

**An Application Layer Approach to Cognitive Radio Networks  
Using SIP Messaging**

By

**Jia Long, B. A. Sc.**

A thesis submitted to the Faculty of Graduate Studies and Research in  
partial fulfillment of the requirements for the degree of

**Master of Applied Science in Electrical Engineering**

Ottawa-Carleton Institute for Electrical and Computer Engineering (OCIECE)

Department of Systems and Computer Engineering

Carleton University

Ottawa, Ontario, Canada, K1S 5B6

September 2011

© Copyright 2011, Jia Long



Library and Archives  
Canada

Published Heritage  
Branch

395 Wellington Street  
Ottawa ON K1A 0N4  
Canada

Bibliothèque et  
Archives Canada

Direction du  
Patrimoine de l'édition

395, rue Wellington  
Ottawa ON K1A 0N4  
Canada

*Your file Votre référence*

*ISBN: 978-0-494-87799-9*

*Our file Notre référence*

*ISBN: 978-0-494-87799-9*

#### NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

#### AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

---

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

The undersigned recommend to  
the Faculty of Graduate Studies and Research  
acceptance of the thesis

**An Application Layer Approach to Cognitive Radio Networks  
Using SIP Messaging**

submitted by

**Jia Long, B.A.Sc.**

in partial fulfilment of the requirements for the degree of  
**Master of Applied Science in Electrical Engineering**

---

Chair, Howard Schwartz, Department of Systems and Computer Engineering

---

Thesis Supervisor, Changcheng Huang

---

Thesis Co-Supervisor, James Yan,

Carleton University  
September 2011

# **Abstract**

This thesis proposes an overlay approach to address some of the implementation issues regarding cognitive radio networks and spectrum management in dynamic environments. By utilizing application layer protocol, Session Initiation Protocol (SIP), a centralized cognitive radio network can be made as an overlay to existing wireless networks such as 3GPP's Long Term Evolution (LTE) network architecture. We propose using two extensions to SIP protocol, Event Notification Framework (ENF) and Event State Publication (ESP), in order to implement our design. We provide event scenarios to demonstrate how our system allows primary users to share their spectrum usage information with secondary cognitive radio users in order to provide better spectrum management. We also provide analytical and simulation results on SIP messaging delay to demonstrate the efficiency of our system in terms of channel utilization. Results from analysis and simulation using OPNET show that our proposed system can be efficient under certain scenarios.

## **Acknowledgements**

I would like to give special thanks to my supervisors, Professor C. Huang, and Professor J. Yan, for their effort and continued guidance throughout my degree.

My sincerest gratitude goes out to my parents as well as my family and friends who supported me through this long period of my life. I will never forget you. You have always been there when I have needed you the most. Thank you all.

We appreciate the financial support from the NSERC Grant #CRDPJ 354729-07 and the OCE Grant #CA-ST-150764-8. This work was also supported in part by Nortel Networks.

# Table of Contents

<b>Abstract</b>	<b>i</b>
<b>Acknowledgement</b>	<b>ii</b>
<b>Table of Contents</b>	<b>iii</b>
<b>List of Tables and Figures</b>	<b>v</b>
<b>List of Acronyms</b>	<b>vii</b>
<b>List of Mathematical Notations</b>	<b>ix</b>
<b>Chapter 1 Introduction</b>	<b>1</b>
1.1 Background	1
1.2 Research Problems	4
1.3 Objectives	6
1.4 Our Approach	7
1.5 Contributions	9
1.6 Thesis Organization	10
1.7 Publications	12
<b>Chapter 2 Literature Survey</b>	<b>13</b>
2.1 Fundamentals	13
2.2 Cognitive Radio Classifications	15
2.3 Spectrum Sensing	17
2.3.1 Energy Detection	18
2.3.2 Feature Detection	19
2.3.3 Matched Filter Detection	20
2.4 Cooperative Sensing	22
2.5 Spectrum Scanning Techniques	24
2.6 Dynamic Spectrum Allocation	26
2.7 Spectrum Mobility	28
2.8 Similar Works	29
2.9 Implementation Challenges	30
<b>Chapter 3 Application-Layer System Solution</b>	<b>33</b>
3.1 Solution Architecture	33
3.1.1 System Entities and Overall Network Overlay	34

3.1.2	Advantages of our Solution . . . . .	40
3.1.3	Disadvantages of our Solution. . . . .	45
3.2	Messaging Model . . . . .	46
3.2.1	SIP Overview . . . . .	47
3.2.2	Messaging Entities . . . . .	49
3.2.3	Messages for Proposed Solution . . . . .	51
3.3	Auctioning Scheme . . . . .	54
3.4	Criteria for Evaluation . . . . .	55
<b>Chapter 4 Analysis and Simulation Evaluation . . . . .</b>		<b>59</b>
4.1	Delay Components . . . . .	59
4.1.1	Access Network . . . . .	63
4.1.2	First TCP Connection . . . . .	64
4.1.3	Other TCP Connections . . . . .	68
4.1.4	Overall Delay . . . . .	70
4.2	Message Flows . . . . .	70
4.2.1	Secondary User Arrival . . . . .	71
4.2.2	Insufficient Channels . . . . .	74
4.2.3	Primary User Arrival . . . . .	76
4.2.4	User Departures . . . . .	79
4.2.5	Spectrum Handoff . . . . .	81
4.2.6	Subscription Timeout . . . . .	84
4.2.7	User Mobility . . . . .	85
4.3	Channel Efficiency Analysis . . . . .	86
4.4	Motivation for Testing . . . . .	89
4.5	Simulation Model . . . . .	90
4.6	Simulation Results for Messaging Delays . . . . .	95
4.6.1	Secondary User Arrival . . . . .	95
4.6.2	Primary User Arrival . . . . .	97
4.6.3	Secondary User Pre-empted and Handoff . . . . .	98
4.7	Simulation Results for Channel Efficiency . . . . .	104
4.7.1	Analytical and Simulation Results . . . . .	105
4.8	Simulation Results for Auction Mechanism . . . . .	108
4.9	Comparison of Results . . . . .	112
<b>Chapter 5 Conclusion and Future Work . . . . .</b>		<b>115</b>
<b>References . . . . .</b>		<b>118</b>

# **List of Tables and Figures**

## **List of Tables**

Table 2.1: Advantages and Disadvantages of Sensing Methods

Table 4.1: SIP Message Sizes

Table 4.2: LTE component delays

Table 4.3: Auctioning Simulation Parameters

## **List of Figures**

Figure 1.1: Wireless Technology Spectrum Ranges

Figure 3.1: Relationship between Entities

Figure 3.2: Overall Network over LTE Architecture

Figure 3.3: SIP Network

Figure 3.4: Protocol Stack

Figure 3.5: ESP and ENF Network Entities

Figure 3.6: Detection Cycle

Figure 3.7: Working Model

Figure 4.1: Network Delay Components

Figure 4.2: Secondary User Arrival and Notification

Figure 4.3: Secondary User Wait when no Channels Available

Figure 4.4: Primary User Arrival pre-empts Secondary User

Figure 4.5: Notifying Multiple Secondary Users

Figure 4.6: Spectrum Handoff

Figure 4.7: Subscription Timeout

**Figure 4.8: Simulation Entities**

**Figure 4.9: Delay for varying FER of secondary user SUBSCRIBE**

**Figure 4.10: Delay of various FER for primary user PUBLISH**

**Figure 4.11: Handoff delay for secondary users**

**Figure 4.12: Delay with varying network entities and loads**

**Figure 4.13: Simulation delays for secondary user arrival**

**Figure 4.14: Simulation Delay with Varying Network Entities and Loads**

**Figure 4.15: Channel Efficiency with Arrival Rate 60/sec**

**Figure 4.16: Channel Efficiency with Arrival Rates 2/sec**

**Figure 4.17: Average Winning Bid**

**Figure 4.18: Revenue Gain over 1 second**

# List of Acronyms

3G: Third Generation

3GPP: Third Generation Partnership Project

eNB: Evolved Node B

ENF: Event Notification Framework

ES: Event State

ESP: Event State Publication

E-UTRAN: Evolved Universal Terrestrial Radio Access Network

IETF: Internet Engineering Taskforce

IMS: IP Multimedia Subsystem

IP: Internet Protocol

LTE: Long Term Evolution

MAC: Media Access Control

MME: Mobility Management Entity

OFDM: Orthogonal Frequency Division Multiplexing

PDN: Packet Data Network

QoS: Quality of Service

RTS: Request to Send

SAE: System Architecture Evolution

SIP: Session Initiation Protocol

S-GW: Service Gateway

SNR: Signal to Noise Ratio

TCP: Transmission Control Protocol

UDP: User Datagram Protocol

URI: Uniform Resource Identifier

VCG: Vickrey Clarke Groves

WLAN: Wireless Local Area Network

XML: Extensible Markup Language

## List of Mathematical Notations

$b$  : The set of available channels

$a$  : The set of users

$V_a^b$  : The maximum total value achievable for all combinations of  $M$  and  $N$

$V_{a-j}^b$  : The maximum total value achievable with bidder  $j$  removed

$V_{a-j}^{b-i}$  : The value of the auction with bidder  $j$ 's bid for channel  $i$  removed

$T'$  : Radio link delay from user to eNodeB

$T''$ : Radio link delay from eNodeB to user

$Q'$ : Estimated queuing/processing delay of Evolved Packet Core

$Q''$ : Estimated queuing/processing delay of eNodeB

$U$ : Queuing delay at user node

$\mu_r$  : SIP messages service rate at the user node

$\mu_s$  : SIP message service rate at SIP proxy and spectrum manager

$\lambda_r$  : SIP message arrival rate at the user node

$\lambda_s$  : SIP message arrival rate at SIP proxies and spectrum manager

$\lambda_0$  : Non-SIP message arrival rate at the spectrum manager

$A'$ : Delay from user to the first SIP proxy with no retransmissions

$A''$ : Delay from the first SIP proxy to the user with no retransmissions

$T'(l)$  : Initial retransmission timer for transmission between user to first SIP proxy

$T^x(l)$ : Initial retransmission timer for transmission between SIP network entities

$D$  : propagation delay over access network

$\tau$  : Inter-frame time

$k$  : number of frames in a TCP segment

$p$  : probability of frame error

$q$  : probability of retransmission  
 $T^P$ : queuing delay at the SIP proxies  
 $B$ : one way delay between SIP proxies with no retransmissions  
 $C$ : one way delay from last SIP proxy to the spectrum manager with no retransmissions  
 $M$  : the maximum amount of retransmissions  
 $\rho_s$  : load at the SIP proxies and spectrum manager of SIP messages  
 $\rho_0$  : load at the spectrum manager of non SIP messages  
 $I$ : Internet delay  
 $L$ : packet loss probability  
 $X_0$ : Mean service time of Non-SIP messages at the spectrum manager  
 $X_s$ : Mean service time of SIP messages at the spectrum manager  
 $J_0$ : Number of non-SIP messages waiting in the queue at spectrum manager  
 $J_s$ : Number of SIP messages waiting in the queue at spectrum manager  
 $R$ : Mean residual service time for messages being served  
 $T^0$ : Mean queuing delay at spectrum manager for non-SIP messages  
 $T^m$ : Mean queuing delay at spectrum manager for SIP messages  
 $S$  : Number of network entities  
 $N$  : Number of messages sent during session  
 $T_1$ : Average delay from user to first SIP proxy server  
 $T_2$ : Average delay between SIP proxy servers  
 $T_3$ : Average delay between last SIP proxy and spectrum manager  
 $T^A$ : Average end-to-end message delay for secondary user arrival  
 $T^B$ : Average end-to-end message delay when insufficient channels on arrival  
 $T^C$ : Average end-to-end message delay when subscription timeout

## **Chapter 1**

### **Introduction**

This chapter will provide background information on cognitive radios. It will describe the features they possess, discuss their increasing relevance in wireless mobility for the future, as well as solving emerging problems within the wireless community. It will then briefly introduce the traditional methods of implementation that rely on spectrum sensing in order to realize the potential of cognitive radio devices. From these methods, we form our research problems that still need to be addressed. This provides us with the motivation to develop our method of using a centralized server and higher layer protocols as an alternative to spectrum scanning. A summary of our contributions and the problems solved is provided along with the thesis organization and our publication at the end of this chapter.

#### **1.1 Background**

The recent increase in wireless mobile devices has led to a surge in growth of wireless communication systems. With even larger wireless usage expected for the foreseeable future, problems begin to arise from the limitation of the wireless spectrum. The unlicensed wireless bands in particular are seeing dramatic usage as new technologies are introduced, including 3G, WiFi, Bluetooth, etc as shown in Fig 1.1. As more devices

using these bands are created, the availability of spectrum bands becomes increasingly limited until there are none remaining.

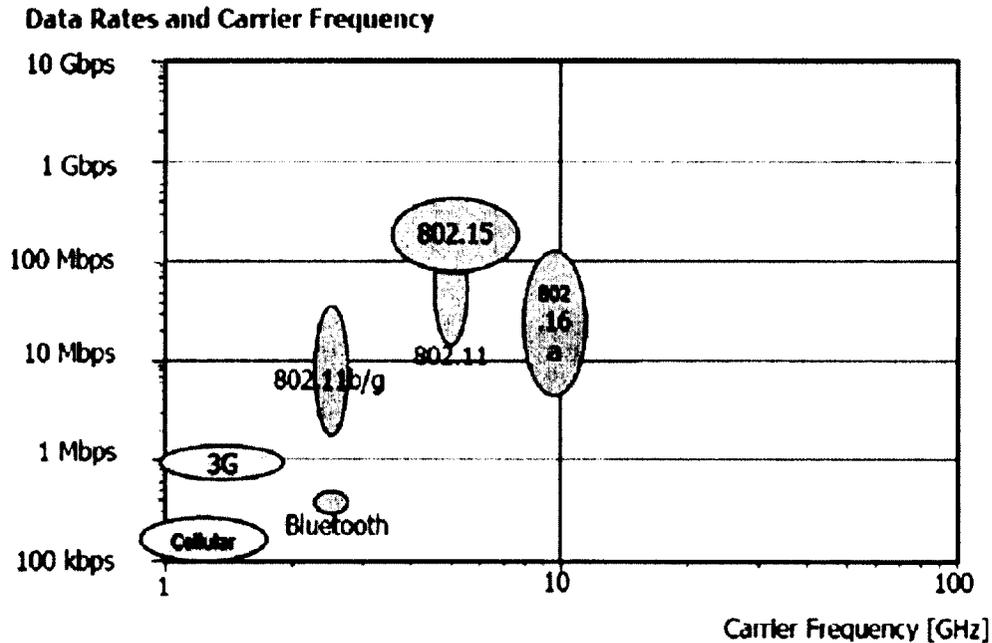


Figure 1.1 Wireless Technology Spectrum Ranges

In contrast, licensed bands have been showed to be largely underutilized compared to their unlicensed counterparts. Licensed bands are parts of the spectrum that have been auctioned off for commercial purposes by a governing body (such as the FCC in the US) to a buyer. Only licensed holders may offer services and build infrastructure for devices operating within the licensed bands. A new trend towards efficient spectrum utilization will be needed for future growth in wireless communications. This requires the use, in a dynamic manner, of these licensed bands without causing interference to the bands' primary users.

A survey in [1] shows several important applications of cognitive radio devices. Cognitive radios and their ability to dynamically access the wireless spectrum has been

considered in changing our current power grid and electrical delivery system to a 'Smart Grid' that incorporates wireless networks for metering, and billing purposes. Public safety networks have also been proposed for emergency services such as fire and police workers to respond to incidents quicker using wireless handheld devices. Other uses explored include the creation of a wireless medical network that would allow hospitals to monitor patient's vital signs using wireless devices, and improved cellular networks that could support more users and wireless devices in an age where users expect to be always connected.

Cognitive radios are a promising way to solve the problem of efficient spectrum usage. Devices using cognitive radio technology require several functions that allow it to use unoccupied parts of the licensed spectrum band. These include: sensing the spectrum for white space (a portion of frequency band that is not currently being used), managing spectrum usage to meet the user's quality of service requirements, and providing a fair spectrum scheduling method to share with other users [2]. Cognitive radio devices may also have spectrum mobility functions that provide the means of exchanging frequencies of operation allowing seamless transition between spectrum bands without degradation to communications when required [3]. These functions allow opportunistic use of these unoccupied channels while ideally avoiding interruption to the licensed users of that part of the spectrum [2].

## 1.2 Research Problems

The spectrum sensing function that cognitive radios require in order to detect other users still has many open issues. Accuracy in detecting signal presence is a major issue with current spectrum sensing devices especially in situations with high noise. Designs of higher quality spectrum sensing devices as well as algorithms for exchanging spectrum sensing data between cooperative users have been a major focus of cognitive radio improvement [4]. The current trend calls for designing more sophisticated spectrum sensing algorithms as well as increased sharing of spectrum data between cooperative users to produce a more accurate picture of the surrounding spectrum bands [5][6].

Another problem comes from spectrum scanning. Spectrum scanning is a required part of the sensing function. It can be highly costly in terms of both time and energy. Depending on the range of frequencies that is needed to be scanned, a large amount of time may be consumed scanning before users can begin use [7]. This is a major issue facing cognitive radios and efficient spectrum utilization. In cases with highly dynamically changing spectrums, spectrum sensing data from scanning may be outdated very quickly. This could lead to situations in which certain white spaces become unavailable after detection. Periodic or even constant spectrum scanning would be needed in order to provide an accurate and up-to-date picture of the spectrum [8]. This would not only increase turnaround times, but also would be a constant drain on the devices' battery power.

Even when usage is detected, it is difficult to classify different users [9]. In the case of cooperative sensing where information from multiple cognitive radio users are

incorporated for primary user detection, the issue remains in how devices using different wireless communication technologies, i.e. WiMax, WiFi, etc, are able to communicate with each other. Assuming that they are able to do this, users of different service providers are not presented with any incentive to help each other. However, cooperative approaches introduces additional overhead and complexities into cognitive radio devices including the creation of a wireless network for devices, synchronization of nodes, choosing the best nodes for spectrum sensing, etc [10]. More information on spectrum sensing is provided in later chapters.

The allocation of channels is difficult to achieve efficiently when the environment is highly dynamic. Spectrum conditions may change drastically over a short period of time [8]. Devices need to adapt to different channel conditions. This requires the optimal allocation of channels given the users service requirements while avoiding collisions with the primary users. A collision caused by failure of detection is costly in terms of spectrum usage. Open issues include providing fair scheduling methods to allocate channels between secondary users in a dynamic environment, as well as collision avoidance with the primary users.

The return of primary users to take over channels used by secondary users is also an issue that needs to be addressed. Detection of signal presence on secondary user's channels may indicate a returning primary user. However, it is difficult to classify the signal as either a primary user or some other unknown source [9]. In addition, there are spectrum mobility functions to consider in order to provide uninterrupted service when this occurs.

In many cases, deployment of cognitive radio technology is hampered due to costs associated with the infrastructure. Cognitive radio devices require an underlying infrastructure in place in order to work. However, primary users may not allow them to use their own. A separate network needs to be built for them. Given the cost to implement a separate network solely for cognitive radio use may be considered unfeasible to develop.

### **1.3 Objectives**

The main objectives of the thesis are to address the problems in the last section. These include the discussion on spectrum sensing and scanning, exploration of application layer solutions for channel allocation, and development of an overlay to existing network to support the application layer approach.

- In this thesis, we plan to replace the spectrum sensing and scanning parts of traditional cognitive radio with a system that provides the same functionality under the assumption that primary users will share usage information with the secondary user. This solves many problems that are associated with the use of spectrum sensing and scanning that are explained in more detail in the next chapter.
- This thesis also explores the use of message signalling at the application layer in order to share spectrum usage information between users, as well as to allocate, and manage channels in the licensed band. The use of application layer message

signalling can let us look at the overhead due to delays they create and how they ultimately affect the efficiency of channel utilization.

