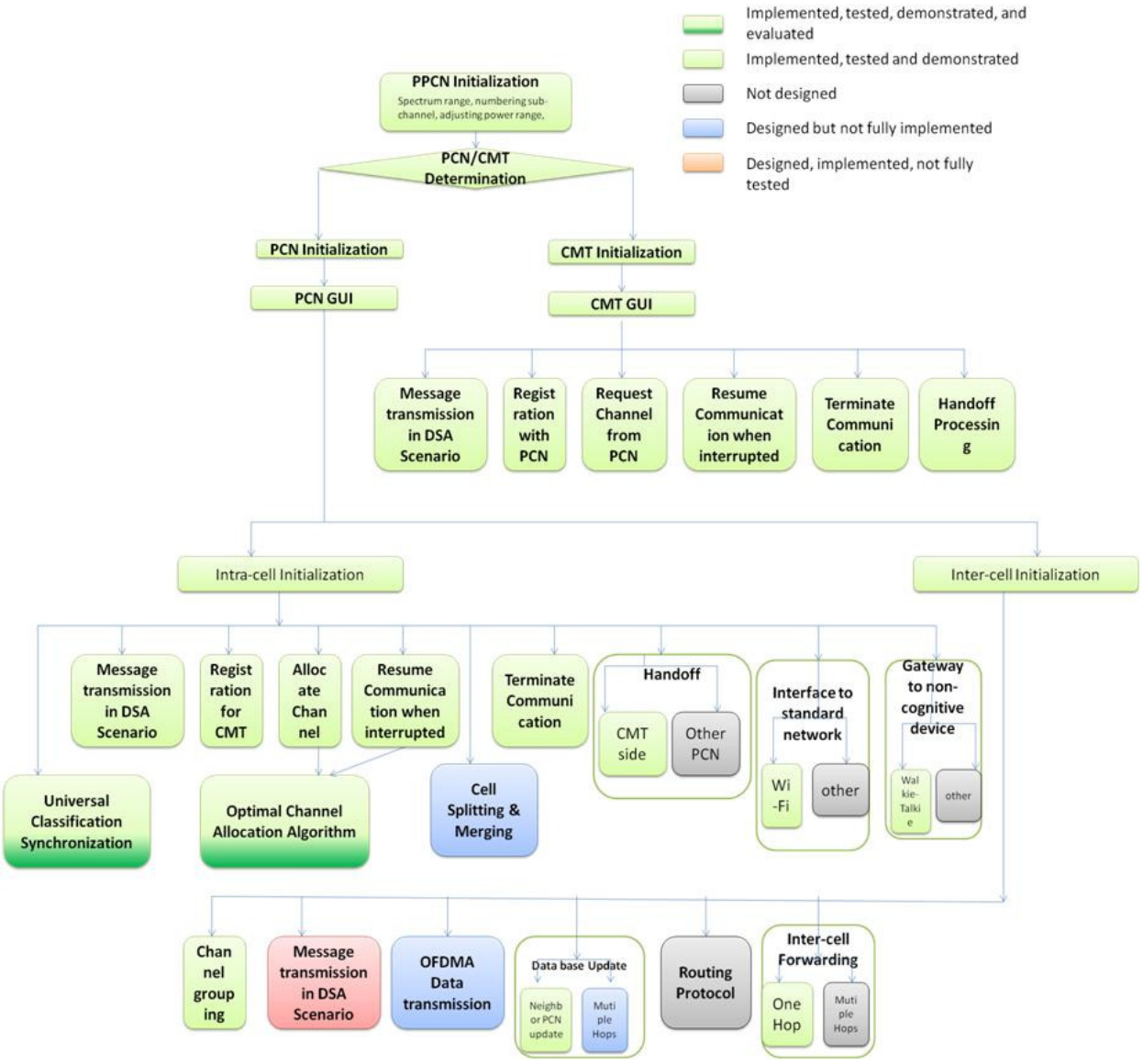

Figure 52: Current development progress of DCCS

7.2 Prototype for PCN

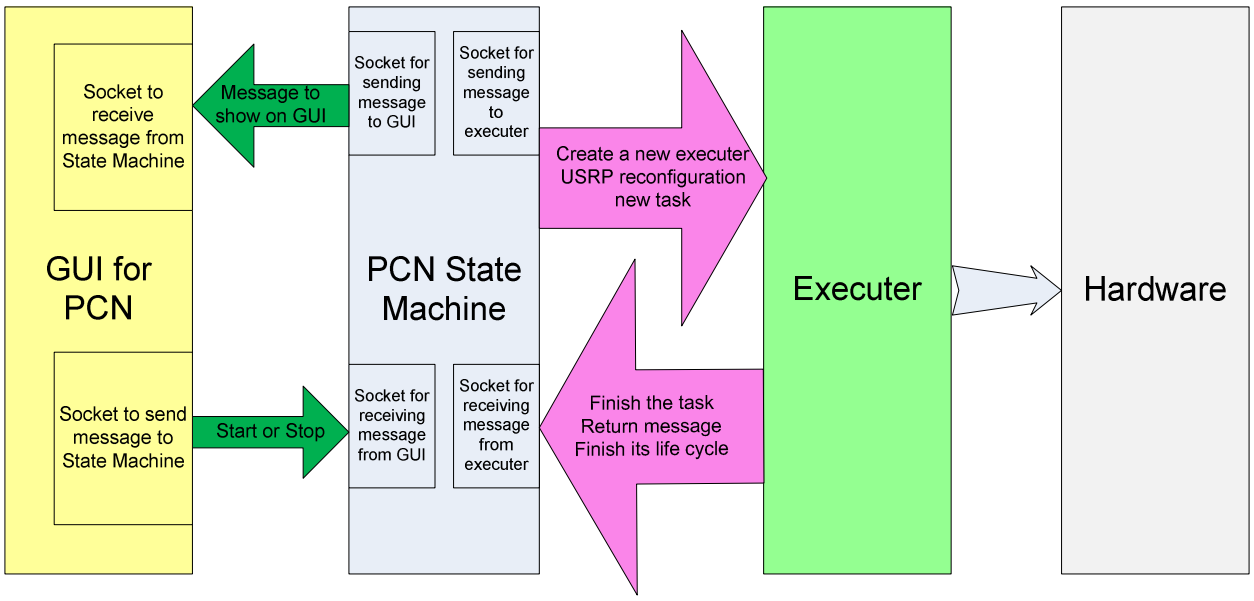

Figure 53: Software prototype of PCN

Figure 53 shows the structure of a PCN from a software point of view. The user's commands are passed to the PCN state machine by the GUI through a socket. The user's commands include start and stop. Other processing by the PCN doesn't require an operator. The main purpose of GUI is to display information to observers. Once the PCN receives the command "start" from GUI, the PCN state machine gets into the initialized status and will create an executer process. The executer process is the part that interfaces with the hardware, performs spectrum sensing, and communicates with CMTs or other PCNs. The executer will generate a transmit flow graph and a receive flow graph. In the GNU radio and USRP version we use to implement DCCS, it is not allowed to fully configure the hardware while the flow graph is still running. Thus, an executer will return a message to the state machine once it finishes everything that needs to be done in one channel and ends its life cycle. The message will include the information the state machine needs to update GUI, database and make decisions re which channel to move to. After it delivers the message to the state machine, its cycle is finished. The state machine will analyze the received message and makes the next step decisions, including sending messages to the GUI to show them to the users, determining with which channel is the next channel to access and what task to perform, and then generating a new executer to configure the hardware and perform the tasks. The information displayed on GUI includes the current channel status, associated CMTs' status, neighbor PCNs' status, and current on-going intra-cell communications.

The life cycle of the executer is 0.05s. This value is a test result based on the shortest time to reconfigure a USRP. If the life cycle is set much shorter, then, because of the latency of the USRP and the General Purpose Processor (GPP) processing speed, it will often appear that the old life cycle is not finished using USRP before the executer has already begin to use USRP for the next cycle. Then, hardware's being busy will cause an error in this system. When the life cycle increases, the sweeping time for a PCN through the entire spectrum will increase which means a long waiting time for a CMT to send a request.

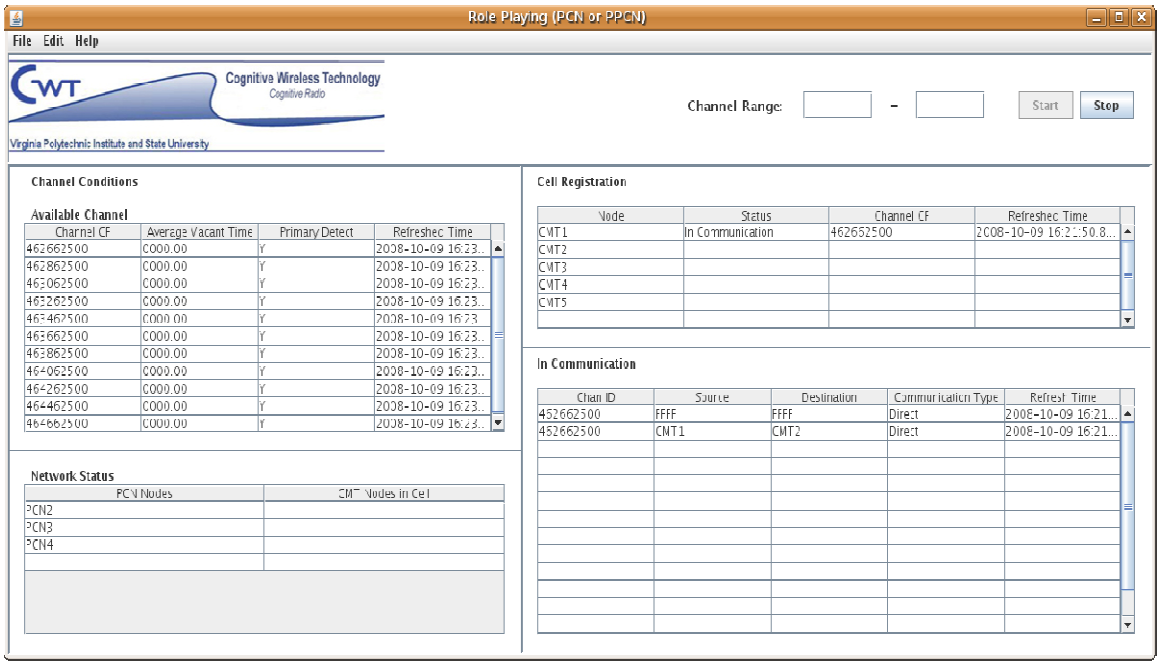

Figure 54: GUI for PCN

Figure 54 shows the GUI for PCN. The column in the upper left corner is the listing of sub channels. When PCN goes through each channel, besides processing the request from CMT, it is also executing a process of detecting the channel utilization information. In each channel, it will record the average time that a channel is occupied by the primary user, whether the channel is available at this moment, and the last time a primary user is detected. It doesn't require the continuous coverage of a channel by a receiver. The trust factor we used in Section 4.3 has taken the period of coverage of a channel into consideration. This information is used to determine the channel as shown in Section 4.3. (Note that these features were implemented in the demonstration.) The network status item shows the direct connection with other PCNs and their

associated CMTs. On the right side of the GUI, the cell registration shows this PCN's associated CMTs' information. The information includes whether a CMT is in communication or just registered but not in communication, which channel this CMT is currently in, and the last update time. This form tracks the CMT records. Under this item, there is another item called "in communication". This item shows the information about ongoing communications in this channel. It contains the identities of transmitter and receiver, whether the communication is a direct connection or the PCN needs to forward the message to other cells, and the last update time. All of this information keeps the user updated on the channel status and cell conditions.

The flow graph of the intra cell communication PCN state machine is shown in Figure 55. From the flow graph, we can see that this block of the PCN is not directly working with hardware or in communicating with other nodes. It assigns tasks to the executer and maintains the data base and GUI update. A task generally includes all the work needs to be done in one channel. When executer has finished the task, it will report the information to PCN and finish its life cycle. When the state machine calls the executer, two threads are executing at the same time. However, during this time, the state machine doesn't do anything except wait for the report from the executer. Thus, the PCN is a single thread in the demonstration. It will be more efficient if the PCN is multitasking. Because of the complexity in hardware and software, this is left for future improvements. When the state machine receives "time out" from the executer, it means that an executer has been listening in a vacant channel for 0.05s, and nothing happened. The state machine will update the channel information in GUI and move on to the next channel. When the state machine receives "primary user", it means an executer has detected the primary user that exists in this channel. The state machine will updated the channel information in GUI and move on to the next channel. The other four cases are signs for finishing the hand shaking with CMTs. It is defined that in each channel, PCN only processes one request at any one time. This definition will guarantee that the time period for a PCN to revisit a channel is less than $0.05N$s, where N is the number of channels. In the demonstration, the value of N is 10. If the finished task is user registration, or termination of using the channel, then the state machine will generate a command to ask the executer to move to the next channel. If the finished task is a request for a channel, PCN will need to look up the channel location where the intended receiver is located, and generate the command to the executer to move to that channel. If the intended receiver is

outside of the cell, the inter-cell function will be called, executer will end its life cycle, the state machine will call executer, and hardware will be configured to move the next channel. If the finished task is to finish informing the intended receiver, then the PCN will move to the next channel.

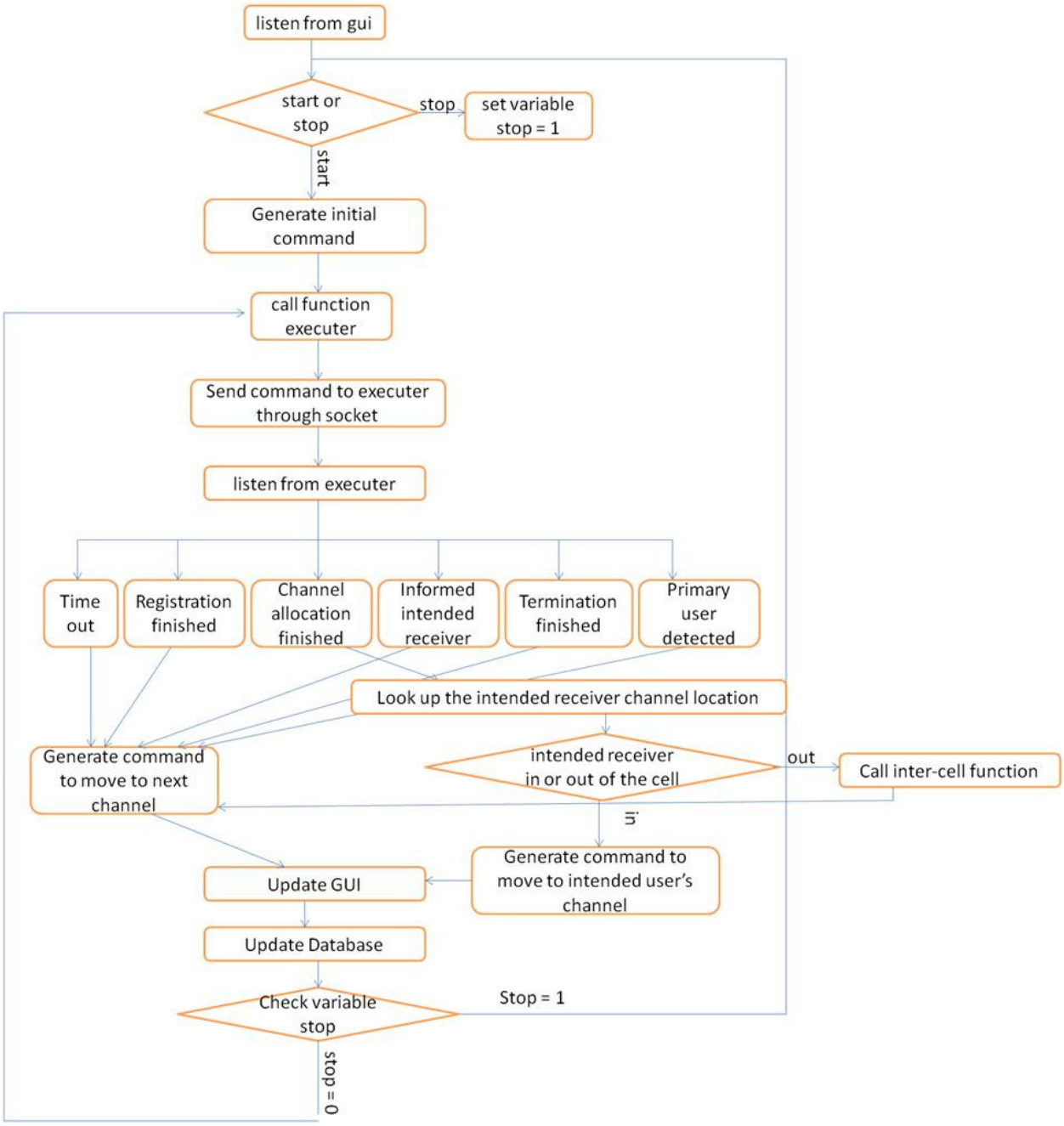

Figure 55: Flow graph for PCN state machine

In Figure 56, the flow graph of the executer is shown. Each path in this flow graph ends with reporting to the state machine. From the beginning to the end is a life cycle. If we consider a PCN as a structure, the GUI will be the eyes and ears; they hear and see the requirements, and transmit the information to the brain, which is the PCN state machine. It will make the decision and assign the tasks to the hands, which are the executer.

One special task the executer does is to allocate a channel. If the request is to ask for a channel, the executer will look up the available channel, choose a channel with a large throughput using the method explained in Section 4.3. The reason to ask the executer to call this function is to include the channel information in the respond message, and thus, overhead is reduced and the user waiting time is also greatly reduced.