- By providing a system model of our approach, we can share a view of our implementation at the application layer involving an overlay to existing networks. This can demonstrate the feasibility of our method as well as an alternative implementation of cognitive radios.

## **1.4 Our Approach**

The problems facing cognitive radio require additional development in order to create a viable technology capable of realising the potential of efficient spectrum usage. Issues facing spectrum sensing, channel allocation and network implementation are all problems we wish to solve.

Our approach to cognitive radios involves the use of a centralized server to help allocate channels within the licensed band to licensed users and opportunistic users equipped with cognitive radios. We use an application layer solution to communicate between cognitive radio users and this allocation server. An application layer solution means we are using an approach to cognitive radios that involves the use of an application layer protocol. Specifically, we are referring to the use of Session Initiation Protocol (SIP). SIP is an application layer protocol meant for real-time media access. This is not a session layer protocol.

This method is meant to forgo the use of spectrum sensing required by traditional cognitive radios. The physical aspect of cognitive radio is instead replaced by the application layer to perform the function of channel allocation and management. In this way, we can eliminate the need for spectrum sensing and scanning and move the responsibilities away from the physical layer towards the application layer. A more detailed explanation is presented in Chapter 4.

Opportunistic users equipped with cognitive radio as well as licensed users provide their usage information to this server before allocation decisions are made and permission granted to cognitive radio users. This approach has the benefit of avoiding costly spectrum scanning as well as collisions that may occur between both types of users. In addition, it solves the inherent problem in cooperative sensing mechanisms where users of different wireless technologies are not capable of communicating with each other. By using an application layer protocol, we can assume devices can communicate with this allocation server as long as they are Internet Protocol (IP) based. Also, since we use the application layer in our communications we can use an overlay of an existing IP-based communication infrastructure. This gives a very economic alternative to actually building a separate network infrastructure solely for cognitive radio users or resort to increasingly complex devices to handle cooperative methods. In this system, the allocation server remains in control of the spectrum and can easily track the spectrum usage among all types of users.

Centralized approach has been explored before in works such as [11]. The approach of using a messaging protocol to provide information from primary users has not been explored before. By providing information on their spectrum usage to the allocation

server, the problem of primary user detection is solved. However, primary users must gain something in order to share their usage information. Otherwise there is no incentive to help secondary users. We use an auction mechanism to provide primary users with some monetary compensation for their usage information. This auctioning mechanism will be used to allocate channels among winning secondary users. This technique not only provides us with an allocation mechanism, but also provides incentive for primary users to share information with secondary users.

## **1.5 Contributions**

The focus of our contributions lies with the implementation of cognitive radios. Our method provides solutions to problems associated with spectrum sensing, network implementation, and communication between different wireless technologies.

This thesis provides the following major contributions:

- Specifications pertaining to a system model based on primary users' willingness to share their spectrum usage information to a centralized server in order to manage spectrum allocation among cognitive radio users
- Communication between the server and users using application layer protocol. Our method employs Session Initiation Protocol (SIP) and two extensions to adapt the SIP protocol to meet our needs.

- Adoption of an auctioning mechanism in order to distribute spectrum among cognitive radio users as well as provide primary users with incentive to share their usage information.
- Performance analysis, both analytical and through simulation, of our proposed solution in order to validate the effects of delay as well as spectrum efficiency.

These contributions provide benefits for cognitive radio devices in the following forms:

- Eliminate the need for spectrum sensing and in turn remove the possibility of false or missed detection of primary users.
- Optimization of channel allocation centralized to remove the need for more complex computations and devices associated with cooperative sensing.
- Overlay of an existing mobile network. This communication using SIP can be accomplished by using existing mobile access networks that support IP such as WiMAX or Third Generation Partnership Project (3GPP)'s Long Term Evolution (LTE). This eliminates the need to deploy a different network infrastructure solely for cognitive radio use. We merely need a server that provides spectrum management and allocation capabilities.

## **1.6 Thesis Organization**

The remaining chapters are organized as follows:

**Chapter 2:** Literature survey gives a state of the art overview of current literature that is relevant to this topic. This includes background information on cognitive radio

technology, trends in development, and other approaches to solve challenges facing cognitive radios.

**Chapter 3:** Our system model will be presented here. This includes the entities that will be the core of our system, a description of the SIP protocol along with its extensions will, the messaging model and the SIP messages that will be used. A messaging model will show how a channel allocation protocol can be developed using SIP and how this model can be used in an overlay of an existing network infrastructure. We also provide an auctioning mechanism that can distribute channels among competing users.

**Chapter 4:** Here we derive an analytical model for the delay we expect from our system solution presented in the previous chapter. This section will also describe sessions and protocol management. It will provide a detailed explanation of the various scenarios that will take place and the amount of delay incurred in each case according to our analysis. It also shows analysis of the efficiency given the previous derivations on mean delay. Simulation results are also presented in this chapter. A detailed look at the network model used for the simulation and procedure, assumptions used, as well as analytical model will be given before the results of the simulation. Simulation results show overhead incurred due to messaging delays, channel utilization percentage relative to user behaviour changes as well as auction results.

**Chapter 5:** A conclusion will be presented along with suggestions for future research.

## **1.7 Publications**

J. Long, C. Huang, J. Yan, "Network assisted auctioning for cognitive radios," *2010 IEEE 3rd International Workshop on Selected Topics in Wireless and Mobile computing (STWiMob 2010)* , vol., no., pp.669-674, 11-13 Oct. 2010

J. Long, C. Huang, J. Yan, "Application-Layer Approach to Cognitive Radio," *2012 IEEE International Conference on Communications (Next Generation Networks)*, pending review

## **Chapter 2**

# **Literature Survey**

This part of the thesis surveys recent advances in research related to cognitive radios. The architecture of traditional cognitive radio network and its applications are first introduced. The existing works in spectrum sensing are reviewed, and important issues in dynamic spectrum allocation and sharing are investigated in detail. A section will be dedicated to the lessons learned from these works and an ideal solution to cognitive radio implementation.

### **2.1 Fundamentals**

Radios were once devices that could tune to specific frequencies to send and receive wireless information; such as a radio station for music. However, if the channels conditions were poor there was no other way to avoid the interference other than to wait it out. Advances within software have made radio devices smarter. Now they may have the capability to intelligently detect which of the channels are in use and which are not. This would allow them to use the vacant channels while avoiding those that are already in use. Not only does this optimize the use of the spectrum bandwidth, but it also minimizes the interference to other users.

Cognitive radios may be defined as any wireless communication devices that opportunistically uses the white space of the spectrum and includes technology that prevent interference to other users such as spectrum sensing. In order to perform these functions, the cognitive radio device has several duties that are needed in order to detect the unused parts of spectrum, use it, and avoid interference with other users.

A typical duty cycle of cognitive radio includes detecting spectrum white space, selecting the best frequency bands, coordinating spectrum access with other users and vacating the frequency when a primary user appears [12]. Such a cycle is supported by the following functions:

- Spectrum sensing
- Spectrum management
- Spectrum allocation

Through the spectrum sensing function, cognitive radios can detect occupied channels from the spectrum white space. After recognizing the spectrum white space by sensing, spectrum management function of cognitive radio enables secondary users to choose the best frequency band to meet their requirements while still providing fairness to other users.

## **2.2 Cognitive Radio Classifications**

With the development of cognitive radio technologies, secondary users who are not authorized with spectrum usage rights can utilize unused licensed bands owned by the primary users. Therefore, in traditional cognitive radio network architecture, the components include both a secondary network and a primary network [13]. Both secondary users and secondary base stations are equipped with cognitive radio functions. If several secondary networks share one common spectrum band, their spectrum usage may be coordinated by a central network entity, called a spectrum broker. The spectrum broker will be responsible for collecting information from secondary users and allocate channels for efficiency and fairness [14].

Primary users access the spectrum using their own networks. This consists of user equipment as well as one or more base stations. They are not assumed to possess cognitive radio functionality. When primary users and secondary users share the same spectrum bands, the responsibility falls on the secondary users and the secondary network to do the detection and ensure that primary users are not interfered with. Therefore, the secondary network must both detect spectrum white space as well as detect the presence of primary users.

In a hierarchical access model, secondary users and primary users share and co-exist within the same licensed spectrum bands. Secondary user access technology within this model can be divided into two groups: spectrum underlay and spectrum overlay [15].

In spectrum underlay secondary users are allowed to transmit their data in the licensed spectrum band when primary users are also transmitting. The interference temperature

model is imposed on secondary users' transmission power so that the interference at a primary user's receiver is within the interference temperature limit and primary users can deliver their packet to the receiver successfully [15].

Unlike spectrum underlay, secondary users in spectrum overlay will only use the licensed spectrum when primary users are not transmitting. There is no interference temperature limit imposed on secondary users' transmission. Secondary users are required to sense the licensed frequency band and detect the spectrum white space, in order to avoid interference to primary users [15].

Another classification is based on the network architecture. If there is a central entity that controls and coordinates the spectrum allocation and access of secondary, such as a spectrum broker, then the spectrum allocation is centralized. If there is no such central controller, it is distributed [16] [17]. In distributed spectrum sharing, each user makes their decision about spectrum access strategy, based on local observation of the spectrum band.

A third way of classifying cognitive radio systems is according to the access behavior of secondary users. They can be either cooperative or non-cooperative. In cooperative spectrum sharing, users coordinate their allocation for the benefits of everyone. This is usually the case when users are of the same service provider [18]. A non-cooperative approach assumes users do not cooperate to achieve the same objective. They are assumed to be maximizing their own benefits, even to the detriment of everyone else.

## 2.3 Spectrum Sensing

Spectrum sensing allows cognitive radios to examine the surrounding radio environment for signs of primary users as well as spectrum white space. With this information, a cognitive radio user can adapt itself to specific parameters such as frequency, transmission power, etc., to utilize the available spectrum dynamically without causing interference to primary users. Spectrum sensing is the first step in ensuring that dynamic spectrum management is possible. There are three main techniques employed by traditional cognitive radio technologies to provide spectrum sensing.

- Energy Detection: Used to measure interference levels at receivers to detect primary users
- Feature Detection: For sensing certain features commonly associated with users (mostly cyclostationary).
- Matched Filter Detection: Match a predefined template of user signal with an unknown one.

These techniques will be discussed briefly and a summary of their advantages and disadvantages will be provided. Spectrum sensing can be performed in the time, and frequency domains. If primary users do not broadcast in all directions, secondary users may utilize the same frequency in an area that the primary user does not. They may also transmit on the same band if they use a code orthogonal code as primary users. This opens up more opportunities in the spatial domain as well as the code domain [19] [20].

### **2.3.1 Energy Detection**

The most common form of spectrum sensing is energy detection. Energy detection is an attractive spectrum sensing method in practice, primarily due to the facts that it has low implementation complexity and it does not depend on any deterministic knowledge about primary signals. A good energy detector needs to have a high detection probability so that it does not produce false detections that could lead to less efficiency as well as providing some level of primary user protection.

Several approaches used to improve energy detection based spectrum sensing include the following.

In [21], methods for energy detection are proposed, where the noise power is unknown and is adaptively estimated. [22] Introduces a wideband spectrum sensing technique based on energy detection, which mutually detects the signal energy levels across multiple frequency bands in order to improve the throughput of cognitive radios and reduce their interference to primary users.

Although energy detection has benefits for cognitive radio design, there remain open issues regarding their implementation. Energy detection relies on two critical assumptions, 1) noise power is perfect and known beforehand and 2) test statistics can be accurately modeled as independent and identically distributed Gaussian random variables [22]. However, in practice, noise power can vary considerably as time progresses in a dynamic environment. In addition, the noise power threshold for determining primary users must be known prior. This may not always be true. This threshold is used to compare received signal energy in order to determine a primary user presence. However,

detection of other unknown signals in the area that meet that threshold may trigger false detections that may or may not be a primary user. Since the detection threshold depends on noise power, detection may become unreliable in some situations, such as when noise power in the area is very high.

### **2.3.2 Feature Detection**

Features refer to any inherent characteristics related with a primary user's transmission. Statistics of transmitted signals in many communication paradigms are periodic because of the inherent periodicities such as the modulation rate, carrier frequency, and others. These features are often associated with the information transmission of a primary user and are usually viewed as cyclostationary features. Most communication systems have signals that are modulated with sine wave carriers, while white space is wide-sense stationary. Since there is no correlation between the two, a cyclostationary detector can be used to determine spectrum white space from user presence. These features may then be classified and used to determine whether a signal is in fact the presence of a primary user.

Compared to energy detectors that are prone to high false alarm probability due to noise uncertainty and cannot detect weak signals in noise, cyclostationary detectors become good alternatives because they can differentiate noise from primary users' signal and have better detection robustness in low signal to noise ratio (SNR) situations. [23]

There remain challenges to feature detections that need to be solved before implementation. A high computational complexity is needed for cyclostationary detectors.

This may prove to be unfeasible given the cost that is associated with cognitive radio implementation. In order for cyclostationary detectors to work, they require some partial information from primary users on their cyclic frequencies. Several wireless communication standards have begun using orthogonal frequency-division multiplexing (OFDM) as part of their air interface. The migration towards OFDM poses a problem for cyclostationary detection. Identification of different systems may become challenging, since the features of OFDM signalling are likely to be close or even identical [24].

### **2.3.3 Matched Filtering Detection**

If secondary users know prior information about a primary user's signal, then the optimal detection method is the matched filtering [25]. A matched filter is obtained by correlating a known signal template of a primary user, with an unknown signal to detect the presence of the template in the unknown signal. This is equivalent to convolving the unknown signal with a time-reversed version of the template provided by primary users.

The advantage of matched filtering is the short time it requires to achieve a certain detection performance such as a low probability of missed detection and false alarm [26], since a matched filter needs less received signal samples.

**Table 2.1: Advantages and Disadvantages of Sensing Methods**

	Advantages	Disadvantages
<b>Matched Filter</b>	<ul style="list-style-type: none"> <li>• Low Complexity</li> <li>• Can provide optimal detection</li> </ul>	<ul style="list-style-type: none"> <li>• Requires perfect knowledge of primary user's information</li> </ul>
<b>Feature Detection</b>	<ul style="list-style-type: none"> <li>• Robust to low SNR</li> <li>• Robust to interference</li> </ul>	<ul style="list-style-type: none"> <li>• Computationally complex</li> <li>• Requires some information from primary users</li> </ul>
<b>Matched Filter Detection</b>	<ul style="list-style-type: none"> <li>• Low Complexity</li> <li>• Can provide optimal detection</li> </ul>	<ul style="list-style-type: none"> <li>• Requires perfect knowledge of primary user's information</li> </ul>

There are still many issues facing matched filter implementation. Matched filtering requires perfect knowledge of the primary user's signal. This information must be provided from the primary users. However, the means or the incentive for primary users to share information is only assumed and never presented. The case also exists where if false or erroneous information is used for matched filtering, detection performance will be degraded dramatically. In addition, the required number of signal samples also grows as the received SNR decreases, so there is a SNR wall for a matched filter. Since the matched filter needs receivers for all types of signals and corresponding receiver algorithms to be executed, its implementation complexity and power consumption is also considered to be high [27].

## 2.4 Cooperative Sensing

Cooperative spectrum sensing is proposed to improve the reliability of spectrum sensing, increase the detection probability to better protect a primary user, and reduce false detection to utilize the idle spectrum more efficiently. In centralized cooperative spectrum sensing, a central controller, e.g., a secondary base station, collects local observations from multiple secondary users, decides the available spectrum channels using some decision rule, and informs the secondary users which channels to access. In distributed cooperative spectrum sensing, secondary users exchange their local detection results among themselves without requiring a backbone infrastructure. Relays can also be used in cooperative spectrum sensing, such as the cooperative sensing scheme proposed in [28], where the cognitive users operating in the same band help each other relay information using amplify-and-forward protocol.

Co-operating all secondary users in spectrum sensing is not optimal. Instead, optimum detection is achieved by co-operating a group of users who have higher SNR [29]. Detecting a primary user costs battery power of secondary users. Therefore, an optimal selection of secondary users for cooperative spectrum sensing is more advantageous. Different algorithms based on different amount of available information are proposed to select a proper set of sensors. Secondary users collaboratively estimate the location and transmit power of the primary transmitter to determine their maximum allowable transmission power, and use the location information to decide which users should participate in collaborative sensing in order to minimize correlation among the secondary users.

Once the results of spectrum sensing from selected users have been taken, decisions must be made using all the aggregate data collected to form a clear picture of the spectrum. Different combination rules have been studied such as using logic OR rule, counting rule, voting, etc [30] [31] [32]. Cooperative sensing is coordinated over a separate control channel, so a good cooperation schemes should be able to use a small bandwidth and power for exchanging local detection results while maximizing the detection reliability. Performance of cooperative spectrum sensing depends on the accuracy of the local sensing data reported by the secondary users.

Several challenges face cooperative sensing. Secondary users can be low-cost devices only equipped with a limit amount of power, so they cannot afford very complicated detection hardware and high computational complexity. In wideband cooperative sensing, multiple secondary users have to scan a wide range of spectrum channels and share their detection results. This results in a large amount of sensory data exchange, high energy consumption, and wasteful data throughput. If the spectrum environment is highly dynamic, the sensed information may even be stale due to user mobility.

Additionally, in order to coordinate the cooperation in spectrum sensing, a lot of information exchange is needed among secondary users, such as their locations, estimation of the primary user's location and power, which users should be clustered into a group and which users should perform cooperative sensing at a particular time [33]. Such a large amount of information exchange brings a lot of overhead to the secondary user. If cognitive users belong to different service providers, they tend to contribute less in sensing in order to increase their own data throughput.

## 2.5 Spectrum Scanning Techniques

Wideband spectrum scanning with multiple channels has been recently studied in order to produce the best and maximum throughput performance. Scanning has been proposed as both sequential and multi-channel joint detection which uses a larger number of detectors that operate simultaneously. However, it is often difficult for cognitive radio devices to be equipped with such a large number of detectors and operate them simultaneously [34]. Using a single detector has two scenarios of note. When the detector is given infinite time to scan the spectrum, or when given a strict delay constraint on how long it may scan.

Scanning takes the following sequence of steps to determine channel usage: 1) scan current channel 2) decide if channel is free 3) scan current channel again and go back to step two 4) move to new channel to scan.

Several scanning methods have been proposed in [35] as alternatives to unidirectional sequential spectrum scanning. Bidirectional scanning scans in both directions on the initial channel to find the closest available channel whereas a sequential scan may miss a free channel close to initial scanned channel that is scanning in the opposite direction. If starting frequency is  $f_x$ , then the scanning sequence would be  $f_x, f_{x+1}, f_{x-1}, f_{x+2}$ , etc. A dual scanning method initiates scans in two different parts of the spectrum that may be randomly selected and then scanned sequentially alternating between each other. If starting frequencies are  $f_x$  and  $f_y$ , then the scanning sequence would be  $f_x, f_y, f_{x+1}, f_{y+1}, f_{x+2}$ , etc.

There are several factors that contribute to the overall scanning delay times of cognitive radio systems. In practice, cognitive radios may choose to set a maximum allowable scanning time in order to meet requirements. By setting an upper bound on scanning times, users can experience more consistency at the cost of detection reliability. When switching to another channel to scan, tuning times factor into the overall delay. In wideband scenarios with many channels, tuning times can play a significant part in delay.

Scanning times are also influenced by the sensing method employed (e.g. energy detection has less overhead delay than cyclostationary detection). The sensing methods may be classified by their average detection times. A fast sense is usually employed using energy detection that has average detection delays of around 1 millisecond. Fine sensing that employs detection methods with costlier delays (such as cyclostationary detection) may run scanning delays that average 25 milliseconds.