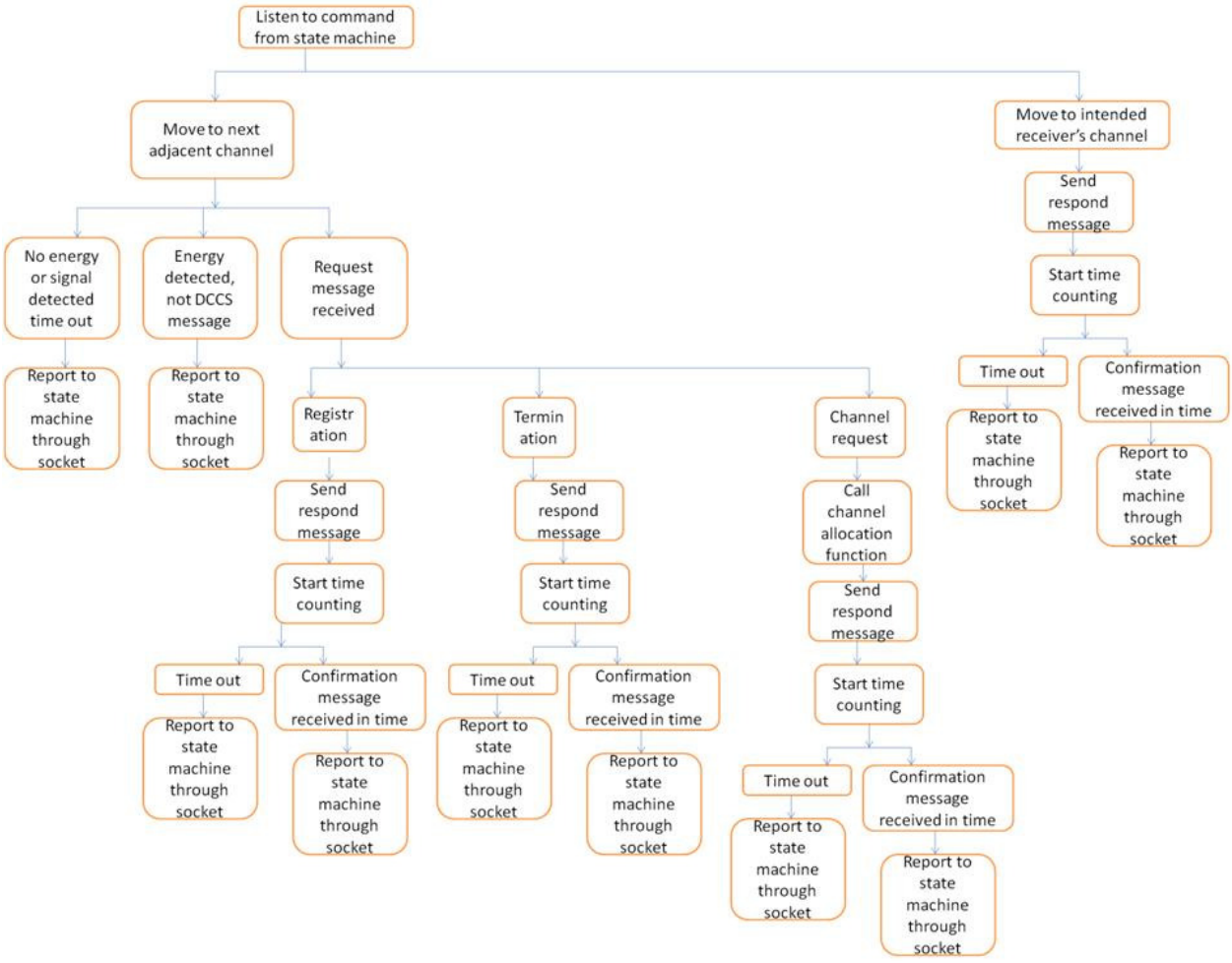

Figure 56: Executer of PCN

In Figure 55, a function called inter-cell is called when the intended receiver is outside of the cell. Because of the limitation of time and available equipment, I only implemented one-hop inter-cell communication. In demonstration, the hardware to perform inter-cell communications is another USRP board connected to the same GPP. In this way, the intra-cell management function is still functioning, and the data forward function for a CMT to communicate with a node outside of the cell is also performing. The flow graph of inter-cell function is shown in Figure 57.

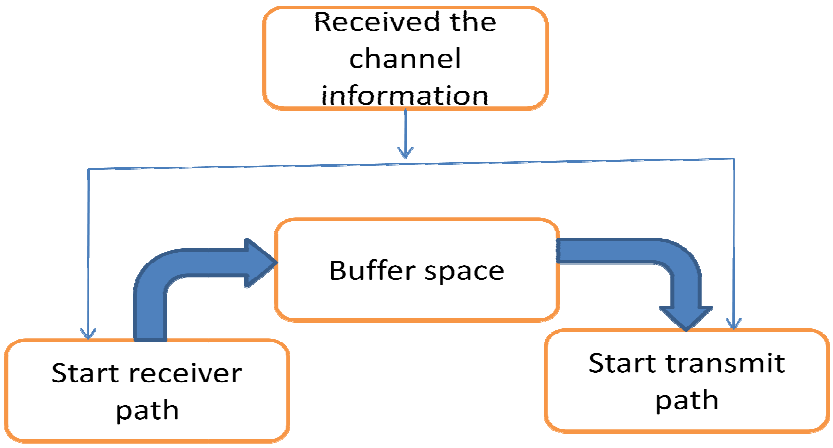

Figure 57: Forwarding in PCN

A more hardware efficient way is to have two daughter boards on one mother board performing transmitter and receiver separately. We did not implement that in the demonstration. This is left for future work.

The algorithm for channel allocation has been discussed in the previous chapter, and in this section, I will analyze the performance and compare the result with other channel allocation algorithms. In the most simple case, when there is a request, one option is to choose the channel with the largest throughput, and the other option is to randomly choose an available channel. Essentially, it is to compare the expected value of the maximum throughput and the expected value of the average throughput.

In Section 4.3, we have derived the algorithm for finding the maximum throughput.

$$Objective: \max(S_i \cdot (1 - F_i) \cdot T_i \cdot C_i)$$

$$s.t.: \qquad F_i = 1 - e^{-\lambda_{chi} t_{chi}}$$

$$T_i = -\frac{\ln(1-pa)}{\lambda_{chi}}$$

$$C_i = B_i log\left(1 + \frac{P_{maximum}}{p_{chi}}\right)$$

$$i = 1, 2, 3, \dots, N$$

Suppose $EN_{throughput}$ is the throughput of a channel (E represents r entropy.). Then, if a channel is available

$$EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi}) = e^{-\lambda_{chi} t_{chi}} \left(-\frac{\ln(1-pa)}{\lambda_{chi}}\right) B_i log\left(1 + \frac{P_{maximum}}{p_{chi}}\right)$$

One assumption is made that $\lambda_{chi}, p_{chi}, t_{chi}$ are independent uniform distributions. We consider the bandwidth of each channel to be the same.

Then the expected value of $EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi})$ is :

$$E\left(EN_{throughput}\right)$$

$$= \iiint_{\lambda_{chi}, t_{chi}, p_{chi}} cdf'_{entropy}(\lambda_{chi}, t_{chi}, p_{chi}) \left(e^{-\lambda_{chi} t_{chi}} \left(-\frac{\ln(1-pa)}{\lambda_{chi}}\right) B_i log\left(1 + \frac{P_{maximum}}{p_{chi}}\right)\right) \mathrm{d}\lambda_{chi} dt_{chi} dp_{chi}$$

For unlicensed devices in white space, the maximum transmission power allowed by FCC is 20dBm. Thus, we define $P_{maximum} = 20dBm$. p_{chi} is the noise level in a channel. The white noise level is usually low, around $-90dBm$. The main noise source in a channel is the leakage from signals in other channel. The threshold for energy detection to determine whether a channel is vacant in our demonstration is l_{th} = -20dBm. Then, p_{chi} is a uniform distribution in dB at range$(-60dBm, -20dBm)$. λ_{chi} is a uniform distribution between $(1, 2)$. In the demonstration, the period of going through all the channels is about 0.5s, the number of channels is $N = 10$. Thus, t_{chi} is a uniform distribution between$(0.05s, 0.5s)$.

The expected value of $Max(EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi}))$ is also determined by the number of samples. Since the number of the channels is N, the number of samples is also N. The cumulated

distribution function of $Max(EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi}))$ is $F_{maximum}$, the cumulated distribution function of $E(EN_{throughput})$ is $F_{entropy}$, then,

$$F_{maximum} = F_{entropy}^{N}$$

$$F_{entropy}(z) = \iiint_{EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi}) < z} f(\lambda_{chi}, t_{chi,} p_{chi})\, \mathrm{d}\lambda_{chi} dt_{chi} dp_{chi}$$

$$= \int_0^1 \int_{0.05}^{\min\left(0.5, \frac{\ln\left(\frac{\log(10^8)(-\ln(1-pa))}{z\lambda}\right)}{\lambda}\right)} \int_{-60}^{10\log\left(10^{\frac{z}{\ln(1-pa)}\lambda e^{\lambda t}} P_{maximum}\right)} dp_{chi} dt_{chi} d\lambda_{chi}$$

Where $f(\lambda_{chi}, t_{chi,} p_{chi}) = \frac{1}{0.45*1*40}$

Then, the expected value of the maximum throughput is

$$E(Max(EN_{throughput}(\lambda_{chi}, p_{chi}, t_{chi}))) = \int_z z f_{maximum}(z)\, d$$

In the simulation, we did not use the derived expected value of the maximum throughput but list the throughput with all the combinations of the parameters.

To satisfy different users' requirements is important to a wireless communication network. In this channel allocation algorithm, pa represents for the tolerable probability of interruption. This parameter is set differently for different services. For example, for voice transmission, frequent interruption is not acceptable because of the real time requirement. The requirement on SNR, on the other hand, is not strict. Thus, a channel that has higher noise floor and smaller primary user visiting rate is a better choice for voice transmission. When transmitting data packets, the preference is different. The BER performance, which is mainly determined by SNR is more important to data packets than the interruption probability.

Without giving a detailed application scenario, it is difficult to show each parameter's influence to the throughput performance. Thus, a special method is used to explain the result. There are 4 parameters considered:

$SNR = \frac{P_{maximum}}{P_i}$ $in\ range\ (10dB \sim 20dB) with\ step$ 2 dB,

$\lambda_i\ in\ range\ (0.1 \sim 0.4) with\ step$ 0.1,

$pa\ \ in\ range\ (0.1 \sim 0.2) with\ step$ 0.01,

$t\ in\ range\ (0.05 \sim 0.5) with\ step$ 0.05,

and the bandwidth is 100kHz. We listed all the possible combinations of the parameters in one dimension. SNR is in the outer loop, and t is in the inner loop, λ and pa in between. The throughput for every combination is shown in Figure 58. Each point in x direction represents one possible combination. Axes in y direction represents for the throughput.

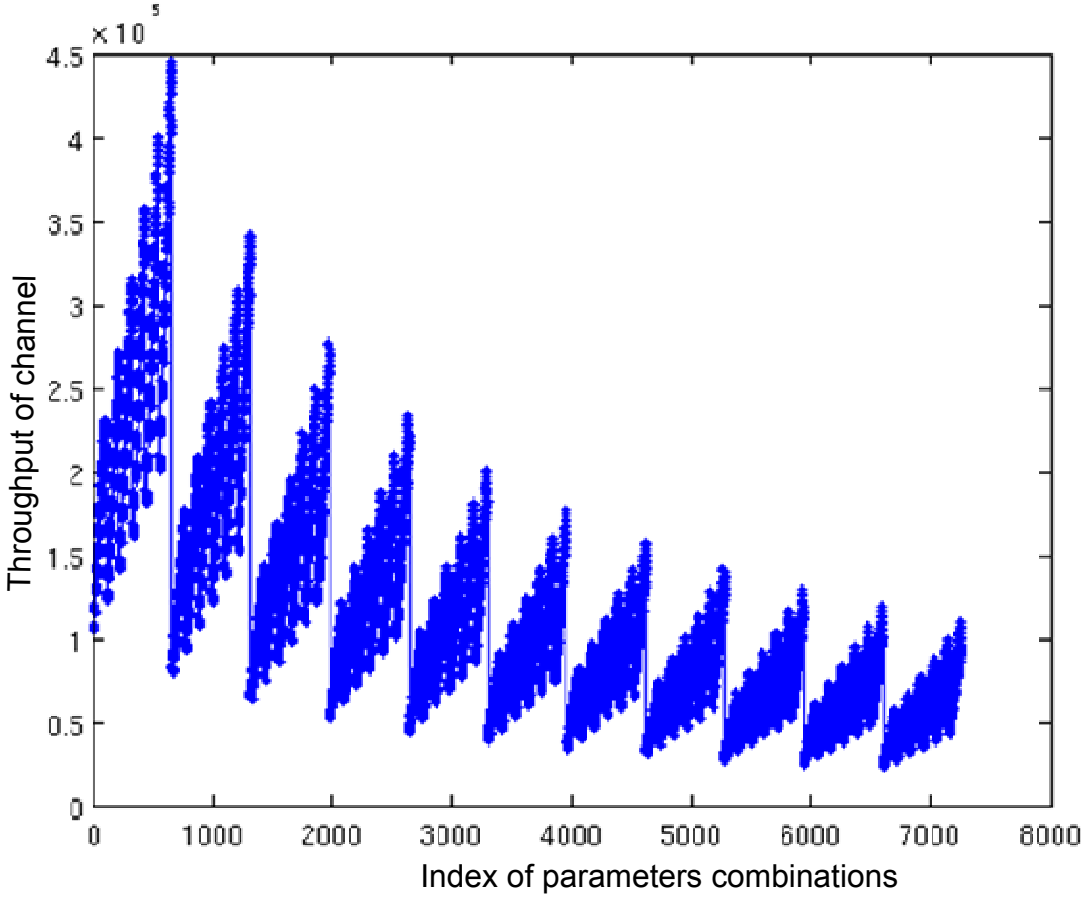

Figure 58: Throughput influenced by 4 parameters

The purpose of this figure is to show that the channel throughput changed very significantly when the parameters change. The ranges of the parameters are estimated based on the prototype. In Figure 59, the simulation result is shown. From the combinations of the

parameters, 100 combinations are randomly selected and form a group. The same process is repeated for 60 times, therefore there are 60 groups. The x direction coordinates represent for the index of the groups. Within a group, the throughput for each combination is calculated. The maximum throughput is shown in the y direction coordinate of a red dot. The throughput of the combination with the minimal primary accessing rate is shown in blue dot. The throughput of the combination with the maximal SNR is shown in green dot. The throughput of the combination with minimal time interval between the detection time and determination time is shown in black dot.

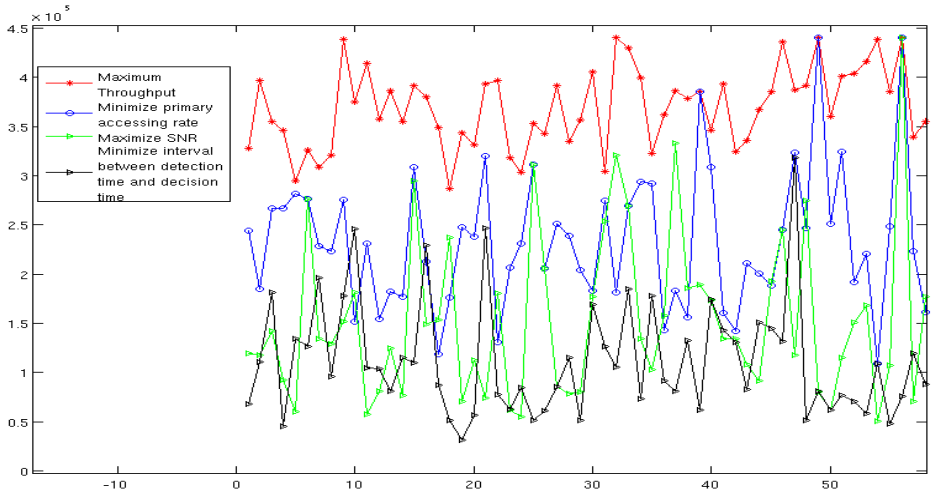

Figure 59: Simulation result of maximal throughput channel allocation

From this figure, we can get two conclusions.

1. Based on this set of parameter ranges, the most influential parameter is the primary accessing rate.
2. The maximum throughput method can increase the throughput around 1.5 times of that using the minimal primary user accessing rate.

7.3 CMT Prototype

CMT is designed in a corresponding manner to that of PCNs. The software prototype structure is similar to what it is in PCN. This is shown in Figure 60. The GUI for CMT is shown in Figure 61. Users of CMT will usually have three types of requests; one is to register with a PCN, another is to request a channel to communicate with an intended user, and the third request is to

terminate communication, which is to release the channel. In the dialog box below, the confirmation or feedback from PCN will be shown. The dialog box at the bottom shows the alert information. For example, intended user is not in the service area, communication interrupted, or no respond from the PCN is received. If a channel has been allocated for data communication, and it is interrupted by the primary user, then resuming communication happens automatically. A message showing "resuming communication" will be shown in the upper box.