Noise can lead to false or missed detections. When SNR is low, the probability of these cases occurring increases. Scanning times are usually adjusted to compensate for this by increasing scanning times in order to meet their minimum detection probability. Studies have shown [36] that depending on channel conditions, some sensing methods are better suited than others. Very high SNR scenarios make energy detection very attractive since detection probabilities are high. In low SNR scenarios, energy detection becomes infeasible because detection probabilities are too low. It is recommended to use more robust sensing methods with longer scanning times such as cyclostationary detection in these cases.

The use of cooperative sensing has shown the ability to reduce scanning times among cooperating users while maintaining high detection probabilities. [37] shows the scanning times of local secondary users using energy detection that meets an overall detection sensitivity of 99 percent given the amount of cooperating users. These scanning times range from high in the case of only one user, 100 milliseconds, to low in the case of ten users, 0.1 milliseconds.

In all these cases, a trade off exists between scanning delays and probability that the cognitive radio makes a false or missed detection of primary user. Errors probabilities can be made very small given infinite scanning time. However this incurs significant delay that is unfeasible. The opposite is true for very short scanning times may not offer appropriate protection of primary users or could also lead to inefficiencies in channel usage due to false detection case.

## **2.6 Dynamic Spectrum Allocation**

The availability and quality of a spectrum band may change rapidly with time due to primary users' activity and competition from other secondary users. In order to utilize the spectrum resources efficiently, secondary users need to be able to address issues such as when and how to use a spectrum band, how to co-exist with primary users and other secondary users, and which spectrum band they should sense and access if the current one in use is not available.

The medium access control (MAC) provides the policy on which secondary users should access the spectrum. By using the features of cognitive radios, such as spectrum sensing,

new MAC protocols have been proposed to offer spectrum access that is both fair and efficient. MAC protocols using stochastic modeling have been proposed as a means to share spectrum bandwidth between cognitive radio and WLAN users in [38] based on spectrum sensing. A cognitive MAC is proposed in [39], which takes into consideration hardware constraints and optimizes the spectrum sensing decision by formulating sensing as an optimization problem.

The research challenges and issues still facing MAC protocols for cognitive radios include:

Traditional cognitive radio networks frequently use a common control channel between users and base station. The event may occur when control channel saturation occurs and a user must wait before it can successfully send its Request to Send (RTS) frame. This can lead to significant and unnecessary access delay when the data channel itself is being underutilized.

Coordination among distributed cognitive radio networks is difficult. Synchronization of nodes within the network must take place for accurate sensing data to be aggregated among users.

Information dissemination and the overhead it creates may interfere with other users who are using channels. Managing the transfer of information between nodes while avoiding interference with users and providing low overhead is difficult.

In cooperative cases, the cooperation overhead generally increases with the number of cooperating users due to the increased volume of data that needs to be reported and processed. There exists a tradeoff between local processing overhead and cooperation

overhead as they both add to the total delay. This tradeoff may be balanced by finding optimum levels of processing and cooperation in order to minimize the overall delay.

The underlying protocol used for cooperation has an underlying efficiency that must also be considered. Cooperation may be conducted by polling each user for their sensing data. However, the overhead associated with this method increases linearly with the number of users. Other protocols have been created that allow for more efficiency at the cost of protocol complexity.

Optimizing sensing duration and allocation is also a question that must meet both high data throughput, meet quality of service (QoS) requirements of users and provide protection of primary users. This balance is an issue that also must be addressed.

## **2.7 Spectrum Mobility**

When channel conditions worsen, or primary users reclaim occupied channels, secondary users must stop data transmission and find other available channels before resuming. This is called spectrum handoff in cognitive radio networks. Each time a user changes its frequency of operation, the network protocols may require modifications to the operation parameters. Spectrum mobility management in cognitive radio networks ensures a smooth and fast transition with minimum performance degradation during a spectrum handoff. A good spectrum handoff mechanism should provide secondary users with frequency transition that has the least latency. One suggestion to lighten the performance degradation from lengthy delays is to reserve some number of channels for spectrum handoff [40]. When secondary users need to switch frequencies, they can instantly pick a

channel from the reserved bands. Problems arise when secondary user reserve too much bandwidth for spectrum handoff. The throughput may be unnecessarily low, because the primary user may not reclaims their licensed band regularly. Therefore, there is a tradeoff in optimizing the channel reservation.

Another proposed method of avoiding interference with primary users is controlling the transmission power. The impact of transmission power on spectrum opportunities is investigated in [41], and it is shown that the optimal transmission power of secondary users decreases with the traffic load of the primary network. An opportunistic power control strategy is proposed in [42], which enables the cognitive user to maximize its transmission rate while guaranteeing that the outage probability of the primary user is not degraded.

Cognitive radio networks adapt to the wireless spectrum based on the available bands. Since channel availability changes over time, providing a level of QoS in this environment is difficult. The available bands also changes as users move from one place to another. Hence, continuous allocation of spectrum is a major challenge.

## **2.8 Similar Works**

In [43], Xu, et al discuss reconfigurable wireless networks that use cognitive radio technology. In this paper, a network architecture based around 3GPP's IP Multimedia Subsystem (IMS) and WiMAX is used as a basis for spectrum management and allocation between cognitive radio devices.

This paper shares similarities with the concepts in this thesis. They propose the use of a signalling protocol at the application layer to exchange information between users which is similar to our contribution of using SIP. In the paper, they propose the use of Simple Network Management Protocol (SNMP) to contact any radio device using IP to inform each other of the spectrum environment and usage conditions.

Network architecture types are discussed in more general terms. A centralized server model is discussed where a central management server located within the IMS has a shared database that cognitive radio users may access.

Spectrum sensing still occurs by a spectrum management entity that continuously scans the radio environment and updates this information to cognitive radio users. They do not consider the primary users to be part of this system and do not ask for their spectrum usage information.

This still leads to issues in spectrum sensing that were discussed previously. However, by shifting the spectrum sensing part to the spectrum management entity, they remove the responsibility from the user which may greatly reduce the complexity of the user devices. This thesis takes the idea one step further by introducing the idea of spectrum information sharing across all user types and thus eliminating the need for spectrum sensing.

## **2.9 Implementation Challenges**

This section will provide a summary of the challenges that still face cognitive radios and their implementation and an ideal solution that is strived for. In the next chapter, we can

then show how our approach aspires to this ideal solution and how they answer some of the questions that other methods do not.

There exists three current methods to provide spectrum sensing features of cognitive radios 1) Energy detection 2) Feature detection and 3) Matched Filter detection. All three methods have their advantages and disadvantages. For energy detection, they have low complexity, but are not well suited for use in low SNR scenarios. Feature detection of cyclostationary signals provides a more robust detection mechanism in this case, but overly complex hardware and computation may lead to additional overhead. Matched filter detection provides the most optimal way to detect signals, but requires that primary users provide information about their signal before it can work. In addition to the problems associated with these sensing techniques, spectrum sensing in general incur an overhead due to the turnaround time associated with scanning. Certain scenarios such as wideband lead to large scanning times which directly affect both spectrum efficiency as well as power consumption.

Cooperative sensing attempts to solve some of the major issues facing spectrum sensing, but introduces its own problems. With the use of distributed network to disseminate information, large amount of sensory data must be exchanged between nodes. This leads to high energy consumption and wasteful data throughput with the amount of overhead that it generates.

A trade off exists between scanning times and detection reliability. Shorter scanning times can lead to better utilization, but the chance of false or missed detection increases. This may not be enough protection for primary users. The reverse is true for longer

scanning times which leads to under utilization for secondary users. Scanning methods need to achieve more reliable results with lower scanning times in order to fully realize spectrum utilization. Advancements to scanning times must also be made in the case of wideband scenarios with many channels.

New MAC protocols need to be developed to make essential decisions regarding optimization of sensing and allocation times to meet efficiency, how to select cognitive radio users for scanning duties, how to disseminate data without interfering other users, provide primary user protection, meet user QoS requirements, and many more. Achieving these purposes in a highly dynamic spectrum scenario could prove to be a very difficult task. The overhead in terms of delay caused by saturation of control channels have yet to be answered.

Handoff scenarios also face issues involving providing seamless transition to other available channels when pre-empted by primary users. Traditional methods must have pre-empted users conduct spectrum sensing which incurs scanning delays before finding another available channel. Using reserved channels for the handoff scenario means less overall channels available for users.

## **Chapter 3**

# **Application-Layer System Solution**

The proposed system model will be based around three network entities and the exchange of information between them in order to allocate channels and provide control over session management. This chapter will introduce three core network entities of our system model, their integration in an overall network architecture based on LTE, an introduction to SIP protocol and the extensions used, and a messaging model including auctioning mechanism will be created.

### **3.1 Solution Architecture**

In our method, we propose that primary users of the licensed band share their spectrum usage information with secondary users. This coordination between the two user types in the licensed band can yield all the information necessary in providing an accurate picture of the spectrum assuming there are no other unauthorized users. When both primary and secondary users collaborate by sharing spectrum usage information, the need for spectrum sensing can be removed. This would solve the various problems associated with spectrum sensing.

### 3.1.1 System Entities and Overall Network Overlay

An application server can be used to aggregate the information provided by primary users. This server can be located somewhere close to the users themselves, but can also be flexible enough to be placed anywhere in the internet. By providing channel usage information to this server from primary users and channel usage requests from secondary users, enough information is available to provide dynamic spectrum access to secondary users without the need for spectrum sensing. Other issues with spectrum sensing including robustness in low SNR, complexity, overhead, and device power are all removed.

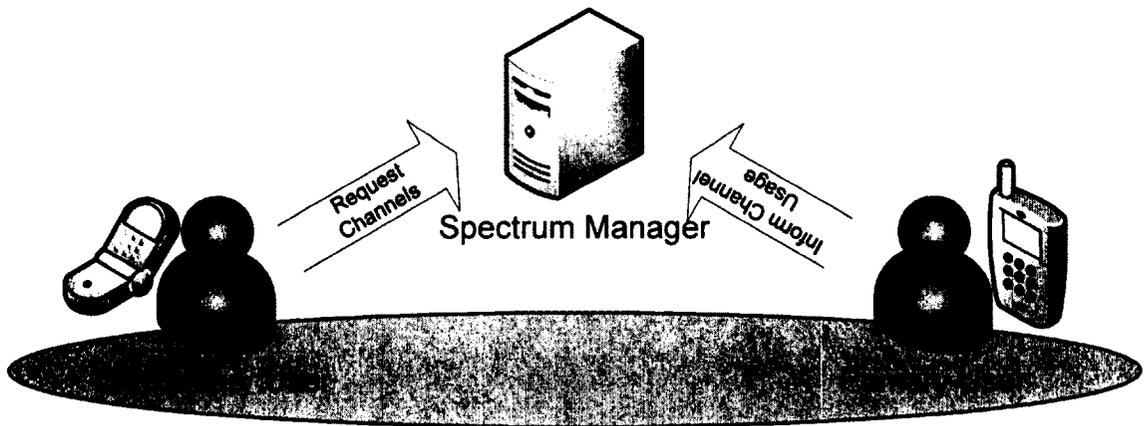
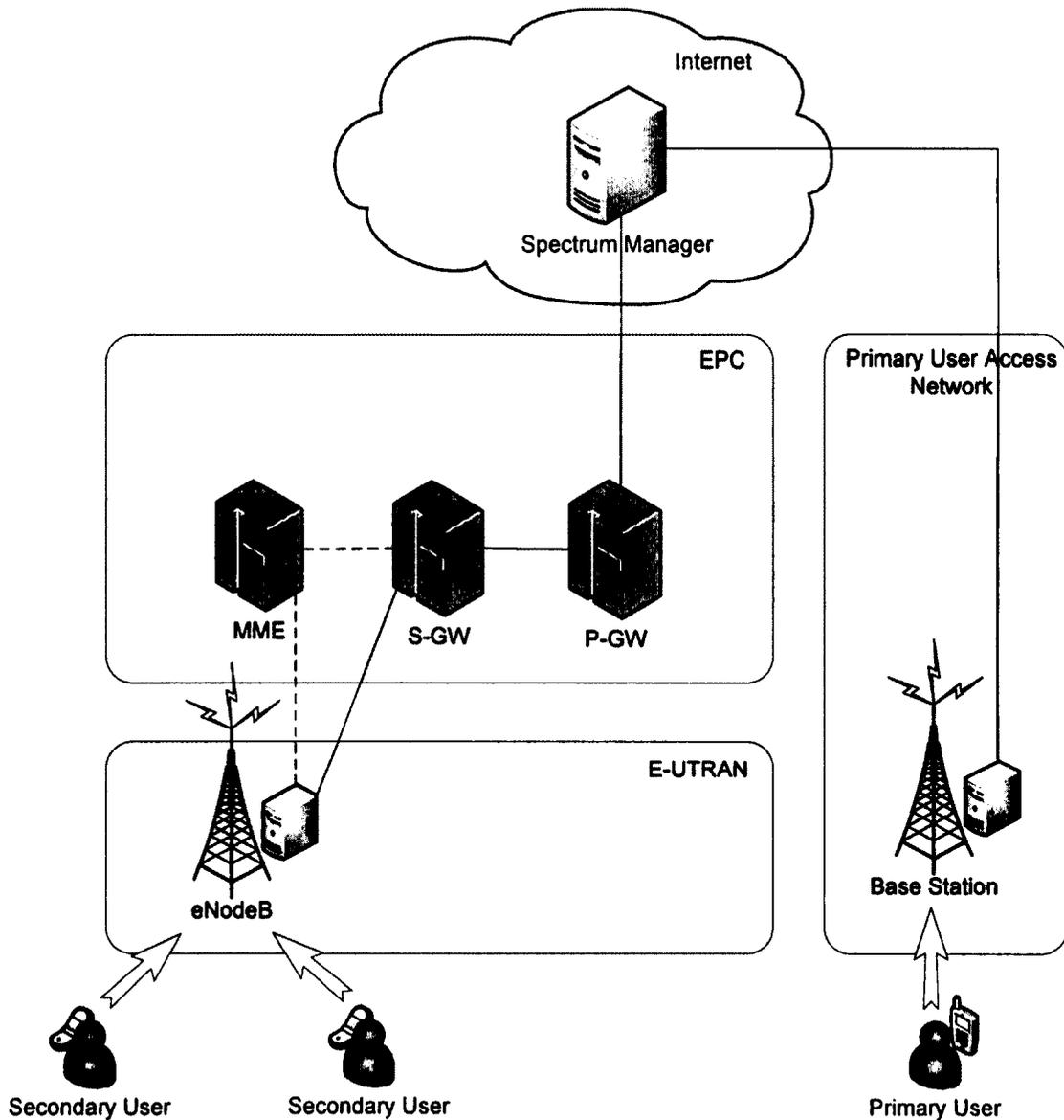


Figure 3.1: Relationship between Entities

In our model, three unique entities will be part of the system. A spectrum manager will maintain information about spectrum usage in the area as well as coordinate channel usage. Primary users are defined as licensed users who have priority to use channels. Secondary users are opportunistic users equipped with cognitive radios. The secondary users will first contact the server to secure unused licensed channels. Figure 3.1 shows the relationship between the three entities. When primary users use channels, they will

share their channel usage information with the server. Responsibilities of each entity are defined as follows:

- **Spectrum Manager:** This server will be the main control unit of our three entity model. Its responsibilities include connection establishment/termination, session management between the two other types of user devices, channel allocation, and billing information.
- **Primary Users:** We assume that these users have their own means of communicating with the spectrum manager. A main goal of this system model is to make it as transparent as possible to primary users. In order to accomplish this, a minimal amount of work should be put onto primary users. The main responsibility that they will have to do is to provide the spectrum manager with their channel usage information.
- **Secondary Users:** These are the opportunistic users that are equipped with cognitive radios who wish to use part of the spectrum. Secondary users will have to make a connection with the spectrum manager first and wait until granted permission. Responsibilities include providing information to the spectrum manager on what channels they want to use, amount they are willing to pay, and numerous other information deemed necessary. They will also provide occasional updates to the spectrum manager with information pertaining to channel usage status, requests to change channels, and channel termination.



**Figure 3.2: Overall Network over LTE Architecture**

Figure 3.2 shows the overall network that uses LTE architecture as an example of how our system can be used as an overlay to existing networks. Primary users are assumed to have access to the internet through their own access networks. Secondary users are equipped with cognitive radios. However, primary users are not necessarily equipped with cognitive radios. Both types of user devices will communicate with their

corresponding base station located locally. Base stations will route messages from the user devices to the spectrum manager that administers the usage of channels for the respective areas of each base station. The LTE is a fully packet-switched domain that can provide connectivity to the Internet. The spectrum manager can be placed anywhere in the Internet and still be reachable. LTE network entities shown in the above figure include:

**eNodeB:** Evolved Node B (eNodeB) serve as base stations for secondary users in their respective areas. The eNodeB is the end point for all radio related protocols. It lies in the air interface portion of LTE's infrastructure known as Evolved Universal Terrestrial Radio Access Network (E-UTRAN). The E-UTRAN is a mesh of eNodeBs connected to nearby eNodeBs.

**MME:** the Mobility Management Entity (MME) serves as the control element in the LTE architecture. This can be a secure server located within the Evolved Packet Core Network (EPC). Some of its main functions include: authentication and security, tracking locations of users, and managing user profiles.

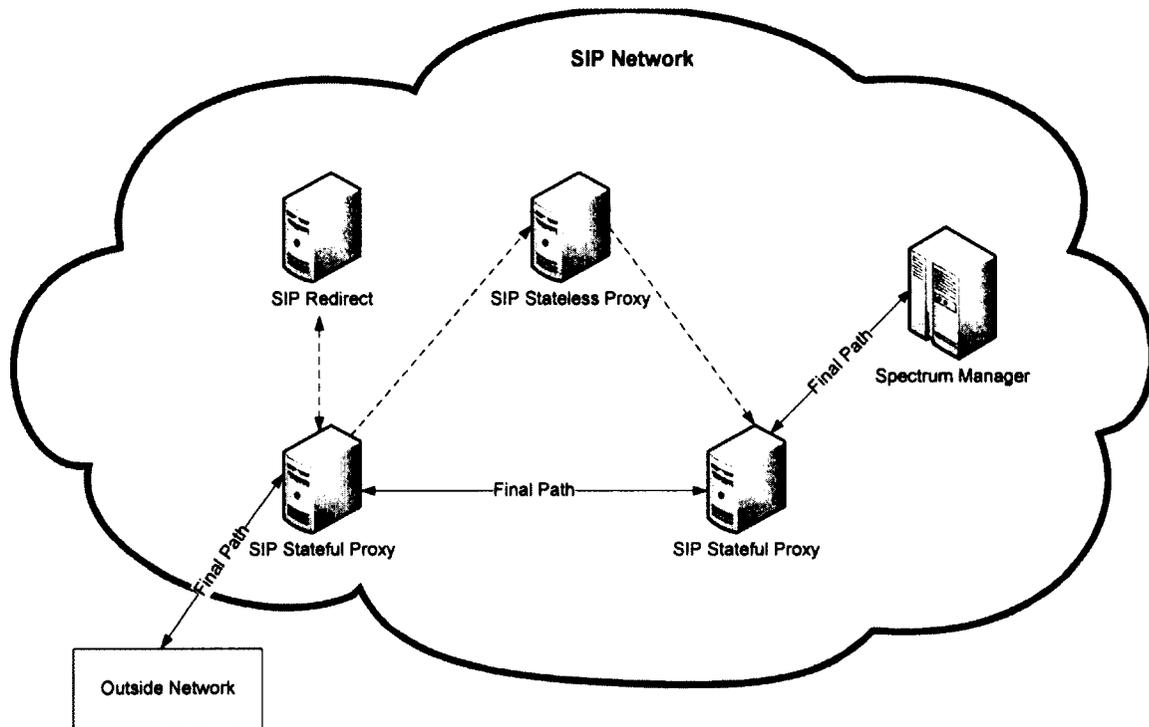
**S-GW:** Serving Gateways (S-GW) is mainly used for IP tunnelling and switching it has very little role in any control functions.