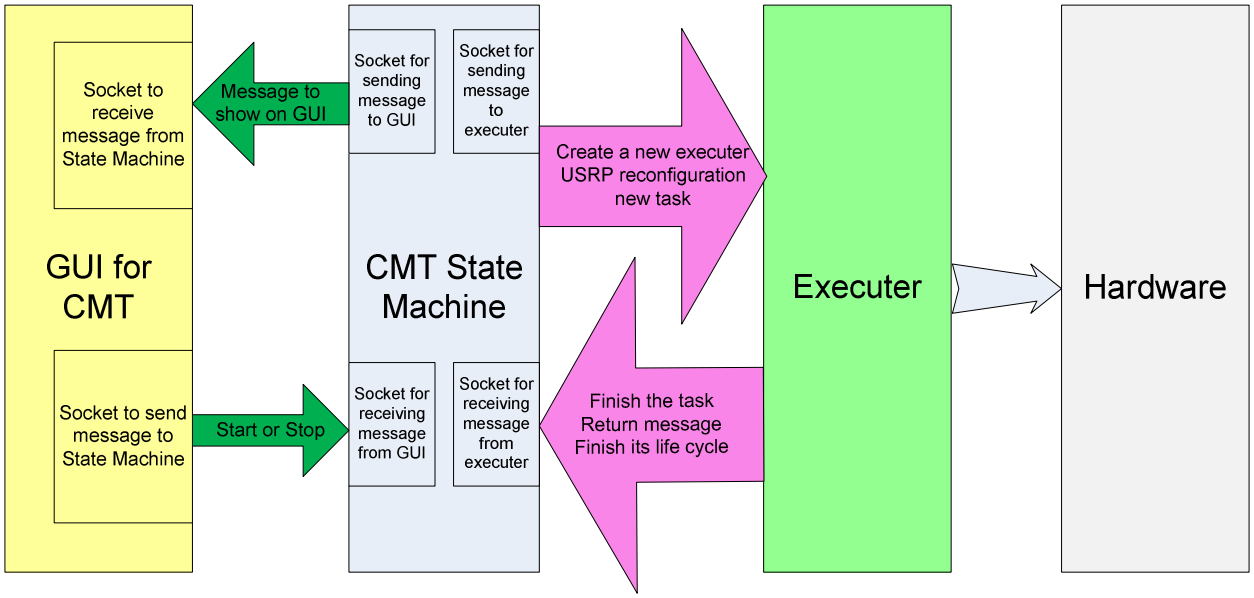

Figure 60: CMT prototype structure

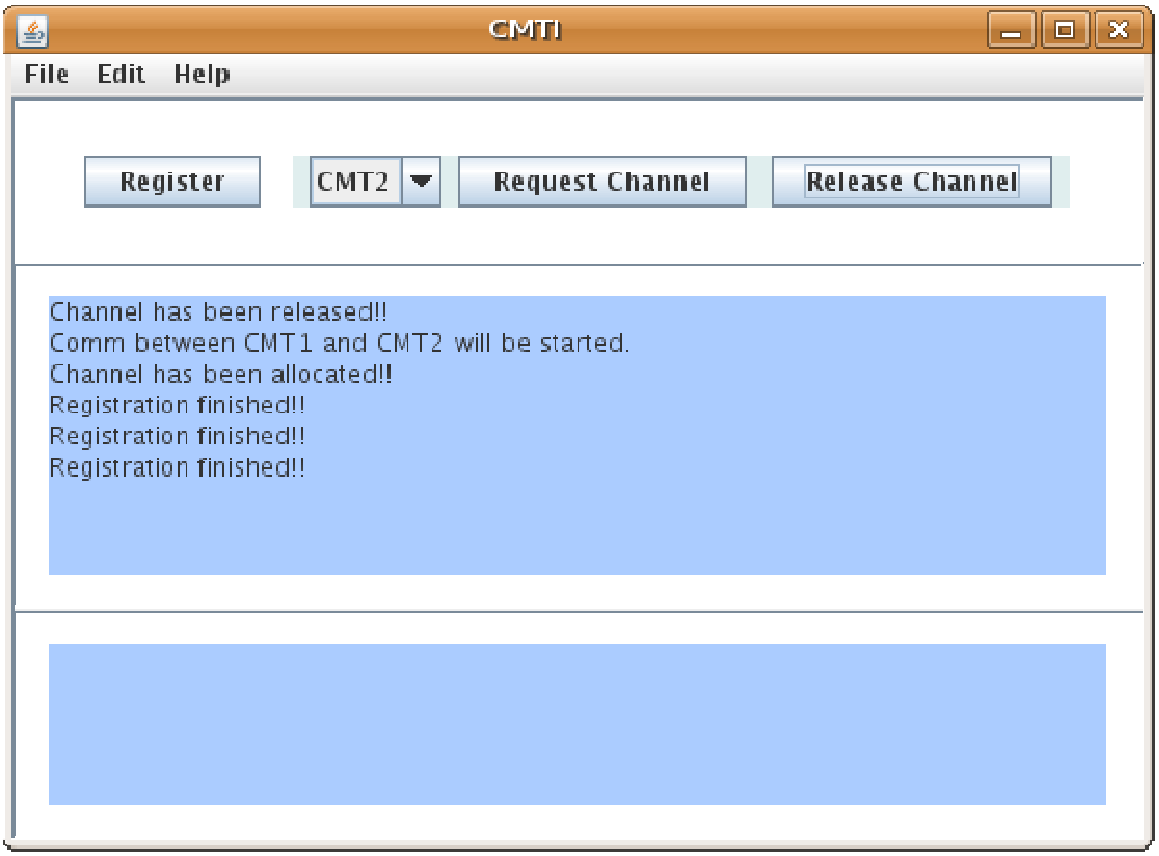

Figure 61: GUI for CMT

The following graph of CMT represents the combination of the state machine and the executer because I want to show a complete process for the energy detection strategy in two-way or three-way hand shaking. The flow graph is shown in Figure 62. One basic principle is detect before transmit, always. In GNU radio and USRP platform, the transmit path and the receive path are automatically switched. Once the payload in the buffer is transmitted o, then, the software and hardware will automatically be switched to the receive path. In the receive path mode, three conditions can occur, energy detected and no respond message received, respond message received, and no energy detected.

During the message exchange for a CMT, if energy is detected and it is not the respond message, the CMT will switch to another randomly chosen channel and start over the message exchange process, except during the termination three-way handshaking. If interrupted by primary user, the node will switch to another available channel and start registration instead of repeat termination. This is because PCN will automatically be aware of the termination when the node is registered again.

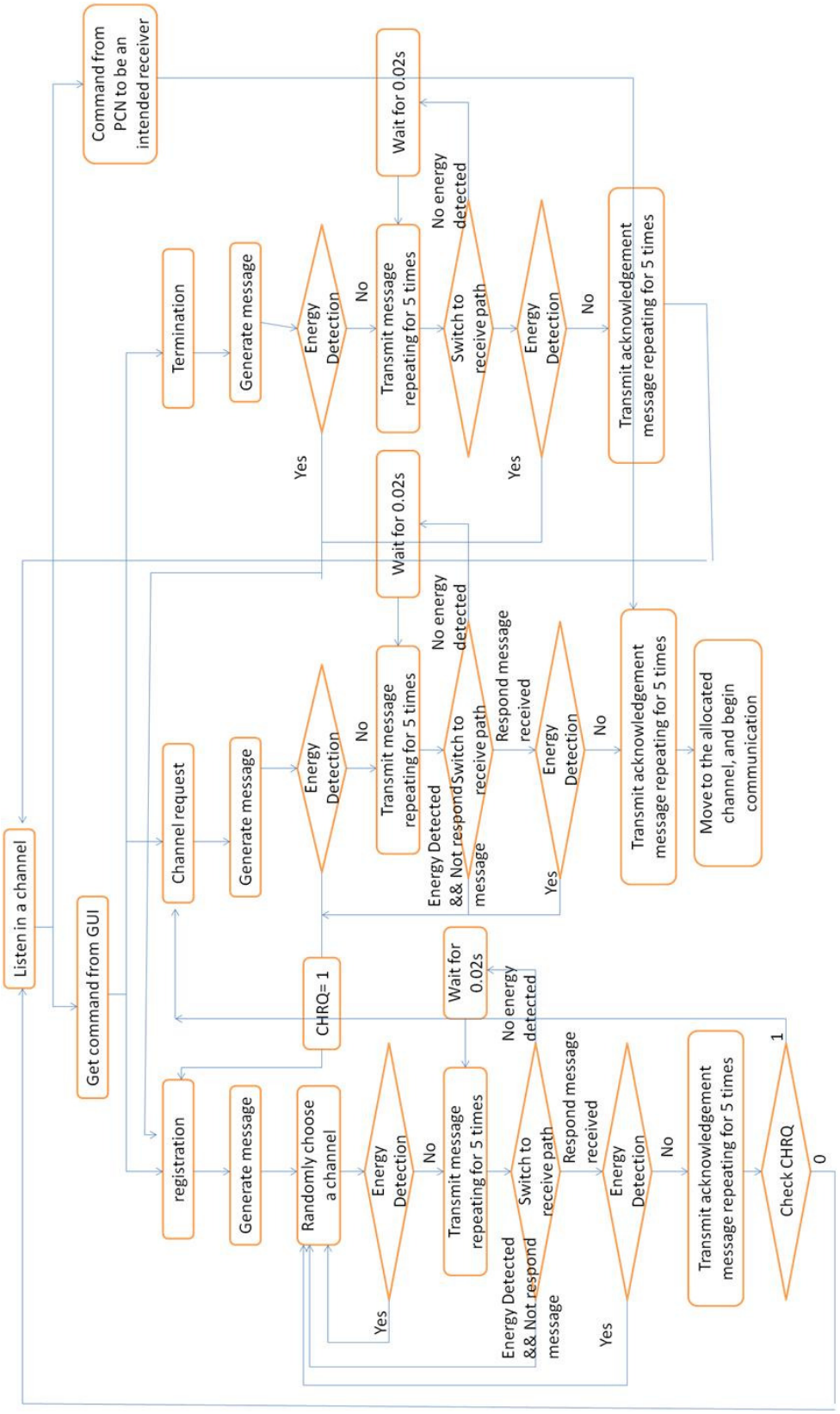

Figure 62: CMT state machine

7.4 Demonstration Setup

Figure 63 gives an overview of the demonstration setting. There are two PCNs and four CMTs. PCNs primarily provide three functions: spectrum management, message forwarding, and interoperability gateway. By performing these three functions, cognitive radio and non cognitive radios are automatically organized to communicate as a secondary user group. Management functions are performed by the USRP on the right side of PCN1 whereas forwarding and gateway functions are performed by the USRP on the left side as shown in Figure 63. Because of the limitations of hardware in our demonstration, this part is shown using simulation. The management function includes CMT registration, request for a channel, and termination of a channel. For communication across different cells, forwarding functions will be activated, and for communication between incompatible radios, gateway functions will provide interoperability.

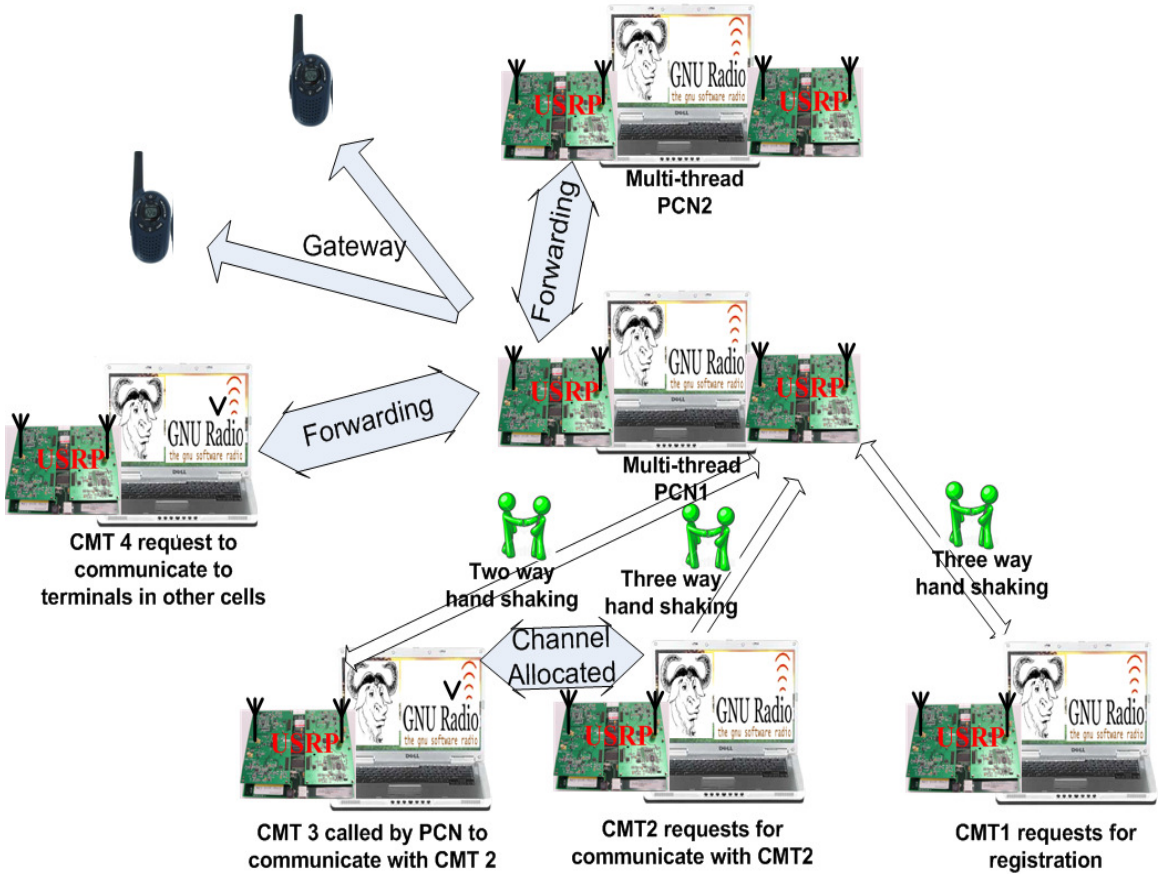

Figure 63: Demonstration setup

Demonstration Activities

During the demonstration, PCNs scan sub channels sequentially, receive requests, allocate and manage channels, record events and optimize channel utilization based on experience and knowledge of channel. PCN management is through a three-way handshaking message exchange. The message types are defined in Table 12.

Table 12: Message definition

	DST-ID	SRC-ID	MSG-TYPE	MSG-STATUS	RX-ID	Ch-Num
REGI (register)	FFFF	CMT#	RQST	REGI	FFFF	#########(462662600)
	CMT#	PCN#	RSPD	REGI	FFFF	#########
	PCN#	CMT#	ACKW	REGI	FFFF	#########
CHAN (request channel)	PCN#	CMT#	RQST	CHAN	CMT#	FFFFFFFFF
	CMT#	PCN#	RSPD	CHAN	CMT#	#########
	CMT#	PCN#	RSPP	CHAN	CMT#	#########
	PCN#	CMT#	ACKW	CHAN	CMT#	#########
	PCN#	CMT#	ACKP	CHAN	CMT#	#########
RESU (resume communication)	PCN#	CMT#	RQST	RESU	CMT#	FFFFFFFFF
	CMT#	PCN#	RSPD	RESU	CMT#	#########
	PCN#	CMT#	ACKW	RESU	CMT#	#########
TERM(terminate communication)	FFFF	CMT#	RQST	TERM	CMT#	#########
	CMT#	PCN#	RSPD	TERM	CMT#	#########
	PCN#	CMT#	ACKW	TERM	CMT#	#########

Because the whole scenario is secondary user based, Carrier Sensed Multiple Access (CSMA) is adopted and the modified CSMA design based MAC control in our system as shown in Figure 2. After a CMT registers with the PCN, it can then setup communications with other members of the DCCS through assistance from the PCN. For example, when CMT1 requests communication with CMT2, a request message will be sent to PCN1 asking for channel allocation. PCN will then authorize CMT2 to communicate with PCN. Based on the features of CMT2 and CMT1, PCN will take different actions, as enumerated in Table 13.

Table 13: PCN activities in different scenarios

	Cognitive Radio	In PCN1's Cell	PCN1 Action
CMT1	YES	YES	Allocate Channel
CMT2	YES	YES	
CMT1	YES	YES	Allocate Channel and Forwarding
CMT2	YES/NO	NO	
CMT1	YES/NO	YES	Ask Cognitive Radio to reconfigure
CMT2	NO/YES	YES	
CMT1	NO	YES	Gateway
CMT2	NO	YES	
CMT1	NO	YES	Gateway and Forwarding
CMT2	NO	NO	

Forwarding and gateway function of the PCN is performed by the left USRP under the control of the right USRP. Forwarding means to forward the package or signal to other PCNs. Gateway is the function to provide interoperability for two nodes which are not compatible. In the DCCS

demonstration, forwarding is performed as 1-hop. The PCN connections are shown using "network simulator" because of the limitation of the hardware. Gateway function is demonstrated to provide FRS radio the interoperability to other radios.

7.5 System Testing

In this section, the testing of DCCS prototype is discussed. Not all the functions in DCCS have been fully implemented in the prototype. Thus, the test cases are limited to the functions that have been implemented in the prototype. In Table 14, the expected result and experimental result is shown and, for the failed test cases, a brief analysis is given.

Table 14: PCN and CMT function testing

No.	Function to be tested	Expected Result	Experimental Result and Analysis
1	Single CMT registration without interruption	Requesting CMT responds within 0.5 second	Requesting CMT responded Time consumed is between (0.1 – 0.5)
		GUI for PCN is updated, including channel condition and cell registration	GUI is updated as expected instantly
		PCN moves to next adjacent channel after the process	PCN moved to the next adjacent channel as expected.
2	Multiple CMTs happened to request registration in the same channel without interruption	One CMT is responded at one round for all the channels.	PCN respond to one CMT request at each time and move to the next adjacent channel after one request is accepted
		Waiting CMTs should not move to other channels	If the SNR in the channel is low, waiting CMTs might move to other channels. Reason: If energy is detected and

			demodulation is not correct, CMT will consider it as primary user interruption.
3	Single CMT registration with interruption of another modulation type	CMT moves to another channel as soon as the energy detected. PCN did not update GUI since the registration is not completed	As expected, CMT randomly chooses another channel. PCN did not update GUI. PCN moved to next adjacent channel.
4	Single CMT registration with interruption in the same modulation type	CMT moves to another channel as soon as the energy detected. PCN doesn't update GUI since the registration is not completed	CMT does not move to another channel but continues detecting current channel availability. PCN donot update GUI since the registration is not completed Solution: more intelligent verification scheme for CMT.
5	CMT requests channel for an intended user within the cell when channels are available	PCN allocated a channel to the CMT and informed the intended receiver.	PCN allocated a channel to the CMT and informed the intended receiver.
6	CMT requests channel for an intended user within the cell when no channels is available	PCN returns the error message to CMT indicating there is no channel available and moves to the next adjacent channel.	PCN returns the error message to CMT indicates there is no channel available and moves to the next adjacent channel.
7	CMT requests channel for an intended user in a neighbor cell when a vacant channel is	PCN allocates a channel, to CMT, updates GUI, and launches its forwarding function.	PCN allocates a channel to CMT, updates GUI and launches its forwarding function.

	available		
8	CMT requests channel for an intended user in a neighbor cell when no channel is available	PCN returns the error message to CMT indicating there is no channel available and moves to the next adjacent channel.	PCN returns the error message to CMT indicating there is no channel available and moves to the next adjacent channel.
9	CMT requests channel for an intended user not registered in the system	PCN returns the error message to CMT indicating there is no information about the intended receiver in the system	PCN returns the error message to CMT indicating there is no information about the intended receiver in the system
10	Video transmission between two CMTs in the allocated channel	Observing video quality	Video quality is good
11	Termination request without interruption	CMT get confirmation of termination request PCN updates GUI	CMT gets confirmation for termination request PCN updates GUI
12	Communication interrupted	Resume communication should be performed automatically for CMT and PCN within 0.5 s	Resume communication is performed automatically for CMT and PCN. When no further interruption, the task is finished within 0.5s. With further interruption during the resume, the task takes longer than 0.5s.

7.6 Hardware Settings

The hardware settings are shown in Table 15. It includes frequency range, bandwidth, peak input power to antenna, antenna polarization and gain, and waveform modulation types.

Table 15: Hardware settings

<table>
<tr><td>Frequency Ranges</td><td>
<table>
<tr><th>Channel</th><th>Centre Freq. (MHz)</th><th>Max ERP</th><th>BW (MHz)</th><th>Mobile</th></tr>
<tr><td>1</td><td>231.2250</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>2</td><td>233.0250</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>3</td><td>234.8250</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>4</td><td>236.6250</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>5</td><td>238.4250</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>6</td><td>386.8750</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>7</td><td>396.8750</td><td>10 W (10dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>8</td><td>406.9750</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>9</td><td>408.7750</td><td>10 W (10dBW)</td><td>1.75</td><td>Yes</td></tr>
<tr><td>10</td><td>436.8750</td><td>1 W (0dBW)</td><td>1.75</td><td>Yes</td></tr>
</table>
</td></tr>
<tr><td>Bandwidth</td><td>DCCS can work either as the primary user or as the secondary user:
Primary user: 3.5M
Secondary user: No requirement</td></tr>
<tr><td>Peak Input Power to Antenna</td><td>0 dBm</td></tr>
<tr><td>Antenna polarization and gain</td><td>Vertical omnidirectional,
Gain: 3dB relative</td></tr>
<tr><td>Waveforms (modulation types)</td><td>FM, AM, MPSK, QAM and OFDM</td></tr>
</table>

DCCS is still under developing. More functions will be added into the prototype. The current development status is shown in Figure 52 in Section 7.1

Chapter 8 Conclusion and Future Work

8.1 Summary of Research Results

This dissertation mainly includes the following 5 groups of contributions.