**P-GW:** the Packet Data Network Gateway (P-GW) is the edge router between the LTE network and external networks such as the internet. This usually acts as the IP point of attachment for users. The P-GW typically allocates IP address to users which then use it to communicate with other IP hosts in external networks.

3GPP's LTE wireless communication standard [52] will be used in this thesis to demonstrate how our system model can be used as an overlay to an existing network architecture. Other wireless technologies may also be used to facilitate communication with the spectrum manager.

E-UTRAN allows for high data rate, low latency radio access for wireless device with eNodeBs serving as base stations for secondary users in their respective areas.

The S-GWs are multi-standard access points to the network. The S-GW and the MME together connect users to P-GW which provides connectivity to external networks such as the Internet. System Architecture Evolution (SAE) bearers, IP tunnelling, and IP connectivity from secondary users to the PDN gateway are taken care of by these two entities as shown in Fig.3.4 to provide end-to-end service to the internet. The PDN gateway operates a routing table that will allow IP routing via the Internet. This gateway will route IP packets to the correct eNodeBs that serve the respective user. The user can transmit IP packets to the eNodeB which will route them to the PDN gateway and then to the Internet



**Figure 3.3: SIP Network**

Figure 3.3 shows an example of what the SIP network may look like. The network entities on this side include the following:

**SIP Proxy:** Proxy servers are entities between the user and spectrum manager that act as both server and client for the purposes of making requests for other entities such as users or another proxy. The proxy's primary function is to route messages to another entity closer to the destination. Stateful proxies stay within the final transmission path for message signalling. Stateless proxies do not remain within the final path. They simply forward messages they receive

**SIP Redirect:** This entity generates REDIRECT response messages to requests. This redirect messages tells preceding entity to try a different set of URIs.

**Spectrum Manager:** The spectrum manager is the destination server in our model. This server responds to NOTIFY and PUBLISH messages by users to provide the dynamic spectrum allocation services we want.

The network assisted cognitive radio works over an LTE-SAE architecture in the same way that [53] proposed device-to-device communication as an underlay to LTE networks. After a successful session setup, a secondary user may communicate over the internet using IP.

Once a session is setup, messages can be sent between the server and users for allocation of channels in the licensed bands. Users will keep the session active for the duration of use of the licensed bands in order to keep control and notify terminate of use when finished.

### **3.1.2 Advantages of our Solution**

As described in chapter 2, there remain many challenges facing cognitive radio before they can be fully utilized. The issues facing cognitive radios and their networks include, but are not limited to: spectrum sensing and reliability of sensing data, scanning times and delay overheads, cooperative sensing complexity, the media access control protocol and overhead from data transmission, and handoff penalty to QoS for secondary users.

In our method, we propose that primary users of the licensed band share their spectrum usage information with secondary users. This coordination between the two user types in the licensed band can yield all the information necessary in providing an accurate picture of the spectrum assuming there are no other unauthorized users. When both primary and

secondary users collaborate by sharing spectrum usage information, the need for spectrum sensing can be removed. This would solve the various problems associated with spectrum sensing.

In traditional cooperative approaches, communication between different users is assumed. However, this may not be the case. Users of different wireless technologies or service providers within close proximity to each other do not necessarily have the means of cooperating with each other. Our proposal in using an application layer SIP protocol as the communication protocol between the spectrum server and users solves this dilemma. In this way, there is no need for users to communicate to each other; only to the spectrum manager. As long as both primary and secondary users support IP, they may communicate with the spectrum manager using SIP protocol without the need for hardware changes.

By providing the spectrum manager with spectrum usage information from users, enough information is available to provide dynamic spectrum access to secondary users without the need for spectrum sensing as well as scanning. Since we remove the need for spectrum sensing, other issues related with spectrum sensing including robustness in low SNR, complexity, overhead, and device power are all removed.

Secondary users may simply ask for the spectrum usage information from the spectrum manager to get an accurate up-to-date picture of their immediate spectrum surroundings or be informed when new opportunities arise. This removes the delay associated with scanning times which can be a real problem in wideband scenarios.

A trade off exists between the delays with spectrum scanning and the delays with message transmission. Depending on the amount of time it takes for message transmission versus spectrum scanning times, the results may represent better spectrum utilization.

There already exists a trade off in traditional cognitive radio devices in spectrum scanning between scanning times and reliable detection. Better detection can be made with more time spent scanning. Given infinite time, the probability of false or missed detection can be made negligible. However, the cost incurred by delay prevents this from being feasible.

By allowing primary users to share data with secondary users, we have a complete picture of the surrounding spectrum that is always accurate without the need for spectrum scanning or the costs they demand. The overall trade off between traditional cognitive radios and our method of implementation becomes the amount of delay incurred by messaging delays versus scanning delays and probability of false or missed detections by traditional methods. This is a clear advantage that supports our method.

The concept of cooperative sensing is meant to improve reliability of spectrum sensing and user detection by aggregating results of multiple users into making decisions on channel access.

The advantages of using cooperative sensing include better detection probability and lower scanning times. Users may use a centralized server to aggregate their spectrum sensing results which will make channel allocation decisions for them, or use a distributed network among participating users to share results.

A distributed network still faces many issues in development including disseminating information among users, increased traffic overhead due to large amounts of data transmission, and increased complexity of devices.

A centralized approach using a server to poll results from users has the advantage of not needing overly complex devices and provides a more practical solution in the near term. However, it still requires the need for spectrum sensing from user devices and therefore still suffers from the effects of scanning delays and detection reliability while adding the need for information dissemination among users.

Our approach is similar to the central server method in that it is used as a means of collecting information from its users, but since we assume that primary users share their data with this server with their own access networks, there is no need for the sensing part. Thus we eliminate the overhead of scanning delay and data transmission while providing a clear picture of the licensed spectrum without concern with detection reliability.

This method in comparison with centralized cooperative sensing requires more responsibility on the part of the primary user. However, benefits gained from their help greatly outweigh the task they are required to perform.

The spectrum is highly dynamic and requires very sensitive hardware to not only have high detection reliability, but also to provide this detection information in a timely manner so that it is up-to-date as possible and reduce down times of spectrum bands. The MAC provides the rules for channel access among secondary users. It is imperative that devices understand and operate under the same MAC in order to ensure a working environment. Development in this area is high. There are many new MAC protocols

being proposed that seek to solve problems related to dynamic channel allocation among a group of users.

MAC protocols themselves are increasingly complicated as they attempt to set rules for cognitive radio use. These rules may pertain to channel allocation, spectrum scanning times, information dissemination and combination. They are important to ensure fairness among users, primary user protection, handling overhead, data transmission among users, while at the same time ensuring spectrum efficiency all in a dynamic environment. This lead to increased complexity also leads to unwanted overhead delay.

Our method of using SIP application layer protocol has a much simpler design while still providing the necessary services that are required in our own model. By having a client-server model, we eliminate the complexities associated with information dissemination among users that MAC protocols of traditional methods are required to handle. This removes a lot of overhead in data transmission between users.

In addition to these changes, our approach also has the benefit of removing the delays associated with data transmission between users. Secondary users will only need to provide information to the server. Therefore, there is no need for direct communication among each other.

The spectrum handoff scenario in cognitive radio must ensure that a pre-empted secondary user by a returning primary user maintains low latency while transitioning to another available channel. This can be quite difficult in traditional methods where spectrum scanning must first be done to find a channel. The delay caused by scanning can lead to higher latencies and reduce QoS that the user expects. Reserving several

bands explicitly for the handoff scenario is one solution to this problem, but also has the disadvantage of secondary users reserving too many bands. It also means these reserved channels cannot be used for other purposes. This reduces the efficiency of the overall spectrum.

In our case, handoff is easily accomplished since the server will be notified upon the arrival of primary user. If they match the channel that is currently occupied by a secondary user, it can inform them of a returning primary user and suggest an alternative available channel. This is possible since the server has all the current user spectrum usage information. There is no need for the sensing or the scanning to find another channel. The only delay comes from the messaging that must be sent between the server and user.

### **3.1.3 Disadvantages of our Solution**

Although there are many benefits discussed in the previous section, there still exist disadvantages to using our approach. Often times this may be tradeoffs of disadvantages that exist with traditional cognitive radio approaches. Other times, there exist new problems that need resolving.

A major disadvantage that is covered in more detail in the next chapter is the addition of messaging delay. As mentioned previously, there exists a trade off between our messaging delays and traditional spectrum sensing. Although this delay may be low enough to consider it spectrally efficient, other methods in traditional cognitive radios approaches can achieve lower values.

Requiring a connection be established between a user and the spectrum manager can be seen as a disadvantage to this approach. Frame error rates can affect the connectivity between users and base stations. Messaging delays may become larger in worse conditions and negatively affect the efficiency. This is discussed in more detail in the next chapter.

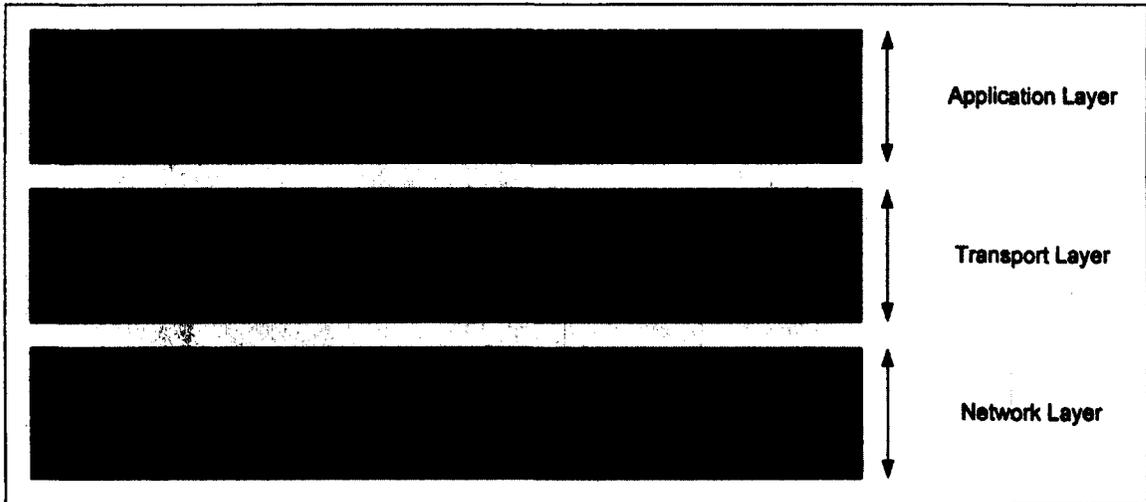
Another limitation is that all users in the coverage area must use this system for it to work at all. The spectrum manager must have complete knowledge of who is using the spectrum in order to manage it. If there are users who use channels and do not report their usage to the spectrum manager, collisions can be caused when the spectrum manager allocates channels.

The security of the system is easily compromised. Rogue users can still use traditional cognitive radios. There is currently no way to enforce the use of this system and have users report their usage to the spectrum manager. This leads to the problem in the previous paragraph, where allocation of channels becomes impossible because there are cognitive radio users who do not report their usage.

### **3.2 Messaging Model**

In this section we will provide a brief overview of the SIP protocol, the extensions that will be used, as well as the SIP messages that will be used in order to facilitate our communication needs.

### 3.2.1 SIP Overview



**Figure 3.4 Protocol Stack**

SIP is an application layer control protocol defined in [44] as an Internet Engineering Task Force (IETF) standard SIP is used as a protocol for multimedia applications over IP-based networks. SIP is also an evolving protocol with new applications such as presence being added into the existing protocol. SIP is flexible enough to support these new applications due to its forward-looking framework and XML-based encoding scheme. Figure 3.4 shows the protocol stack for our model.

SIP is designed to provide the following functionality between two or more end points [45] [46].

- Session initiation
- Session management
- Session termination

These operations are conducted using SIP messages transmitted between end points. Session initiation includes user discovery and requesting session initiation using an INVITE message. Once the user has successfully established a session, a media stream can directly be transmitted between the end points. Modifications can then be made by users regarding an existing session, such as modifying the QoS to meet their requirements. When transmission is finished, session termination will release the session using a BYE message. The SIP messages themselves are independent of the media being transmitted and uses other protocols for the transporting the actual media. SIP is based on a request/response model. The messages are sent between nodes as part of the protocol [47]. The body of a SIP message includes a description of the session. SIP messages can be carried by User Datagram Protocol (UDP) or Transmission Control Protocol (TCP) [48]. When SIP runs over TCP as a transport protocol, the transport layer provides the reliable transmission of messages. When SIP is run over UDP, the responsibility of reliable delivery falls on the SIP. Both UDP and TCP have been used for SIP transport. SIP will be used as a mechanism to implement the spectrum allocation process across heterogeneous wireless networks. Specifically, we will use the SIP extensions of Event Notification Framework (ENF) as well as Event State Publication (ESP). We will show that the SIP ENF with minor modifications can support our proposal effectively.

ESP is a framework for the publication of event information from an event publication agents to an event server using the PUBLISH message. The event server can then distribute this information to watchers. The distribution of event information is standardized in ENF, which is discussed in more. A spectrum manager with this function can collect all the information on channel usage from primary user and share this

information with secondary users so they will not be interfered with. All the information sent to the spectrum manager through the PUBLISH message can then be used to determine the remaining available channels. The collected information published from primary users as well as the information from SUBSCRIBE message of secondary users can be used to distribute channels.

ENF is a framework to notify interested parties of information they have subscribed to and wish to be updated upon. It is a general infrastructure for all classes of SIP event subscriptions and notifications [49]. ENF consists of network entities, watcher and event states (ES). Event states are any information related to a resource. Watchers subscribe themselves to an ES it wishes to receive notification from under certain conditions. The ES will send notifications to the watcher when that event information is updated. ENF allows the watcher to request information from ES, which indicates the occurrence of certain events. We use ENF to subscribe for new channel allocation opportunities.

### 3.2.2 Messaging Entities

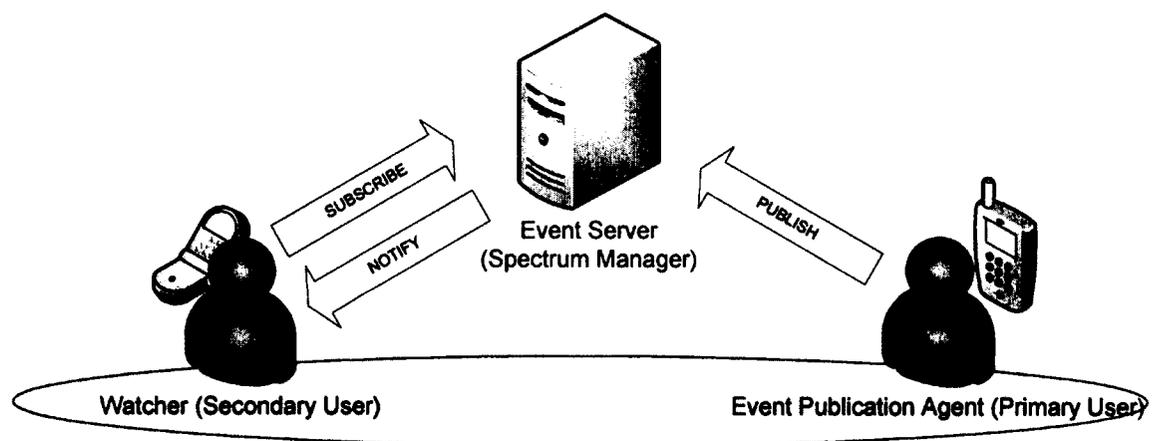


Figure 3.5: ESP and ENF network entities

The model is designed for collecting information sent from primary users to the spectrum manager, as well as to allow secondary users to subscribe to this server for information that primary users provide. The proposed architecture is established based on ESP and ENF frameworks. Figure 3.5 gives a diagram of ESP and ENF networks entities and their relationship with each other. The definitions for the network entities are show below:

**Event Server:** Collects the information sent from Event Publication Agents. The event server will also manage information and notifying the watcher about the requested event information through a NOTIFY messages. In our model, the spectrum manager will serve as the event server as well as the spectrum broker by using the information collected from primary users and then sending the NOTIFY message based on its decision.

**Event Publication Agent:** A network entity that can send event information to the Event Server using the PUBLISH message when the event information needs updating. The primary users will be our event publication agent and provides event information on the channels they want to use using the PUBLISH message.

**Watcher:** A network entity that subscribes to event information from the event server through SUBSCRIBE messages. Secondary users will fill the role of watcher in our model. Upon subscribing to the spectrum manager, they will be sent NOTIFY message when their request for channel has been secured.

### 3.2.3 Messages for Proposed Solution

Using a centralized server to coordinate channel usage means the spectrum manager must have information on all channel usages. This means both primary users and secondary users must share information with the server before using channels. Users will communicate with the server using SIP messages. The following is a list of the SIP messages needed and their functions.

The messages that we propose for our model include the SUBSCRIBE message, the PUBLISH message and the NOTIFY message that are defined by ENF and ESP. These messages give our secondary users (watchers) the ability to receive information on spectrum utilization without the need for constant requesting information with the spectrum manager (event server). By subscribing with the event server, we remove the need to constantly request information. Instead, the event server will contact the watcher when deemed necessary when new information from primary users (event publication agents) become available. This method is dramatically more efficient in terms of message signaling and fits perfectly within our goal of distributing channels among secondary users dynamically. The following describes our proposed messages in more detail.

**SUBSCRIBE:** Gives our secondary users the ability to request notification of an event or set of events at a later time. SUBSCRIBE messages are created by a secondary users, and sent to the spectrum manager when they are requesting use of available channels. When a subscriber wishes to subscribe to a particular state for a resource, it forms a SUBSCRIBE message. This message request will be confirmed with a response. The

200-OK message response indicates that the subscription has been accepted, and a NOTIFY will be sent. A 200 message acknowledges authorization has been granted. Other responses such as 202 indicate an understood message that has not been authorized yet.

The SUBSCRIBE message holds three pieces of information: Uniform Resource Identifier (URI) of the contacting spectrum manager, the event type to subscribe to, and the message body that can contain additional information. The URI will contain routing information to contact the server. The event will be the availability of unused white space in the licensed band. Additional information will contain a bid on the rate to pay, quantity of channels, and any more information deemed necessary.

IETF specifies that SUBSCRIBE messages have expirations. When expire time is reached, they are automatically unsubscribed. Each SUBSCRIBE message contains an "Expire" header that will indicate the duration of the subscription. The 200-OK response will also contain "Expire" header to indicate to the user the subscription time that they have to re-subscribe.

To terminate a subscription, users send a SUBSCRIBE message with "Expire" header value of 0. A NOTIFY will be sent with a "Subscription-State" header value of "terminated".

**PUBLISH:** Allows our primary users to publish event information to the spectrum manager. The published information should include which channels they are about to start using. The PUBLISH message allows a user to create, modify, and remove state in another entity which manages this state on behalf of the user. Multiple users may publish

their own unique states. Then the state compositor can create a composite event state of the resource, i.e.: channel usage information composited to create an overall picture of the spectrum at the spectrum manager.

An initial PUBLISH request is created by primary users and sent to the spectrum manager of the Request-URI header. This initial PUBLISH message contains a body with the published event state. It also may contain an “Expire” header field with suggested lifetime of the event state. If none is provided, the server can choose a suggested one. Publications will expire unless refreshed before its lifetime is exceeded.

Modifying an already published event state is the same as creating an initial PUBLISH message, except the message contains a “SIP-If-Match” header field to match an already existing event state. Successful PUBLISH messages will be responded to with a 200 response

In order to remove an already published event state, primary users will send a PUBLISH message with “Expire” header value of 0 and no additional body. This is the same way subscribers in ENF terminate subscriptions.

**NOTIFY:** An SIP method used by the spectrum manager to notify secondary users of new events, which has been requested by an earlier SUBSCRIBE method. The NOTIFY method contains information about the event which the watcher is interested in.

The next chapter will discuss the analytical part of the delays, and channel efficiencies. We will quantify the delays our model will be dealing with. Chapter five will then present the results of our simulations compared with our analysis from chapter 4.

### **3.3 Auctioning Scheme**

To demonstrate how SIP can be used to create a channel allocation scheme, we provide an example by creating an auctioning mechanism that will grant access to unused channels to secondary users based on their bid. Many other channel allocation schemes can be considered based on other criteria. We choose to use auctioning as an example to show a well known channel allocation method that also integrates a billing plan that provides an incentive for primary users to share their usage information.

By auctioning unused spectrum, primary users can make extra income with very little interference to their existing services. This motivates primary users to share their usage information with secondary users so that their benefits can be maximized.

There are various auctioning mechanisms as discussed in [50]. They can be either sealed or open bids. Open auction needs to go through several rounds of bidding. It is not a good option for our proposal due to its increased delay. In a Sealed bid auction, only one round of bidding takes place and all bids are hidden from everyone else. This will be ideal for our purposes of having low overhead. The auctioning mechanism used for our proposed system includes a first-price sealed bid auction and a second-price sealed bid auction.

A first-price sealed bid auction is a well understood and intuitive way auctions work. The highest bidder wins and pays the highest price. In the second-price sealed bid auction the highest bidder wins the channels they are bidding for, but pays the second highest price. In [50] discusses the advantages of second-price sealed bid auctions over a first-

price sealed bid auction. In general, it gives bidders an incentive to bid truthfully on how much they value the channels.

Amount paid in a multi-item second-price sealed bid auction is determined using the Vickrey Clarke Groves (VCG) mechanism [51]. Let  $B$  be the set of available channels;  $A$  is the set of bidders and  $V_A^B$  the maximum total value achievable for all combinations of  $B$  and  $A$ . Bidder  $j$  pays an amount for channels  $i$  equal to:

$$V_{a-j}^b - V_{a-j}^{b-i} \quad (3.1)$$

where  $V_{a-j}^b$  is the maximum total value achievable with bidder  $j$  removed and  $V_{a-j}^{b-i}$  the value of the auction with bidder  $j$ 's bid for channel  $i$  removed. Therefore, Eq. (3.0) gives the total amount that bidder  $j$  pays for winning channels  $i$ .

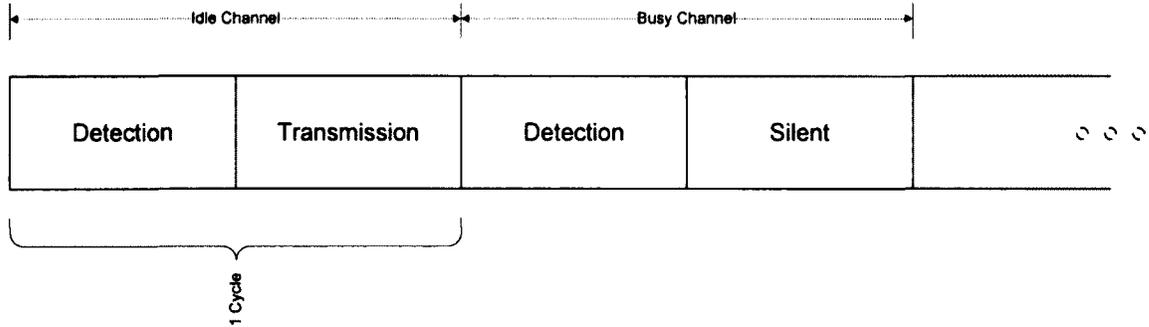
Bids are included along with the number of channels requested within the subscribe message that the secondary users submits to the server. In the event that there are sufficient channels available for all users, they will be charged their bidding price. If insufficient channels are available, the VCG mechanism can be used to determine the winning bidders along with the price they pay.

### 3.4 Criteria for Evaluation

We consider a cognitive radio system where channels are available and two groups of users, secondary users and primary users are present. Primary users are given priority to use channels over secondary users. However, the primary user is not always occupying

channels, which leads to the underutilization in time. Secondary users may only use channels in the absence of the primary users.

In traditional cognitive radio sensing methods, to avoid interference with the primary user, secondary users must first detect the channel for signs of primary user before transmission. This process must be carried out periodically in case of returning primary users. Therefore, secondary users employ cycles of detection and transmission, when no primary users are present. This is the detection cycle as shown in Fig 5.1.



**Figure 3.6: Detection Cycle**

The duration of the detection cycle is decided by the primary user system. A detection cycle is divided among an amount of sampling symbols. Part of the sampling symbols is used for channel detection and the rest for transmission/silent. Let  $N$  be the number of sampling symbols and  $n$  be the number of symbols used for detection. The channel efficiency is then defined as the ratio of time spent detecting over the total idle time of the channel.

$$\text{channel efficiency} = \frac{N - n}{N} (1 - P_{FD}) \quad (3.2)$$

Where  $P_{FD}$  is the probability of false detection. There is a tradeoff between the symbols used for the detection versus the probability of false detection. While decreasing the symbols for detection may increase the number used for actual transmission, this degrades detection performance resulting in higher false detections.

In our model, there is no need for the detection cycle or the possibility of false detection. Rather than time spent on detecting channels for signal presence, overhead comes in the form of message signaling delays from SIP messages. The tradeoff is then between this messaging delay in our model versus the detection time required in traditional cognitive radio methods.

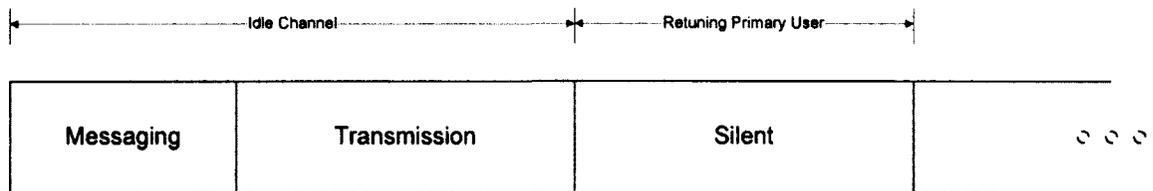


Figure 3.7: Working Model

Our definition of channel efficiency is the ratio of time the secondary user uses the channel for data transmission purposes over the total time. In our system model, the amount of time spent sending SIP messages for channel allocation between users directly affects the utilization. In our working model, a channel operates in two modes, either in use, or not in use. Secondary users may arrive to use the channel at any time. The time it takes to be granted permission to use that channel will be the message sequence delay for a particular scenario. It is important to know the message sequence delays for common scenarios in order to find out how much time the channel is being left idle due to this process. This will allow us to determine the ratio of time spent actively using channel for data transmission over the total time.

Our model of efficiency will be the ratio of time spent by secondary users on data transmission over the total time the user occupies. The total occupied time is the time between when a secondary user arrives looking for channels, to the time it is finished data transmission. Figure 3.7 gives a representation of efficiency.

$$\text{channel efficiency} = \frac{\text{Transmission}}{\text{Messaging} + \text{Transmission}} \quad (3.3)$$

A more complete description of how we quantify and model this is presented in the next chapter.

## Chapter 4

# Analysis and Simulation Evaluation

In this chapter we describe the analysis of our proposed model. We first introduce the types of delays that are factors to consider and the assumptions that we make for our analysis. This chapter will also give equations for overall delay in certain common scenarios that we expect using the delay analysis we find here.

There are several kinds of delays that may affect the overall performance of this system. These delays are affected by several different components as well as the sizes of messages that are being sent. Messages play an important role in determining the overall delay since the scenario that is occurring is based on the type of messages sent. The following discusses various potential delays and factors that can contribute to them as well as various message sizes.

### 4.1 Delay Components

**Queuing:** Queuing delays are experienced by messages at each network entity node that they traverse; starting from the source node, the user device, to each network entity along the way to the destination. When packets arrive faster than they can be processed, they will wait in a queue until they can be processed.

$$queuing = \frac{1}{\mu - \lambda} \quad (4.1)$$

Where  $\mu$  is the service rate at the device,  $\lambda$  is the arrival rate of packets and  $\mu > \lambda$

**Processing:** Processing delay is the time taken by routers to process packet headers. Depending on the type of processing being performed on packets and processor speeds, the amount of time needed for processing may be insignificantly small compared with other forms of network delay.

**Propagation:** Propagation is the delay caused by the amount of time taken for the packet to physically traverse the medium to the destination. This is determined by the length of the medium as well as the propagation speed over the medium.

$$propagation = \frac{d}{s} \quad (4.2)$$

Where  $d$  is the distance between the two points and  $s$  is the speed through the medium. Usually this is the speed of light or close to the speed of light.

**Transmission:** Transmission delay occurs when placing packet bits onto the transfer link. The data-rate of the link and size of the packet determines transmission delay. In our analysis, we calculate the delay over TCP using the transmission and propagation delays.

$$transmission = \frac{b}{r} \quad (4.3)$$

Where  $b$  is the number of bits within the packet and  $r$  is the transmission rate.

These are the typical packet transfer delays introduced by using SIP message signalling as opposed to spectrum scanning. Factors that determine these values may include server processing power, length of queue, and network throughput.

SIP messages requests and responses, are the basic unit of SIP communication. At minimum, messages contain the following fields: 1) Start Line 2) At least one header 3) empty line 4) body (may be empty). Messages sizes can range widely depending on the number of headers, body content, and length of URI. Table 4.1 gives the typical sizes of SIP messages including overhead and the amount of frames they take up. The number of frames is determined by the message size, the duration of a frame, and the channel capacity.

For our delay analysis, we are interested in finding the delay at certain parts of the transmission path that we expect our messages travel. This includes delays at different nodes as well as delays for the different network along the path. These networks include the LTE access network and the outside network that the SIP proxy servers and spectrum manager resides in. The summation of the individual delay components will give us an overall delay expected from our solution. Figure 4.1 gives an overall picture of the delays along the transmission path that we expect to encounter. The following subsections will derive the expressions for each part of the delay.

**Table 4.1: SIP Message Sizes**

<b>Message</b>	<b>Payload Size (bytes)</b>	<b>Message Size (bytes)</b>	<b># Frames (9.6Kbps)</b>	<b># Frames (19.2Kbps)</b>	<b># Frames (2Mbps)</b>
<b>TCP SYN/ACK</b>	74	114	6	3	1
<b>SUBSCRIBE</b>	569	609	30	16	1
<b>SIP NOTIFY</b>	718	758	37	19	1

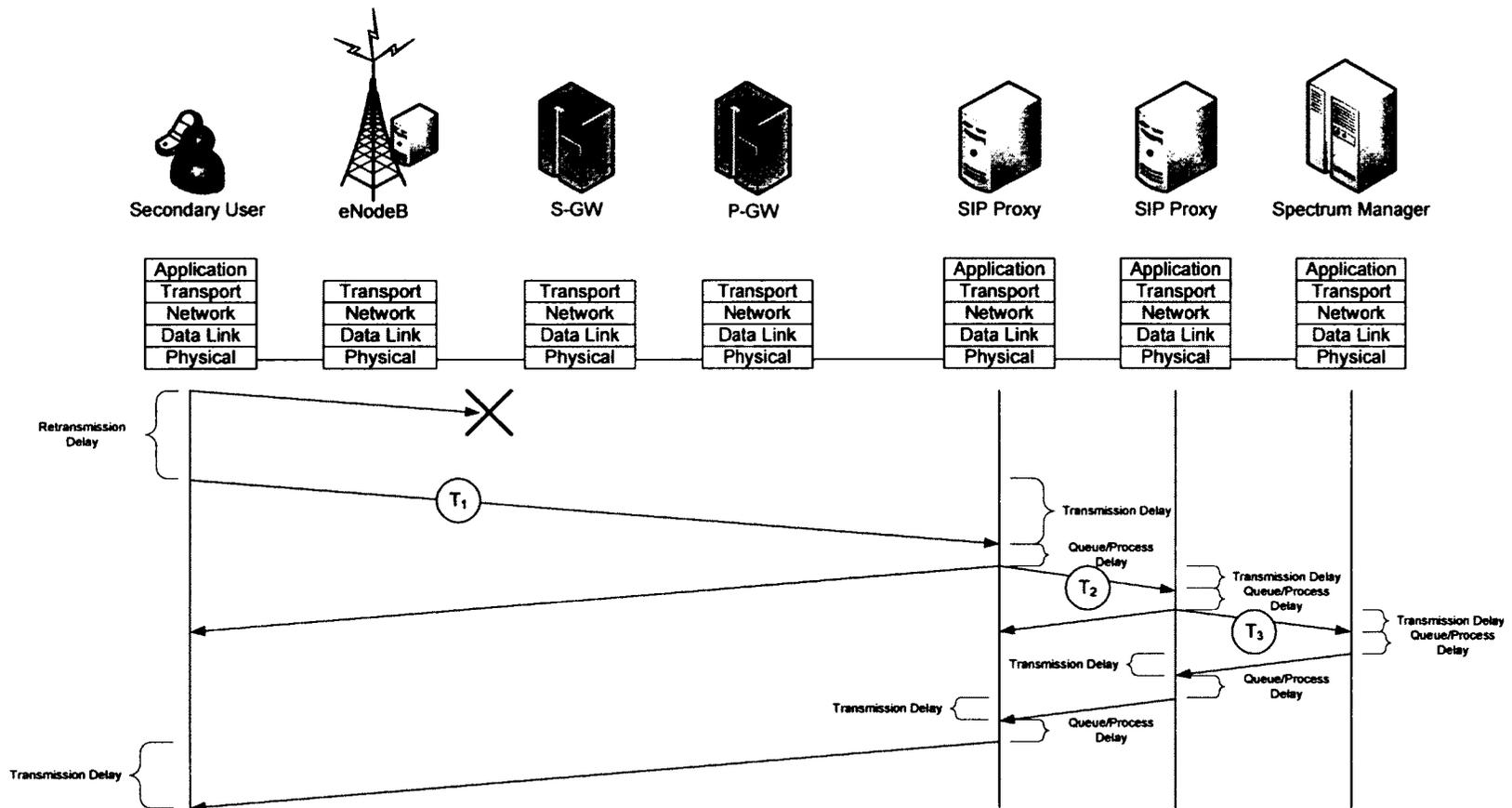


Figure 4.1: Network Delay Components

### 4.1.1 Access Network

Access Network Delays include the transmission delay (both uplink and downlink), the queuing/processing delay of network entities, and retransmission. The number of frames depends on packet sizes, link capacity, and frame length. One way transmission through the access network depends on the number of frames, the inter-frame time, and propagation delay. End-to-End TCP delay from the user to the first SIP proxy would include this transmission delay, and queuing/processing delays at each network entity along the path.

**Table 4.2: LTE component delays [57]**

Delay Component	Delay Value (ms)
Propagation delay	0.5
eNodeB delay estimate $Q''$	2
Processing delay	varies

We first examine delay in the air link between users and the eNodeB. This delay depends on the message size, link capacity, frame length, interframe time and propagation delay of the message across the radio link.

The radio link delay for a message with  $k_1$  frames and the reply with  $k_2$  would thus be:

$$T' = D + (k_1 - 1)\tau \text{ and } T'' = D + (k_2 - 1)\tau \text{ respectively}$$

Where  $D$  is the frame propagation delay over the access network for the message,  $k_1$  and  $k_2$  is the number of frames for the message and its response and  $\tau$  is the inter-frame time.

Another component of the access network includes the delay caused by queuing/processing at the user node. If we assume that user equipment are sufficiently powerful that processing delays are marginal compared with queuing, we can focus on queuing model for our delay analysis. An M/M/1 queuing model is used for our analysis since we assume packets are served single and sequentially, messages arrivals are Poisson distributed, and service times are exponential. The queuing delay at the user node will be:

$$U = \frac{1}{\mu_r - \lambda_r} \quad (4.4)$$

Where  $\mu_r$  is the service rate and  $\lambda_r$  is the arrival rate of messages both at the secondary user node.

Other delays within the access network include the queuing/processing delays at the eNodeB, and the EPC which includes the S-GW and P-GW. Table 4.2 gives delay values for these other components in the LTE architecture. Altogether, the one-way delay for the LTE access network is:  $T' + Q' + Q'' + U$

#### **4.1.2 First TCP Connection**

For our analysis, we plan to use TCP as our transport protocol. Retransmission will be handled at this layer with SIP retransmission disabled. Using TCP as the transport protocol introduces some additional delay due to the handshake of TCP during connection establishment and possible buffering in the TCP stack. TCP protocol runs between SIP network entities. The first one runs between the user and first SIP proxy

and subsequent TCP connections exist between SIP proxies until reaching the destination spectrum manager.

TCP prioritizes reliability over low latency. Higher latencies may be acceptable under our circumstances, since they are not used for a real time application such as VoIP. If we assume that the underlying network is LTE based in an all IP network with E-UTRAN as the air interface for mobile devices, SIP messages transported over TCP can be sent anywhere in the internet. SIP messages are relatively small and several hundred bytes are the maximum they usually become [47]. E-UTRAN can provide broadband access to mobile devices making messaging more feasible.

We start by looking at delay of the first TCP connection. This will include the delay caused by the access network and any delay caused by the Internet. The one-way delay from user to the first SIP proxy server with no retransmissions is:

$$A' = U + T' + Q' + Q'' + I \quad (4.5)$$

And the one-way delay for SIP proxy to user with no retransmissions is:

$$A'' = U + T'' + Q' + Q'' + I \quad (4.6)$$

Where  $I$  is the delay caused by transmission and routing from the edge of the EPC to the first SIP proxy server.

We assume that the SIP proxy servers serve dedicated tasks in SIP message routing and do not perform other functions. They follow general Poisson message arrivals and loads containing only SIP messages.

$$T^p = \frac{\rho_s}{\lambda_s(1 - \rho_s)} \quad (4.7)$$

Where  $\rho_s$  is the SIP message load and  $\lambda_s$  arrival rate at the SIP proxy server with  $\lambda_s \geq \lambda_r$ . The round trip time for a successful transaction includes the message sent to the first SIP proxy as well as the acknowledgement sent back to the user. Messages pass through the Internet and access network twice on their way to their destinations:

$$\text{round trip time} = A' + A'' + T^p \quad (4.8)$$

The TCP transport protocol will retransmit lost frames or frames with errors. Retransmission leads to additional overhead in terms of delay and is one factor that must also be accounted for. Since the round trip time is the expected amount of time to receive a response, we will use this as our initial retransmission timer  $\text{round trip time} = T^r(1)$ . For message retransmission purposes, we will consider an adaptive retransmission timer. The back off timer after the  $i$ th transmission doubles after each retransmission.

$$T^r(i) = 2^{i-1} \cdot T^r(1) \quad (4.9)$$

Retransmission may occur if frames are in error. Let  $p$  be the probability of frame error (FER) in the air link. If we assume that a message sent has  $k_1$  frames and the subsequent response has  $k_2$  frames, then,  $q$ , the probability of retransmission is calculated in [54] as:

$$q = 1 - ((1 - p)^{k_1+k_2}) \quad (4.10)$$

We are interested in finding the average delay for a successful transmission given the number of frames within the message, and the maximum number of retransmission allowed.

Let us first consider the case of a successful transaction. There is no retransmission and the expected delay should be the time for a message sent from a user to the first SIP proxy server. Thus we have:  $(1 - q)(A')$  the probability of successful transmission and the expected delay.

The second scenario to consider is that the first transmission fails and the subsequent retransmission is successful. The expected delay is thus the retransmission timer expiring plus the successful transmission. We then have:  $(1 - q)q(T^r(1) + A')$ .