1. **Innovative cognitive radio network architecture is designed.**

Traditional communication systems like the GSM are based on pre-defined channel allocation and sets of protocols. Relying on fixed base stations reduces robustness in emergently situations. Predefined channel allocation schemes produce resource challenges because of the scarcity of current spectrum. Technology development history decides that cellular system and Wi-Fi infrastructure are independent and are utilizing different spectrum and devices. This generates lack of interoperability among devices and wastes spectrum. The architecture of DCCS is designed to solve the above challenges. Similar to a cellular cognitive radio, base station and mobile terminals both exist in DCCS networks. The difference is that the CR nodes in DCCS can switch the roles between base stations PCN and mobile terminals CMT, based on its observation of the network distribution. Each of the nodes appearing in the area has the ability of sensing the environment and being aware of positioning in the cell of an existing PCN. Based on the result, it will automatically decide whether to switch to a PCN status. Cell splitting and merging and cell adjustment are used to guarantee an optimal use of the resources. The dynamic connections among PCNs are wideband; those of a mobile ad hoc network. Within each cell, a PCN manages the spectrum, allowing a group of secondary users accessing the spectrum to avoid interference with primary users and users outside of the cell using different types of devices and modulations. Frequency reuse technology is adopted for efficiently utilizing the spectrum. A PCN can serve as a gateway to other infrastructure, a forwarding relay, a gateway, and a spectrum manager. DCCS can expand to a broad geographical distribution via linking to existing infrastructure. The DCCS can quickly form a network to accommodate a diverse set of devices in natural disaster areas, for example, the areas affected by earthquake and Katrina. It can also recover the infrastructure in a

blind spot, for example, subway or mountain areas. Portable size and cost reduction enable the feasibility of its commercial applications.

2. **A comprehensive and effective DSA scheme is developed.**

A complete DSA scheme including spectrum sensing, information sharing and maximum throughput channel allocation is developed in DCCS. DCCS accesses the spectrum as a secondary network, which means all the members in the network access the spectrum as secondary users. It is important to avoid interference to primary user or any other users outside of the network. It is also important to optimally cooperate in resource utilization among the members within the network. The DSA scheme takes the two issues into consideration. The spectrum sensing in DCCS is mainly through energy detection. A CMT uses energy detection to find an available channel to communicate with a PCN. The PCN has collected current and history information for all the channels within its cell. An optimal channel allocation algorithm is designed. The collected information is the input of the algorithm and the output is the channel with the maximum throughput. A MAC layer protocol is defined to avoid the collision among members who want to access the spectrum at the same time to transmit control messages.

3. **A complete universal classification and synchronization system is designed, implemented, integrated and tested.**

A universal classification and synchronization system is designed, implemented, integrated and tested. It is conceived as such a self-contained system which can detect, classify, synchronize and provide all parameters needed for physical layer demodulation. The designed system has been verified by a prototype using GNU Radio and USRP platform. Performance for key components and the entire system has been evaluated by theoretical analysis, OTA experiments and computer simulations.

The necessity of efficiently utilizing scarce spectrum by implementing DSA and Cognitive Radio technology is gaining rapidly growing attention. During the resource optimization process of DSA and CR technology, transmitter side and receiver side need to be coordinate in order to correctly demodulate the signal. Compared to the other approach “common control channel”, UCS is more reliable, flexible, and is implemented in an inexpensive radio platform. These advantages of UCS provide promising marketing prospects. We systemically analyzed the design of extracting demodulation parameters from a signal itself, practically provided solutions to

implementing entire UCS system in SDR platforms, and comprehensively evaluated the system performance by OTA experiments and theory.

4. **A series of protocols and algorithms are designed to support DCCS functions.**

These protocols include intra-cell three-way handshaking protocol, PCN and CMT message exchange protocol, inter-cell data communication protocol, inter-cell message communication protocol, etc. A channel allocation algorithm is designed to increase the throughput in a dynamic channel environment. Cell splitting and merging algorithm is designed to reduce the handoff and link drop for intra-cell communications. DCCS is an expandable system. To make it more robust and flexible, more protocols and algorithms can be researched and added.

5. **DCCS prototype is implemented.**

A prototype of DCCS is implemented, tested and demonstrated. This prototype provides high quality communications among a diverse set of cognitive radio (CR) nodes while minimizing interference to primary and other secondary users. PCN plays a role similar to a base station with the additional role of registering and directing one or more CMTs. Both PCNs and CMTs operate as secondary users sharing spectrum with legacy analog and digital users coexisting in the band. Adjacent cells operate in different frequency ranges to avoid interference. PCNs provide the management of the spectrum, including registering CMTs, and allocating channels and forwarding packets if the source and destination nodes are located in different cells. PCNs also provide gateway functions for interoperability among cognitive radios and non-cognitive radios. One-hop connection is shown for inter-cell communication. With more hardware support, more functions could be added to the prototype.

Among the listed contributions, a summary of the research contributions is shown in the following:

1. Algorithm

   The cell splitting and merging algorithm is designed for the purpose of efficient cell coverage in DCCS network. A maximum throughput channel allocation algorithm in a DSA scenario is designed to compare the channel quality and rank the preference of the channel. A set of algorithms in universal classification and synchronization (UCS) is co

developed by Qinqin Chen and me. A channel grouping algorithm is adopted for using OFDMA as the channel collection method in inter-cell communications.

2. Performance Analysis

   The performance of maximum throughput channel allocation algorithm is also analyzed. The necessary performance analysis in each step of UCS is shown.

3. Protocols

   Three way handshaking of message exchange in a DSA scenario is developed. A set of protocols to support inter-cell communications and intra-cell communication functions is developed as well.

4. Implementations

   The core functions of DCCS are implemented and demonstrated. UCS is implemented and demonstrated both as an independent system and part of DCCS.

8.2 Future Work

Future work of DCCS includes continuing to implement the previously designed functions and the fully test for each block. From research point of view, there are several opportunities for future academic exploration.

1. Security mechanisms should be considered for future development.

   In the current stage, security has not been considered in the design. The control message is not encrypted. Some of the security designs in IEEE 802.11i, for example, four-way hand shaking, or group hand shaking can be modified and adopted[74].

2. Multi-threading design should be further developed on PCN.

   PCN serves as a small base station. It performs both the intra-cell management function and forwarding function. In current stage of DCCS, the intra-cell management is done in a serial mode. PCN go through each sub-channels. This mode limits the number of sub-channels and the performance speed. Multi-hop forwarding also requires multi-threading.

3. A more complicated prototype including large number of nodes.

   Current DCCS prototype is a simple proof of concept implementation. The future prototype should include 50-100 nodes including different type of devices, and tests in

different scenarios. It should also be able to provide the interface with standard network in billing, roaming and other issues.

Bibliography

[1] "USRP http://www.ettus.com/."

[2] "Lyrtech Small Form Factor (SFF) SDR platform, Available: http://www.lyrtech.com/."

[3] "Gnuradio: http://gnuradio.org/trac."

[4] "OSSIE: http://ossie.mprg.org/."

[5] S. K. Jones and T. W. Phillips, "Plan for Tests of Prototype Personal/Portable TV White Space Devices (Phase II)," FCC, Ed., 2008.

[6] S. Soliman, "Cognitive Radio: Key Performance Indicators," in *BWRC Cognitive Radio Workshop* Berkeley, CA, 2004.

[7] A. Goldsmith, *Wireless Communications*: Cambridge University Press, 2005

[8] "SOFDM: http://www.conniq.com/WiMAX/fdm-ofdm-ofdma-sofdma-01.htm."

[9] D. Jones, "Sprint Eyes WiMax Backhaul," in *Light Reading*, 2006.

[10] "TV white space: http://www.fcc.gov/oet/projects/tvbanddevice/Welcome.html."

[11] IEEE 802.18 Radio Regulatory Technical Advisory Group, "In the MIn the Matter of Unlicensed Operation in the TV Broadcast Bands, Motion for Extension of Time, IEEE 802.18 Radio Regulatory Technical Advisory Group." vol. ET Docket 04-186, 2004.

[12] Y. Wang and C. W. Bostian, "Dynamic Cellular Cognitive System," in *Publication number: 20090215457, U.S. Patent Application No.: 12/392,419*, Patent and Trademark Office, Ed., Feb, 2009.

[13] H. Rheingold, *Smart Mobs: The Next Social Revolution*: Macquarie University, 2002.

[14] O. K. Tonguz and G. Ferrari, *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective*: Wiley, 2006.

[15] M. Conti and S. Giordano, "Multihop Ad Hoc Networking: The Theory," *IEEE Communications Magazine,* vol. 45, pp. 78-86, 2007.

[16] P. Gould, "3G Radio Network Planning Managing Cell Breathing," in *IIR Cell Planning Technical Forum,* Dublin, 2001.

[17] M. S. Gast, *802.11 Wireless Networks*: O'Reilly, 2005.

[18] B. Zhou, K. Xu, and M. Gerla, "Group and Swarm Mobility Models for ad hoc Network Scenario Using Virtual Tracks," in *IEEE Military Communications Conference*, 2004.

[19] M. Jiang, J. Li, and Y. C. Tay, "Cluster Based Routing Protocol(CBRP) Functional Specification," internet-draft, Ed., 1998.

[20] B. Das, R. Sivakumar, and V. Bharghavan, "Routing in Ad Hoc Networks Using a Spine," in *Proceedings of 6th International Conference on Computer Communications and Networks,* Las Vegas, 1997.

[21] R.-H. Hwang, C.-Y. Li, C.-Y. Wang, and Y.-S. Chen, "Mobile IPv6-Based Ad Hoc Networks: Its Development and Application," *IEEE Journal on Selected Areas in Communications,* vol. 23, 2005.