The next scenario has the first transmission and retransmission failing and the second retransmission successful. Thus the delay is equal to the first retransmission timer expiring  $T^r(1)$ , the second retransmission timer expiring  $2T^r(1)$  plus the successful transmission:  $(1 - q)q^2(3T^r(1) + A')$ .

These scenarios continue until the last retransmission attempted is successful where  $M$  is the maximum number of retransmissions:  $(1 - q)q^M \cdot ((2^{M-1} - 1)T^r(1) + A')$ .

Altogether, the average delay for a successful TCP transmission between the user and the first SIP proxy server is:

$$T_1 = \frac{1}{1 - q^M} [(1 - q)(A') + (1 - q)q(T^r(1) + A') + (1 - q)q^2(3T^r(1) + A') + \dots + (1 - q)q^M \cdot ((2^{M-1} - 1)T^r(1) + A')]$$

This simplifies to the expression:

$$T_1 = A' - T^r(1) + \frac{(1-q)(1-(2q)^M)}{(1-q^M)(1-2q)} \cdot T^r(1) \quad (4.11)$$

#### 4.1.3 Other TCP connections

Subsequent TCP connections exist between the first SIP proxy server and the next SIP proxy server that is closer to the spectrum manager. To simplify our calculations, we assume that proxy servers have the same queuing delays as well as the same loads and arrival rates.

Delays to consider include the queuing/processing delay at the first SIP proxy server, and the average delay for transmission between the first SIP proxy and the second SIP proxy. For simplification purposes we will assume the transmission from SIP proxy to the next SIP proxy are the same as the one from the access network to the first SIP proxy. The one-way delay for this component with no retransmissions is:  $B = T^p + I$  and the round trip time for this component that we will use as our initial retransmission timer for this component of the delay is:  $T^x(1) = T^p + 2I$ . If we assume  $L$  is the probability of packet loss between two SIP proxies, the average delay would be calculated in a similar way as before:

$$T_2 = B - T^x(1) + \frac{(1-L)(1-(2L)^M)}{(1-L^M)(1-2L)} \cdot T^x(1) \quad (4.12)$$

The last part TCP connection exists between the last SIP proxy and the spectrum manager. The spectrum manager may not be dedicated services and have non SIP message loads. We consider an M/G/1 non-preemptive priority queue with messages that

have higher priority than SIP messages. Let  $\rho_0$  be the load with higher priority. The load is based on the arrival rate  $\lambda_0$  as well as the mean service time  $\bar{X}_0 = \frac{1}{\bar{\mu}_0}$ . Thus, the traffic for non-SIP and SIP messages is  $\rho_0 = \lambda_0 \bar{X}_0$  and  $\rho_s = \lambda_s \bar{X}_s$  respectively with total traffic  $= \lambda_0 \bar{X}_0 + \lambda_s \bar{X}_s$ .

Let  $R$  be the mean residual service time for jobs currently being served when an arrival of any priority occurs.  $R = \frac{1}{2}(\lambda_0 \bar{X}_0^2 + \lambda_s \bar{X}_s^2)$ .

We are interested in the mean waiting times in the queue for jobs of SIP priority. Let  $J_0$  and  $J_s$  be the number of jobs waiting in the queue for non-SIP and SIP messages respectively. Then, the mean waiting times in the queue for higher priority non SIP messages is  $T^0 = R + \bar{X}_0 J_0$  which simplifies to  $T^0 = \frac{R}{(1-\rho_0)}$

For SIP messages  $T^m = R + \bar{X}_0 J_0 + \bar{X}_s J_s + \bar{X}_0 \lambda_0 T^0$  which leads to  $T^m = \frac{R}{(1-\rho_0)(1-\rho_0-\rho_s)}$

Thus, the average queuing delay at the spectrum manager for SIP messages would be:

$$T^m = \frac{R}{(1-\rho_0)(1-\rho_0-\rho_s)} \quad (4.13)$$

Where  $T^m$  is the queuing delays at the spectrum manager for SIP priority messages. The value of  $R$ , given a 5 percent standard deviation of the service rate mean, is calculated in [54] to be  $R = 0.501[\rho_0^2 + \rho_s^2]$ . For simplification purposes we assume the loads of both SIP and non-SIP messages are the same.

The delay from the last SIP proxy to the spectrum manager is calculated the same as before except the queuing delay from the SIP node is  $T^m$  instead of  $T^p$ .  $C = T^m + I$

$$T_3 = C - T^x(1) + \frac{(1-l)(1-(2l)^M)}{(1-l^M)(1-2l)} \cdot T^x(1) \quad (4.14)$$

#### 4.1.4 Overall Delay

The overall delay caused by messages is the sum of delays incurred by the access network, the internet, SIP proxy nodes, and spectrum manager.

$$T = T_1 + T_2 + T_3 \quad (4.15)$$

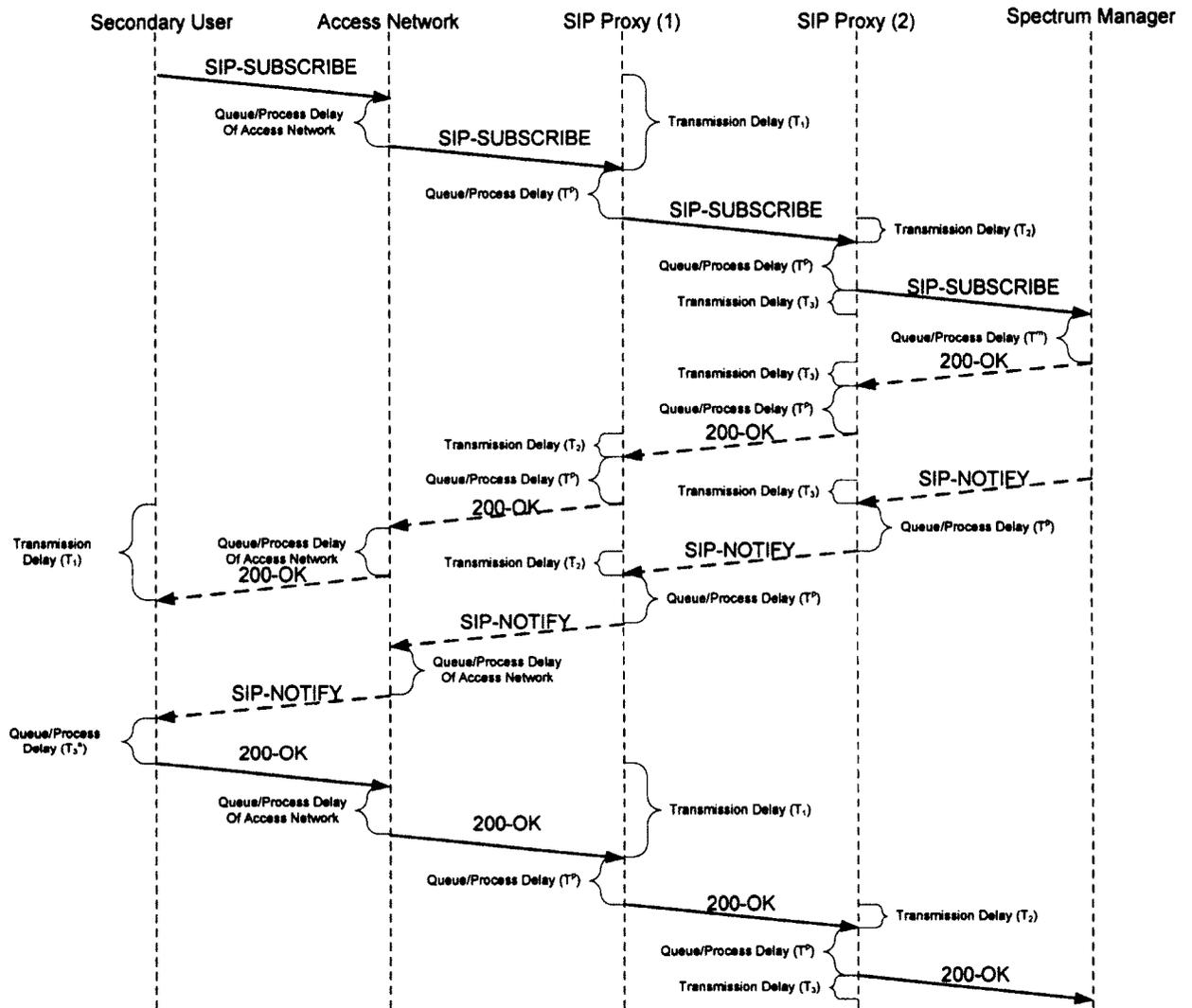
## 4.2 Message Flows

The allocation of channels is handled by the spectrum manager. There are various events that may trigger the allocation/reallocation of channels. These include the cases when a primary or secondary user arrives and when primary or secondary user departs as well as further requests made by users while already using channels. The server will maintain two lists. The user list identifies the users which currently occupy a channel or channels. The subscription list identifies those users who are waiting for channels. The following section will give a description of message exchange in each scenario as well as the expected delay using the equations established in the previous section

#### **4.2.1 Secondary User Arrival**

Upon arrival, secondary users who are interested in using channels will first send a SUBSCRIBE message to the server. This message contains information on the number of channels requested as well as their bid. A successfully sent message will be responded to with a 200-OK message. The 200-OK message is used to confirm reception of the previous message and acknowledge that it has been authorized.

There are two potential scenarios upon receipt of the SUBSCRIBE message. When enough channels are available to accommodate the user, the server will send the NOTIFY message to the user after the 200-OK message and use its bid price as the winning bid. A successfully sent NOTIFY message will be responded to with a 200-OK message from the user. This message will confirm the user's choice in using the channel. The user will then be added to the user list. If no response is received when the subscription expires, the server will assume the current contacted user has left. Figure 4.2 shows the message sequencing chart of this scenario.



**Figure 4.2: Secondary User Arrival and Notification**

In this figure, the message flow is shown as it moves from one network entity to the next. The expected delays are highlighted along the transmission path that each message is expected to take in the LTE network. Each entity has a queuing/processing delay upon message arrival. Messages also experience transmission delay as they are being transmitted. From the user to the destination spectrum manager, messages traverse the different entities. This results in some further delay before it reaches the spectrum manager. The following scenarios assumes that messages originating and destined for

the secondary user follows this path and will be omitted for diagram simplification purposes.

Using the above equations, the analysis for the secondary user arrival scenario would include the TCP setup handshake, subscribe, and notify messages is sent through the access network and the SIP network to the spectrum manager. Therefore, the average delay with SUBSCRIBE, OK, and NOTIFY would be:

$$T^A = 3 \cdot T_1 + 3 \cdot T_2 + 3 \cdot T_3 \quad (4.16)$$

There is a possibility that SIP messages are lost during transmission. This may lead to increased delays. There is also a possibility that the message retransmissions are all unsuccessful and thus the message is not sent. In such a scenario, this may lead to unnecessary downtime of a channel or worse a collision between users.

The first scenario may occur if there are channels available, but the SIP message granting access to channels is not received by the user. In this case, the time spent on retransmission is wasted and the user does not gain permission. The event manager may choose the next user in the waiting list. This may happen during the secondary user arrival scenario.

The probability of this scenario happening would be the probability of message transmission failing. Assuming SIP messages are transmitted seven times before they stop, the probability of transmission fail,  $P_f$ , would be.

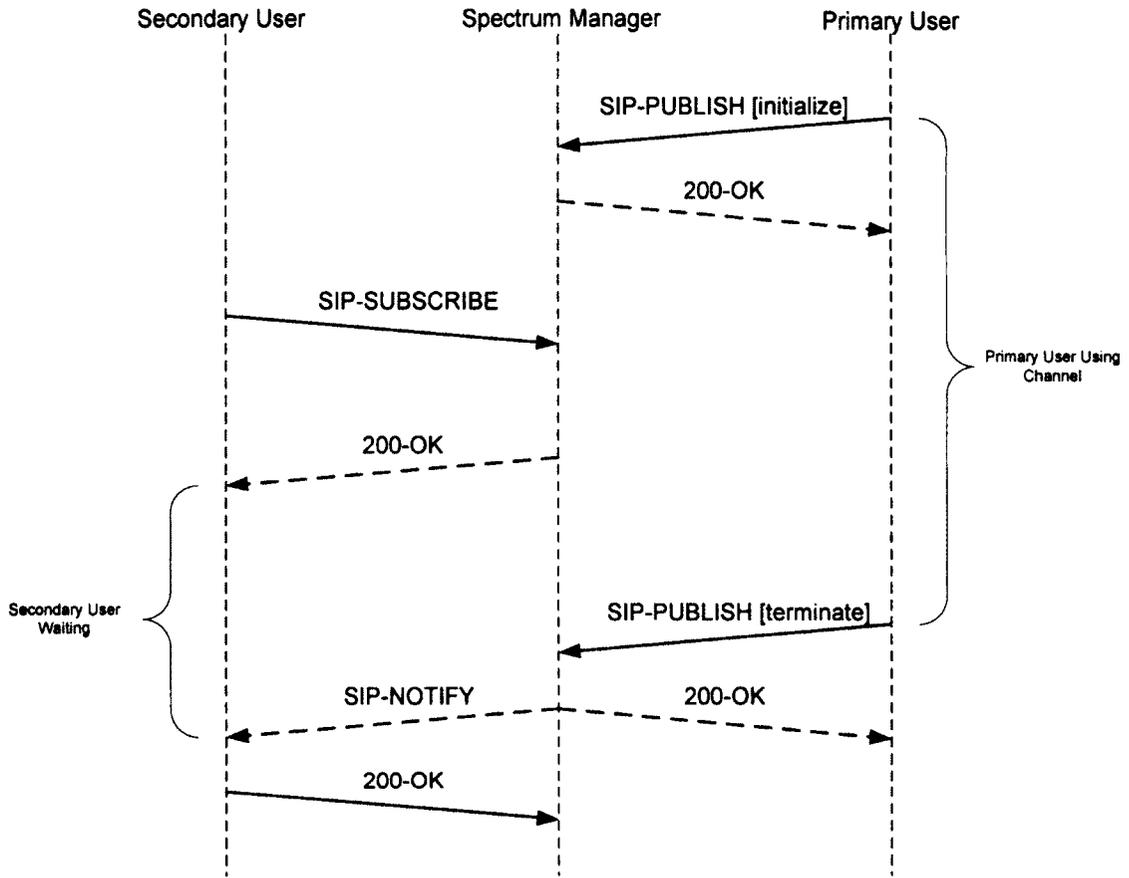
$$P_f = q^7 \quad (4.17)$$

Where  $q$  was previously defined as the probability of retransmission. With a retransmission probability of 30% which is already quite high, the probability that all retransmissions fail is still quite low (0.021%).

With message delay times of  $(2^{M-1} - 1)T^r(1) + A'$  under poor conditions of low transmission rate 9.6Kbps and large message size 758 bytes; the delay would be around 22 seconds after 7 message transmission attempts before giving up. The average delay under this scenario would be several hundred microseconds to milliseconds in scale.

#### **4.2.2 Insufficient Channels**

When there are insufficient channels available, the user will be added to the subscription list until contacted by the server. The secondary user will need to compete with other users through the auctioning mechanism when new channels become available. Users that are waiting can change their bids and number of channel requests at any time by sending another SUBSCRIBE message. In the event that a user sends a new message and is notified of a winning bid with its old SUBSCRIBE message, they will be given the choice to accept or decline the offer. Figure 4.3 shows the message sequencing that occurs when there are insufficient channels available for a secondary user and must wait for more channels to become available.



**Figure 4.3: Secondary User Wait when no channels available**

The secondary user may begin using channels after it receives the SIP-NOTIFY message. It is assumed that this connection will stay for all other transactions until the user chooses to terminate it.

In this scenario, the overall delay will be similar to the previous one except the delay must also take into account the amount of time a user is spent waiting as well as the delay that other users incurs for informing termination of their session.

From the point of view of the overall channel, the time spent unused due to messaging would only include the time spent sending the *i*th message containing the termination message and the subsequent NOTIFY message sent to the waiting secondary user.

$$T^B = 2 \cdot T_1 + 2 \cdot T_2 + 2 \cdot T_3 \quad (4.18)$$

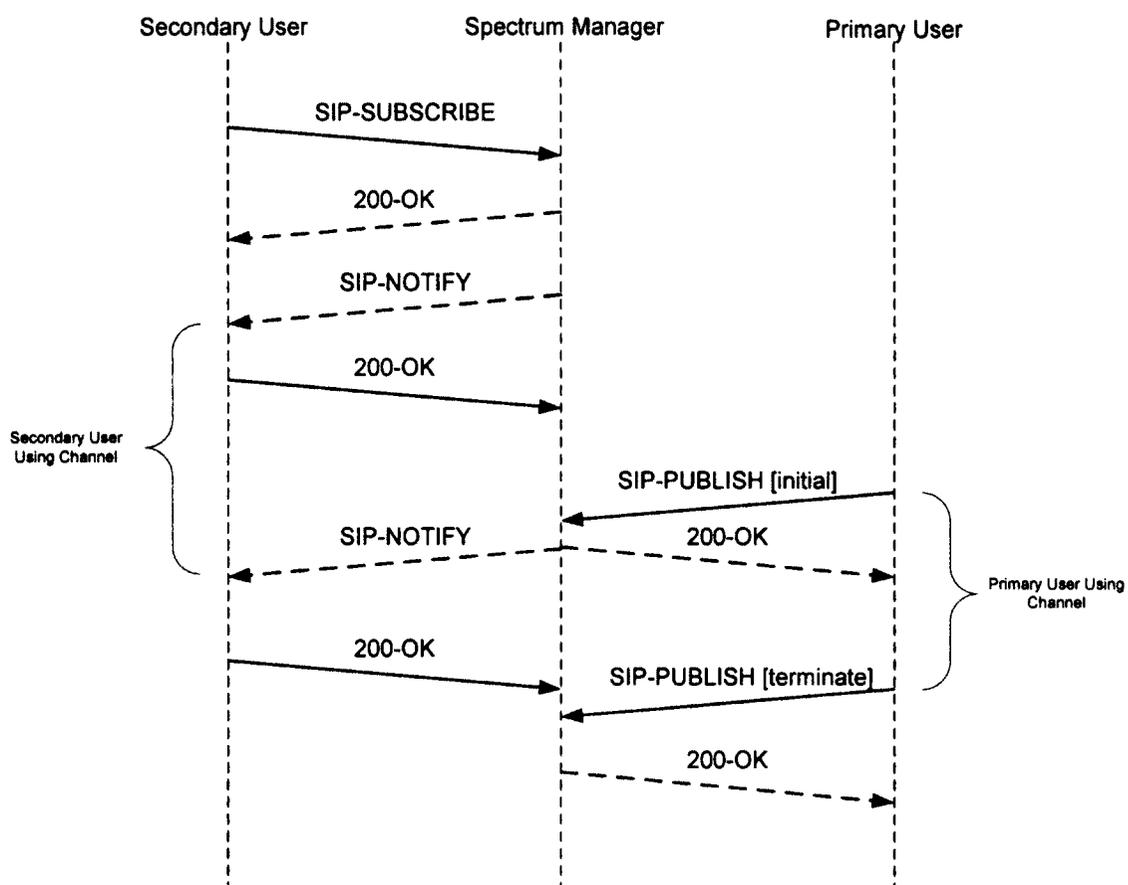
A transmission failure under this scenario is the same as the previous case. However, since all channels are full, it does not reduce channel utilization since all channels are occupied. From the spectrum manager point of view, there is no downtime wasted.

### **4.2.3 Primary User Arrival**

A primary user that arrives will send a PUBLISH message to the server to notify the server of its intention to start using channels. The PUBLISH message contains information in the body on the number of channels as well as which specific channels it will use. Primary users may begin using channels immediately after sending the initial PUBLISH message. It does not need to wait for the 200-OK reply. In the case that a secondary user is incumbent in a requested channel, it will be forced to leave. This is to provide further protection for primary users which the system should be as transparent as possible.

In the scenario where a secondary user is incumbent, the server will send a NOTIFY message to the incumbent user in order to inform them of a primary user arrival on their channel. The secondary user may also have detected increased interference. Secondary users who detect increased interference in the physical layer before receiving NOTIFY message from the server will stop transmission and wait for confirmation from the NOTIFY message of primary user arrival. Upon receiving the NOTIFY message, the preempted secondary will send a 200-OK message to confirm yielding the channel to the

primary user. This scenario is shown in Fig 4.4. Primary users are assumed to have internet access through their own access networks. This may or may not be the same LTE network that secondary users use. Network entities between the primary user and spectrum manager may be very diverse depending on what kind of network they are using. We have omitted them for simplification purposes. If still interested, the secondary user can re-subscribe and compete for the next available channel. It can also choose to end the subscription by sending a terminating SUBSCRIBE message.



**Figure 4.4: Primary User Arrival pre-empts Secondary User**

The arrival of a primary user experiences similar delays to the one for secondary users. A TCP session is setup first between the primary user and the spectrum manger. The

PUBLISH message is sent to the spectrum manager, but the primary user does not need to wait for any NOTIFY message before it begins using the specified channels. It may start using channels when finish sending the PUBLISH message.

The delays that are factors in this scenario include the transmission delay over TCP for the dialog setup as well as transmission of PUBLISH message, the queuing delay of primary user, spectrum manager, and any intermediate network entities.

The equation for delay is the same as the scenario of arriving secondary users. Changes to the variables of number of messages, number of network entities, and queuing and transmission delays may be different.

Since the primary user does not wait for response, there exists a time between when the primary user starts using channel to the time that secondary leaves that is being wasted. Since both users are being denied channel usage due to interference during this time period, the amount of channel usage time being wasted is doubled.

The probability of this scenario occurring must take into account the probability that an arriving user chooses a channel that is currently occupied. Assuming high channel usage, say 90%, and with 40 available channels, the probability of this scenario occurring would be  $0.9^{40} = 0.01478$  or 1.48%

If the conditions for transmission are the same as the previous scenario, the delay would be the same for each user and the time lost is doubled. However, since the probability of this scenario happening are quite low, the overall delay this would cause is still between several hundred microseconds to millisecond range.

#### 4.2.4 User Departures

When users of either user type finish using a channel, they will end by sending a terminating SUBSCRIBE or PUBLISH message respectively to the server. The server that receives the termination message will remove the user from the users list and send a 200-OK message back to the user to confirm that they have been removed. The newly available channels will then be assigned to the next winning bidder.

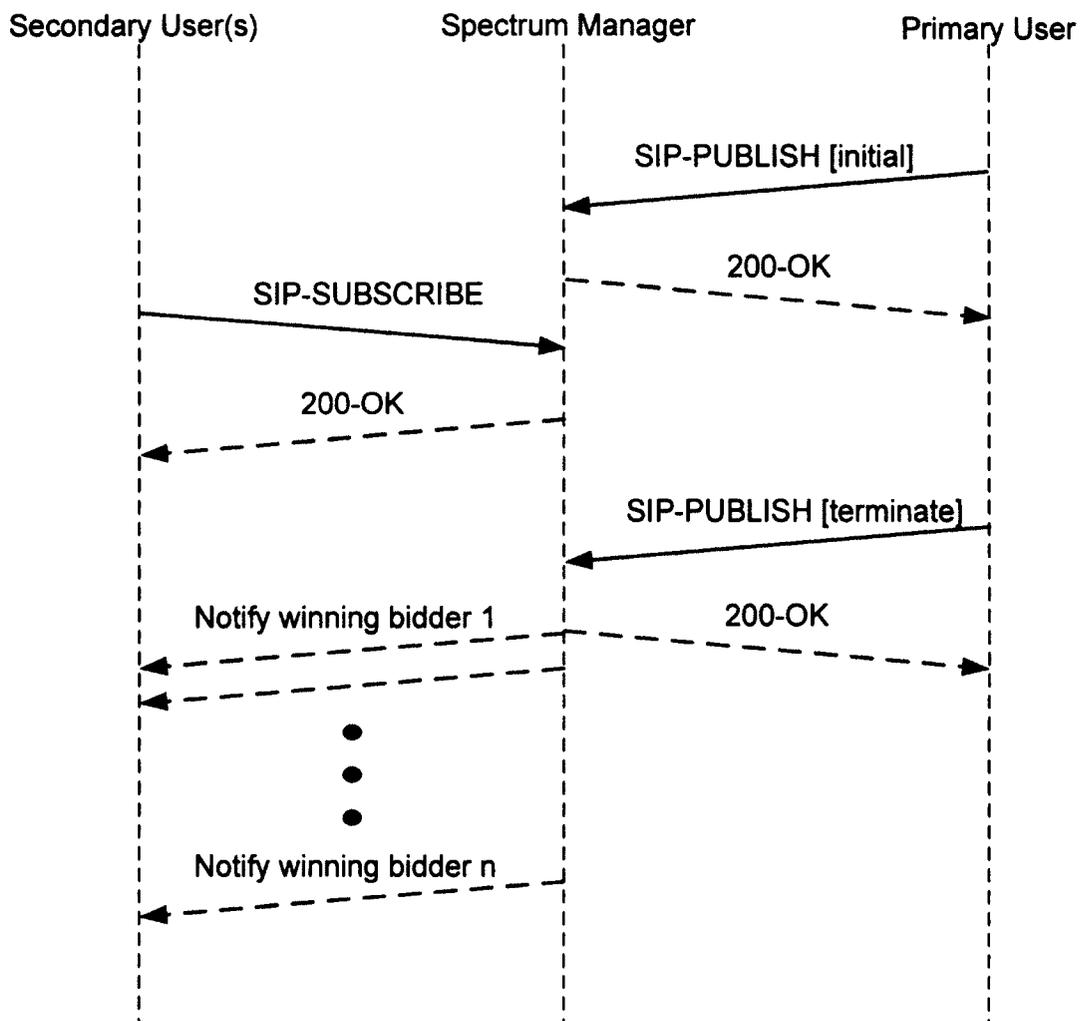


Figure 4.5: Notifying multiple secondary users

Several different scenarios can occur when a user departs from a channel. If not enough channels are available to accommodate the number of channels requested by any users in the subscription list, they will continue to wait until enough channels become available. If one or more secondary users can be accommodated, the server will go through the auctioning mechanism to determine the winners and then notify the winners as. If there are enough channels for all the waiting users, all users will be granted the amount they requested. Fig 4.5 shows notification of multiple secondary users.

If enough channels are available to accommodate multiple users, the server can optimize it based on a number of different criteria. In our case, we demonstrate how to maximize revenue gains based on the information secondary users give to the server with their bid price and number of channels requested. This gives a combination optimization problem based on the unbounded knapsack problem to solve. An approximation to the optimal solution can be calculated by selecting the winners to be users with the highest price to channels requested ratio until no more users can be accommodated.

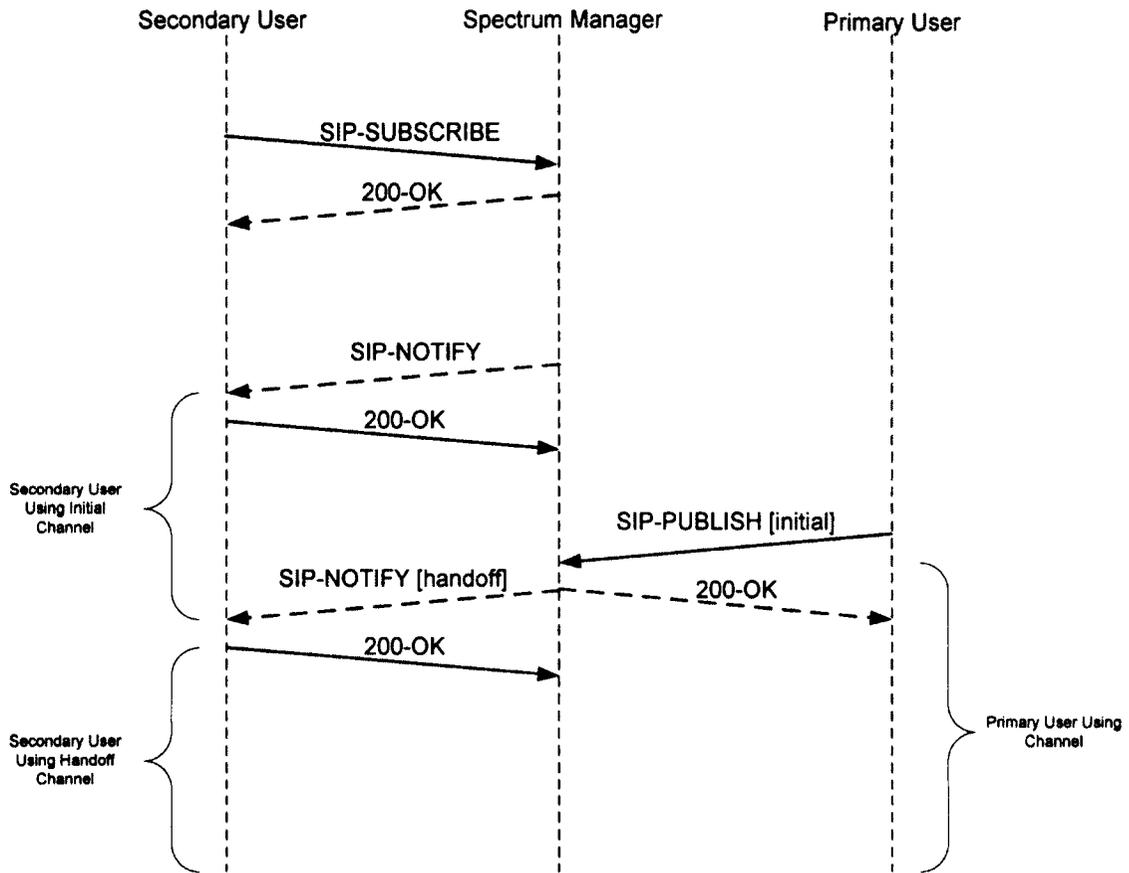
The delay for departing users would include the time to transmit the termination message to the spectrum manager as well as the time it takes for the spectrum manager to send NOTIFY to secondary users. This is the same delay as equation 4.17.

In the event that the terminate message from the user is not received by the spectrum manager, there will be time wasted from channels that are being allocated to users that have stopped using them. Again, the probability for message transmission failure will be the same as the secondary user arrival scenario. After a period of time the users subscription will eventually time out as described in section 4.2.6.

#### **4.2.5 Spectrum Handoff**

When an incumbent secondary user is pre-empted due to the arrival of a primary user, the server has the capability of assisting a secondary user in finding additional free channels that meet its requirements. There are two scenarios that can take place. Either there are sufficient channels and handoff can occur, or there is none available and the secondary user will need to re-subscribe and wait for channels to become available.

In the scenario where there are not enough additional channels to accommodate them, a pre-empted secondary user may re-subscribe and wait for channels to become available or terminate the subscription with the server. A terminated subscription is done using a SUBSCRIBE message with “Expire” header value of 0. If a secondary user wishes to re-subscribe and wait for channels, they will send a SUBSCRIBE message the same way a new secondary user initializes dialog with the server.



**Figure 4.6: Spectrum Handoff**

When there are enough channels available to perform a handoff scenario, the NOTIFY message that the server sends to pre-empted secondary user will contain information on the available channels. This is the same as when their initial SUBSCRIBE message gets notified except it also confirms the arrival of the primary user. They may then choose to either accept or reject the handoff.

In the spectrum handoff scenario, an incumbent secondary user is forced to leave due to the arrival of a primary user on the same channel and subsequent NOTIFY of handoff channels available. This scenario is shown in Figure 4.6. The arrival of a primary user is the same as before with some additional messaging. The delay to consider is the time

between the primary user sending the PUBLISH message until the time that the secondary user receives the NOTIFY message from the spectrum manager.

Similar to the previous scenario where secondary users are waiting, we are interested in the  $i$ th TCP segment that contains the PUBLISH message from the primary user and the  $i$ th TCP segment that contains the NOTIFY message sent to the incumbent secondary user by the spectrum manager for finding the average transmission delays over TCP. The delays experienced include:

TCP transmission delay of PUBLISH message, TCP transmission delay of NOTIFY message. Since the access network for primary user may be wildly different, we will assume that they use the same LTE network as secondary users in order to simplify our analysis.

The overall delay in this scenario would be the same as the one stated in equation 4.18

Transmission failure is a possibility in this scenario. This may happen when a primary user arrives on the same channel that a secondary user is currently occupying. Failure to receive the NOTIFY message from the event server would mean a collision between the two users. The amount of time lost due to message transmission failure in this scenario is compounded by the time each user loses that could have been used for useful transmission. Assuming that a successful handoff would have allowed the secondary user to migrate over to an unused channel, a failure here would lead to time lost by the primary user on the incumbent channel and the time lost by the secondary user on the new channel.

The probability of this happening is the same as the one in the primary user arrival scenario. The handoff is a special case of the primary user arrival event.

#### **4.2.6 Subscription Timeout**

In the event that a secondary user loses connection with the server, the subscription with them will eventually time out. In this scenario, the secondary user who loses connection with the server will stop channel usage when their subscription expires. The server will then re-allocate channels the same as when the secondary user is finished using channels and terminates their subscription. Secondary users who wish to continue using channels will need to successfully re-subscribe. Figure 4.7 shows the message flow for this scenario.

The delay to consider here only includes the NOTIFY message being sent from the spectrum manager to the waiting secondary user. Since there is no message being sent from the expiring subscription, only the one-way NOTIFY message is factored.

$$T^C = T_1 + T_2 + T_3 \quad (4.19)$$

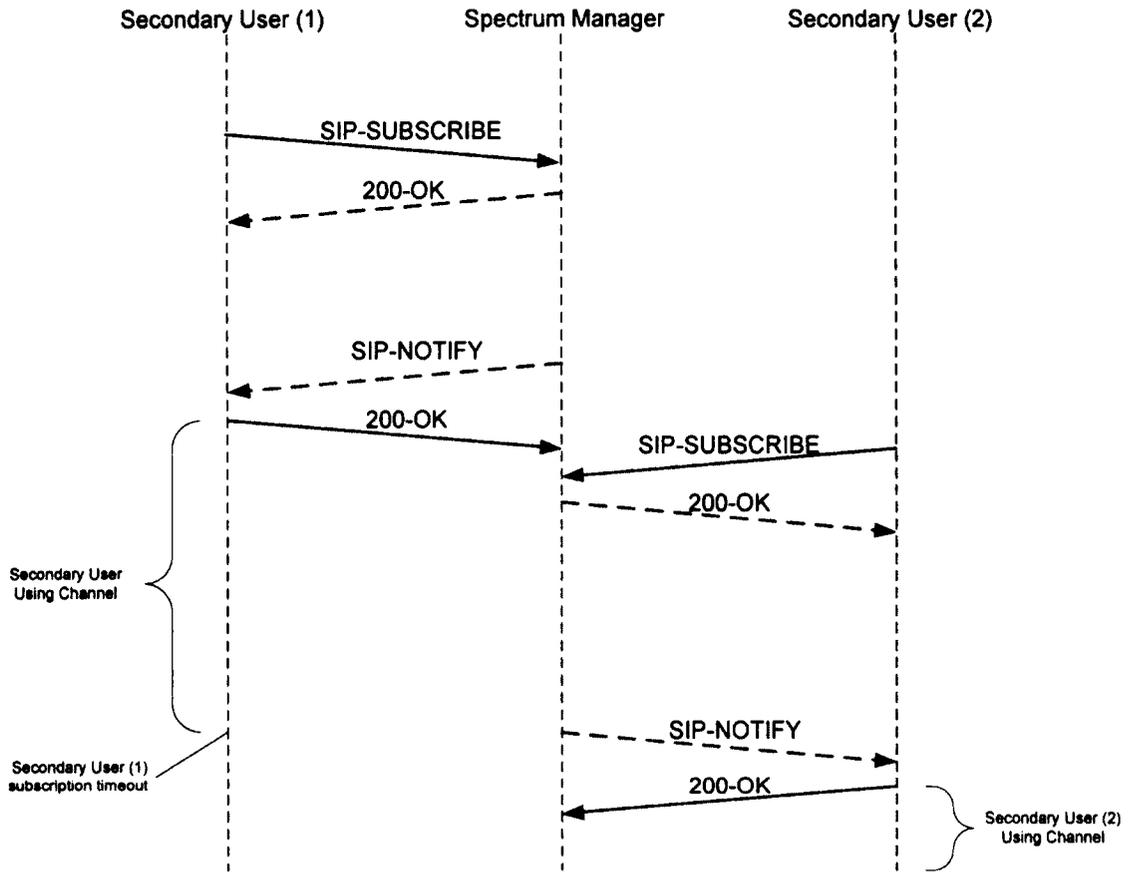


Figure 4.7: Subscription Timeout

#### 4.2.7 User Mobility

In a mobile user scenario, users must always keep their TCP connection alive. That means they must stay within the coverage area of their respective base stations. In the event that they move out of the coverage area, a handoff scenario that passes the TCP connection seamlessly to the adjacent base station that covers the new cell that the user moves into must be accomplished. This may be done the same way that existing LTE user devices switch between different eNodeB's when they move from their coverage area to the next [58].

In the event that the TCP connection is broken due to mobility, the user must stop using the channel and re-establish connection with the event manager before they can start using channels again. The event manager will assume this and reallocate channels to waiting users. This may mean that they need to go through the processes of messaging again before they are granted permission to use channels.

This approach requires that all users have connectivity with the event server to work. In the event that there is no coverage for a certain area the mobile secondary user has moved into, there can be no guarantee to the availability of spectrum as the event manager has no information and cannot service that area. In this situation, the secondary user must rely on traditional cognitive radio sensing methods to find available spectrum.

### **4.3 Channel Efficiency Analysis**

In order to find the efficiency of channel usage using this model we have first determined the delay associated with each scenario. Arriving users must experience a delay associated with messaging according to scenarios established in section 4.3 of this chapter.

The usage of channels by arriving and departing users can be modeled as an M/M/C queue. Where the channels act as servers, arrival rates are arrival rates of users, and service rates are the reverse of the average time spent using channels.

Equations used for the calculation of probabilities for M/M/C model [55] are the following:

The utilization factor,  $\rho$ , is calculated as:

$$\rho_s = \frac{\lambda_s}{c\mu_s} \quad (4.20)$$

The probability that all channels are available is:

$$P_0 = \left\{ \left[ \sum_{n=0}^{c-1} \frac{\left(\frac{\lambda_s}{\mu_s}\right)^n}{n!} \right] + \left[ \left(\frac{\lambda_s}{\mu_s}\right)^c \left(\frac{1}{c!}\right) \left(\frac{c\mu_s}{c\mu_s - \lambda_s}\right) \right] \right\}^{-1} \quad (4.21)$$

And the probability that no channels are available is:

$$P_L = \frac{(c\rho_s)^c P_0}{c!(1 - \rho_s)} \quad (4.22)$$

Where:

- $\lambda_s$  is the arrival rate at spectrum manager
- $\mu_s$  is the service rate of users who are using channels
- $c$  is the total number of channels

We are interested in finding the ratio of time spent utilizing the channels constructively versus the total amount of time spent in both obtaining permission for the channel as well as the actual amount of time spent using them. This will allow us to see what amount of efficiency is given up due to losses associated with the messaging delays.

Let  $T^A$  be the delay calculated in (4.16) of secondary user arrival,  $T^B$  be the delay calculated in (4.18) when channels are unavailable upon secondary user arrival, the total number of channels  $c$ , and  $(1-P_L)$  be the probability that at least one channel is available.