[22] Y. Wang, H. Chen, X. Yang, and D. Zhang, "Cluster Based Location-Aware Routing Protocol for Large Scale Heterogeneous MANET," in *Second International Multisymposium on Computer and Computational Sciences*, 2007.

[23] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wirless network: A survey," *Computer Networks,* vol. 50, pp. 2127–2159, 2006.

[24] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environment," *Proc. IEEE DySPAN,* pp. 131-136, 2005.

[25] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for dynamic spectrum access," *Proc. IEEE DySPAN,* pp. 224-231, 2005.

[26] D. C. S.M. Nishra, C. Chang, D. Willkomm, B. Schewick, A. Wolisz, R.W. Brodersen, "A real time cognitive radio testbed for physical and link layer experiments," *Proc. IEEE DySPAN,* pp. 562–567, 2005.

[27] FCC, "ET Docket No 03-237 Notice of inquiry and notice of proposed Rulemaking," 2003.

[28] A. Sahai, D. Cabric, and R. W. Brodersen, "Spectrum Sensing Fundamental Limits and Practical Challenges," in *Proc. Tutorial IEEE DySPAN,* Baltimore, MD, 2005.

[29] J. S. D. II, "http://www.wirelesscommunication.nl/reference/chaptr04/cellplan/dca.htm."

[30] E. D. Re, R. Fantacci, and G. Giambene, "Handover Queuing Stategies with Dynamic and Fixed Channel Allocation Techniques in Low Earth Orbit Mobile Satellite Systems," *IEEE Transactions on Communications,* vol. 47, 1999.

[31] M. Zhang and T.S.Yum, "Comparisons of Channel-Assignment Strategies in Cellular Mobile Telephone Systems," *IEEE Transactions on Vehicular Technology,* vol. 38, 1989.

[32] H. Jiang and S. S. Rappaport, "CBWL: A New Channel Assignment and Sharing Method for Cellular Communication Systems," *IEEE Transactions on Vehicular Technology,* vol. 43, 1994.

[33] B. Wild and K. Ramchandran, "Detecting Primary Receivers for Cognitive Radio Applications," *Proc. IEEE DySPAN,* pp. 124-130, 2005.

[34] Q. Zhao, "Spectrum Opportunity and interference constraint in opportunistic spectrum access," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP),* 2007.

[35] Q. Zhao and B. M. Sadler, "A Survey of Dynamic Spectrum Access," *IEEE Signal Processing Magazine,* vol. 24, pp. 78-89, 2007.

[36] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE Journal on Selected Areas in Communications,* vol. 25, pp. 589-600, 2007.

[37] T. Luo, T. Jiang, W. Xiang, and H.-H. Chen, "A Subcarriers Allocation Scheme for Cognitive Radio Systems Based on Multi-Carrier Modulation," *IEEE Transactions on Wireless Communications,* vol. 7, pp. 3335-3340, 2008.

[38] M. Y. ElNainay, F. A. Ge, Y. Wang, A. E. Hilal, Yongsheng (Sam) Shi, A. B. MacKenzie, and C. W. Bostian, "Channel Allocation for Dynamic Spectrum Access

Cognitive Networks using Localized Island Genetic Algorithm," in *Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops*, 2009.

[39] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-Throughput Tradeoff for Cognitive Radio Networks," *IEEE Transactions on Wireless Communications,* vol. 7, 2008.

[40] R. E. Irwin and L. A. DaSilva, "Channel Assignment based on Routing Decisions (CARD): Traffic-Dependent Topology Control for Multi-Channel Networks," in *ICC Workshops*, 2009.

[41] H.-j. Liu, S.-f. Li, Z. Wang, W. Hong, and M. Yi, "Strategy of Dynamic Spectrum Access Based-on Spectrum Pool," in *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference*, 2008.

[42] Y. Wang, Q. Chen, and C. W. Bostian, "Universal Classification and Synchronization," *Journal of Autonomous and Adaptive Communications Systems,* 2009.

[43] E. Azzouz and A. K. Nandi, *Automatic modulation recognition of communication signals*. Boston: Kluwer Academic Publishers, 1996.

[44] N. Warke and G. C. Orsak, "'A universal methodology for signal classification in non-Gaussian environments'," in *Sixth IEEE Digital Signal Processing Workshop*, Yosemite National Park, CA, 1994, pp. 101 - 104.

[45] Q. Chen, Y. Wang, and C. W. Bostian, "'Universal classifier synchronizer demodulator'," in *the 1st international workshop on Dynamic Spectrum Access and Cognitive Radio Networks (DSA-CRN' 08) joint with the 27th IEEE IPCCC* Austin, TX, 2008, pp. 366-371.

[46] Y. Wang, Q. Chen, and C. W. Bostian, "Universal classification and synchronization," in *Microsoft Cognitive Wireless Networking Summit*, Snoqualmie, Washington, 2008.

[47] B. Le, "Building a cognitive radio: from architecture definition to prototype implementation," *Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA,* 2007.

[48] A. Polydoros and K. Kim, "On the detection and classification of quadrature digital modulations in broad-band noise," *IEEE Transactions on Communications,* vol. 38, pp. 1199-1211, August 1990 1990.

[49] B. F. Beidas and C. L. Weber, "Asynchronous classification of MFSK signals using the higher order correlation domain," *IEEE Transcations on Communications,* vol. 46, pp. 480-494, 1998.

[50] P. Jahankhani, V. Kodogiannis, and K. Revett, "EEG signal classification using wavelet feature extraction and neural networks," in *IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing*, 2006, pp. 120-124.

[51] A. Prochazka, J. Kukal, and O. Vysata, "Wavelet transform use for feature extraction and EEG signal segments classification," in *3rd International Symposium on Communications, Control and Signal Processing (ISCCSP),* Malta, 2008, pp. 719-722.

[52] S.-Z. Hsue and S. S. Soliman, "Automatic modulation classification using zero crossing," *Radar and Signal Processing, IEE Proceedings F,* vol. 137, pp. 459-464, 1990.

[53] T. Yucek and H. Arslan, "OFDM signal identification and transmission parameter estimation for cognitive radio applications," in *IEEE Global Telecommunications Conference*, 2007, pp. 4056-4060.

[54] K. Kim, I. A. Akbar, K. K. Bae, J.-s. Uml, C. M. Spoonerll, and J. H. Reed, "Cyclostationary approaches to signal detection and classification in cognitive radio," in *IEEE DySPAN*, pp. 212-215, 2007

[55] B. Le, T. W. Rondeau, and C. W. Bostian, "Cognitive radio realities," *Wireless Communications and Mobile Computing,* vol. 7, pp. 1037 - 1048, 2007.

[56] T. W. Rondeau, "Application of artificial intelligence to wireless communications," *Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA,* 2007.

[57] J. G. Proakis and M. Salehi, *Communication Systems Engineering*: Prentice Hall, 2002.

[58] Y. Wang, S. Nair, A. Young, Q. Chen, and C. W. Bostian, "'OFDM signal classification and synchronization for cognitive radio systems'," in *SDR Forum Technical Conference* Washington D.C., 2008.

[59] T. S. Rappaport, *Wireless Communications Principles and Practice*: Prentice Hall, 2002.

[60] J. G. Proakis, *Communication Systems*: Prentice Hall, 2003.

[61] Y. Wang, Q. Chen, B. Le, and C. W. Bostian, "Universal synchronizer design for cognitive radio," in *SDR Forum Technical Conference,* Denver, CO, 2007.

[62] N. M. Blachman, "Gaussian noise part II : distribution of phase change of narrow-band noise plus sinusoid," *IEEE Transactions on Information Theory,* vol. 34, pp. 1401-1405, 1988.

[63] R. Chandra, et al "Adapting channel widths to improve application performance," in *Microsoft Cognitive Wireless Networking Summit* Snoqualmie, WA, 2008.

[64] J. G. Proakis, *Digital Communications*: McGraw-Hill Companies, Inc. , 2001.

[65] S. Haykin, *Communication Systems*: John Wiley & Sons, Inc, 2003.

[66] I. D. Chakeres and C. E. Perkins, "Dynamic MANET On-demand Routing Protocol," *IETF Internet Draft,* 2008.

[67] Z. Han and K. J. R. Liu, *Resource Allocation for Wireless Networks*: Cambridge University Press, 2008.

[68] " IEEE 802.22-09/0029r0: Wireless RANs," 2009.

[69] "Ninth Notice of Proposed Rulemaking," FCC, Ed., 2006.

[70] K. Sohrabi and G. J. Pottie, "Performance Of A Novel Self-Organization Protocol For Wireless Ad-Hoc Sensor Networks," in *IEEE Vehicular Technology Conference*. vol. 2, 1999.

[71] S. J. Shellhammer, A. K. Sadek, and W. Zhang, "Technical Challenges for Cognitive Radio in the TV White Space Spectrum," in *2009 Information Theory and Applications Workshop* San Diego, CA, 2009.

[72] R. A. Santosa, B.-S. Lee, C. K. Yeo, and T. M. Lim, "Distributed Neighbor Discovery in Ad Hoc Networks Using Directional Antennas," *Proceedings of The Sixth IEEE International Conference on Computer and Information Technology,* 2006.

[73] S. A. Borbash, A. Ephremides, and M. J. McGlynn, "An asynchronous neighbor discovery algorithm for wireless sensor networks," *Ad Hoc Networks,* vol. 5, 2007.

[74] "IEEE 802.11i-2004: Amendment 6: Medium Access Control (MAC) Security Enhancements IEEE Standards," 2004.