The average amount of time spent messaging depends on the scenario. Let us assume that the only messaging delays occur during secondary user subscription. There are two scenarios to consider here:

- 1) No channels available. Any arriving secondary users must wait. Assuming that these users wait an unlimited amount of time, they will eventually access a channel with a delay associated with 4.2.2 User Departure scenario
- 2) Channels are available and they may start using channels once the messaging phase has been completed. They experience delay associated with scenario 4.2.1 Secondary user arrival scenario

The average delay exhibited by these two scenarios would include the probability that no channels are available and the delay that scenario represents  $(P_L)T^B$  and the probability that channels are available and the delay that scenario represents  $(1 - P_L)T^A$ . The amount of time spent using the channel is the service time at the spectrum manager  $\frac{1}{\mu_s}$ . The amount of time spent using channels including the delays would be  $(P_L)T^B + (1 - P_L)T^A + \frac{1}{\mu_s}$

Channel efficiency is calculated to be the long-run average time spent using channels by a user (service time) over the total time spent messaging and using a channel including all delays.

$$\frac{\mu_s^{-1}}{(P_L)T^B + (1 - P_L)T^A + \frac{1}{\mu_s}} \quad (4.23)$$

The service rate that the users use channels at is bounded by the arrival rate of users as well as the total number of available channels. The utilization factor,  $\rho$ , cannot exceed one or else the queue is no longer in statistical equilibrium. The requirement to stay in a steady state environment is that  $c > \lambda/\mu$ .

Using the analysis we have discussed in this chapter, we will now perform our simulation and compare them with our analytical results derived here. This will include the delays experienced in each scenario as well as the channel efficiency under different circumstances.

#### **4.4 Motivation for Testing**

Cognitive Radios are envisioned to provide access to unused parts of the licensed spectrums. The goal is spectrum efficiency. Maximizing the utilization of spectrum is of high importance to meet this goal. Our main concern is that the overhead associated with SIP message signaling will negatively impact the spectrum utilization. This is due to the fact that delays discussed in the previous chapter will contribute to the overhead. There is the possibility that in some scenarios the delay associated with this system would be considered inefficient with respect to channel utilization.

We are interested in determining the overhead associated with message delays and how they contribute to channel utilization. Additionally, we look at user behavior in several scenarios that will allow us to determine which cases our system is better suited for. We also provide results for the gains that can be accomplished using our auctioning

mechanisms to give tangible incentive to licensed users to share information with this system.

The remainder of this chapter will provide an overview of the simulation model. The simulation process, results pertaining to the delay in several scenarios, channel efficiency, as well as auctioning gains.

## **4.5 Simulation Model**

Simulation results were obtained using OPNET Modeler. OPNET provides us with the tools needed to build our system model in a trusted simulation environment. At the Node level, we create three distinct nodes that will represent our entities of spectrum manager, primary users, and secondary users.

**Spectrum Manager:** This node receives incoming packets and process them based on the type of SIP message they contain. Events that trigger an action from the spectrum manager are all based on receiving packets. The server maintains three lists, one that tracks the subscribers, one for the primary users currently using channels, and another for primary users. The subscriber list acts as the waiting queue when there are not enough channels to service a particular secondary user. This list contains the users ID, the number of channels they request, as well as their bid. The other lists contain information on the channels they are using for their respective user types.

**Secondary User:** The secondary user node will generate Subscribe packets according to a Poisson distribution. The average can be adjusted. Generated subscribe packets are

given a random unique number that will serve as its ID, a random number between 0 and 1 following uniform distribution to be their bid, and number for the amount of channels requested. A list will keep track of the IDs used and times they successfully started using channels. A timer will determine when a particular ID will send a terminate packet and leave based on Exponential distribution with average configurable.

**Primary User:** Primary user node behaves similar to the secondary user one. Instead of Subscribe packets, it will generate Publish packets. The Publish packet contains information on user ID, and the channels they are planning on using. Primary users do not need further permission to use channels. They are assumed to begin after publishing this information.

**Network Entities:** The number of network entities stationed between users and spectrum management servers may vary depending upon the scenario. In our case, we combine the entities of the access network together into one component. SIP proxies are setup so that the transmission path is always the same. Messages arriving at these nodes experience a queuing delay before moving forward.

Figure 4.8 shows the simulated entities at the node level. Both primary and secondary user nodes act as sources of unlimited users with fixed arrival rates following Poisson distributions.

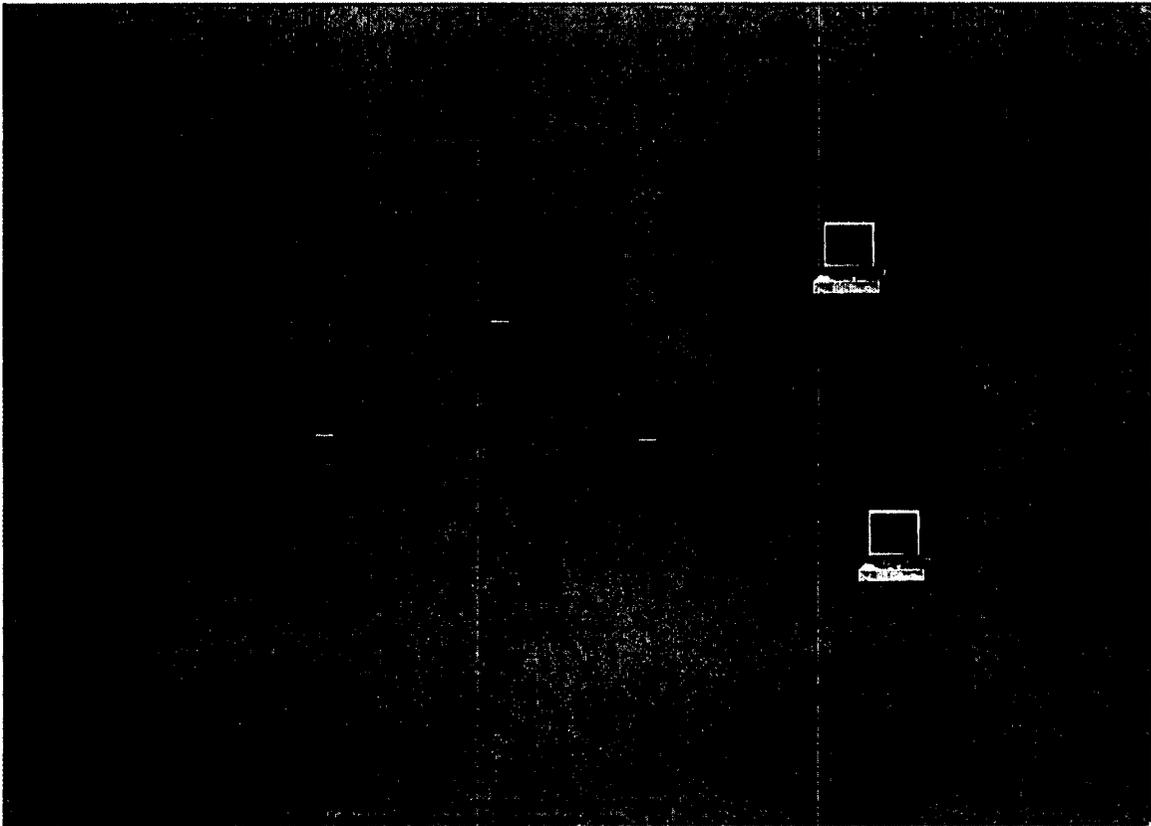
Link properties including errors, collision, and delay can be configured. The server can also be setup within a different subnet that is reachable over an IP cloud to simulate the spectrum manager being located in the internet. In this scenario, the LTE E-UTRAN interface is located at one subnet while the server is placed in another.

Secondary and primary users are connected to a router which will serve as the access network while the other SIP proxies forward packets to their intended destinations. The spectrum manager is located at the opposite end of the network from the two user types.

Assumptions used in the analytical estimates and simulation include the following parameters.

- We assume enough channels are available for the new secondary user and they will not need to wait due to channel unavailability. In the messaging delay part of the analysis, there will be no delays due to channel unavailability when new secondary users subscribe. We first examine the effects of delays and later apply additional waiting delays to our simulation on channel efficiency.
- With a fast enough processor speeds, background processes at the server are assumed negligibly small compared to the other delays experienced. Since the only load on the server of significance is SIP messages being served, we use an M/M/1 queuing model to determine the queuing/processing delay at the destination node. This assumption leads to lower overall delays. If these assumptions are not true, and there are other non-SIP messages also being serviced, or processing times are large, the additional delay will be compounded by the number of messages sent.
- The number of network entities between the user and the edge of the LTE network and the internet include: the access network, the SIP network itself contains a variable number of SIP proxy servers between the access network and the spectrum manager.

- Messages arriving at the spectrum manager may have other Non-SIP messages with higher priorities. Thus, we use a non-preemptive M/G/1 queuing model as discussed in section 4.1.3 at this node.
- Service rate at the proxy servers and the spectrum manager are the same to simplify calculations. Service time values are varied along with the message arrival rates in order to change server load.
- Frame duration and interframe time are set as per LTE standards. Frame duration has frames,  $D$ , at 10ms. Interframe times,  $\tau$ , have values of 10ms.
- Channel capacity affects the overall delay of messages. We give results for varying channel capacities including those from lower values to others that are considered broadband wireless.
- Delays caused by internet traffic are hard to predict. They may depend on a large number of factors including number of forwards, congestion, etc. In our analysis, we use a constant value for our internet delay to simplify calculations. Internet delay,  $I$ , compounds with the number of messages sent between users and the spectrum manager as well as the number of proxies between them. For our simulation we use a constant value of 100ms for simplification purposes. A larger or smaller internet delay changes the overall delay linearly with respect to the number of messages.



**Figure 4.8: Simulation Entities**

Results for messaging delays and overall delays for the secondary user arrival scenario are collected at each user node. Counters begin upon sending of SUBSCRIBE message and end when the expected reply is received.

Results for primary user arrival scenario are collected at the primary user node. Counters begin when the PUBLISH message is sent and end when they receive the expected reply from the spectrum manager.

The spectrum manager itself also keeps results for message delay of the handoff scenario. Counters measure the time when a NOTIFY is sent to secondary user until the reply is received from the secondary user.

Delays of these three scenarios are recorded for later analysis. The analysis of data and statistics are obtained using MATLAB. Data files of the delays are entered into MATLAB. Some basic M-files are used to obtain the averages as well as the confidence intervals of each set of data.

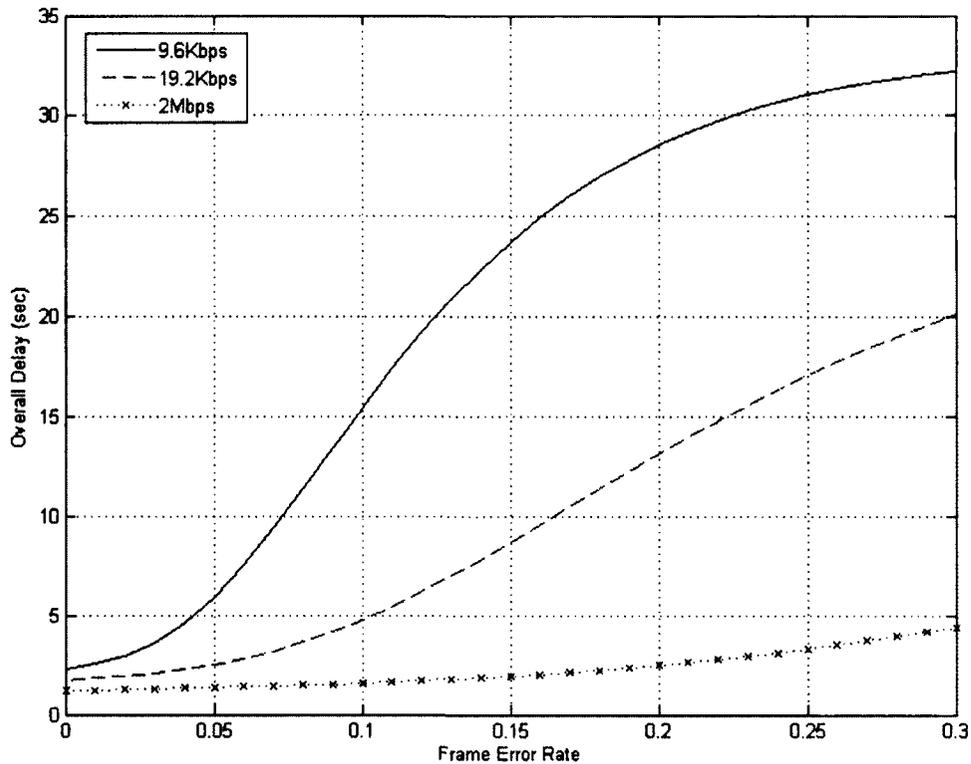
The simulation parameters such as FER, number of SIP proxies, and network loads are then changed to obtain different results under different scenarios. The same process of obtaining the statistics is used afterwards. The results of the simulation are seen in the next section.

## **4.6 Simulation Results for Messaging Delays**

The following delay analysis is based on the analytical model of scenarios proposed in section 4.2. Assumptions used are described in that section unless otherwise stated specifically for each scenario.

### **4.6.1 Secondary User Arrival**

Fig 4.9 shows analytical results for overall delay of TCP session setup between secondary user and server spectrum manager and the NOTIFY message that grants permission for channel use. Results assume 10 percent server loads, two SIP proxies, no waiting delays, and varying channel capacities and frame error rates. Results show that as frame error rate increases, the overall delay also increases. With lower channel capacities, delay can become unfeasible.



**Figure 4.9: Delay for varying FER of secondary user SUBSCRIBE**

A dramatic increase in delay occurs when frame error rate is high increasing the probability of message retransmission. This problem is more apparent in lower channel capacities where there are a larger number of frames are required to deliver SIP messages. The amount of frames has a large influence on the retransmission delay. The problem is less prominent with larger channel capacities where a message can be transmitted using a single frame. Although there are still increases in delay given a high frame error rate, the overall delay remains low enough to be considered efficient in some channel allocation scenarios.

#### 4.6.2 Primary User Arrival

In the scenario of primary user arrival, the PUBLISH message sent from primary users to the spectrum manager do not necessarily travel along the same transmission path, but they experience the same types of delays that secondary users who send SUBSCRIBE messages to the spectrum broker server experience. Namely, TCP transmission delays, internet delays, queuing delays of the user, network entities and spectrum manager.

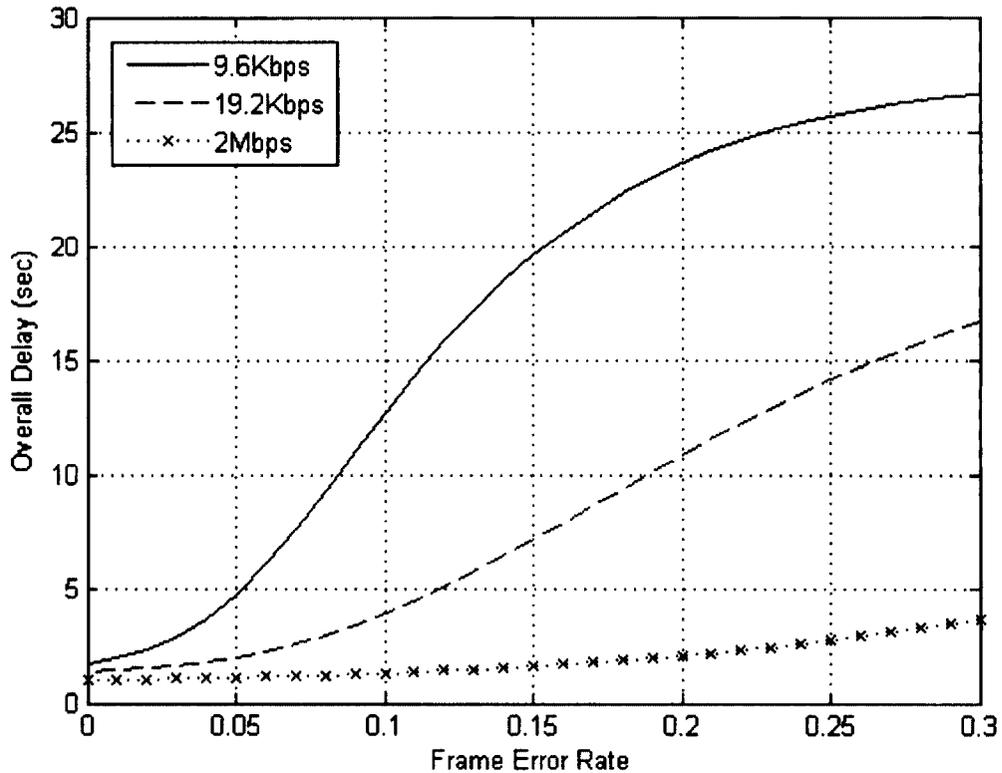


Figure 4.10: Delay of various FER for primary user PUBLISH

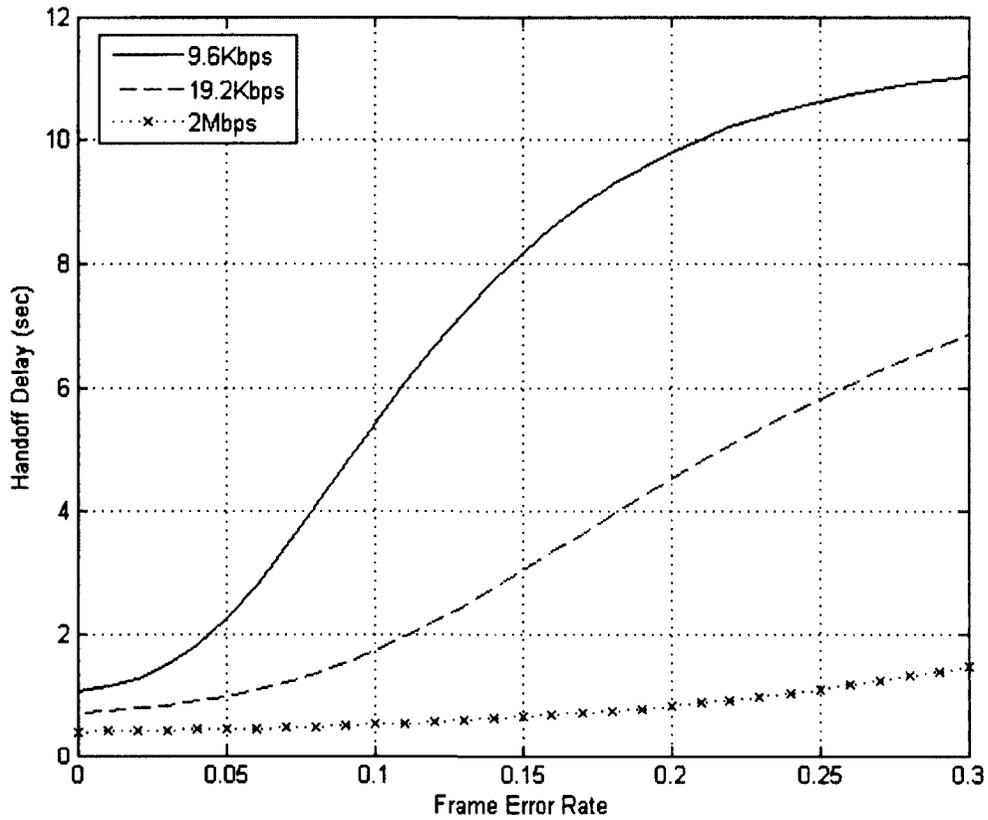
Fig 4.10 shows the delay experienced by primary users for sending PUBLISH message. The graph shows delay values between the time primary users establishes a TCP session

with the spectrum broker server and the 200OK response. Assumptions are identical with the scenario with secondary user arrival.

Results show that delays are similar with the previous scenario with secondary user's arrivals. Since the delays are the same for both scenarios, the graph shows they follow the same curve. The difference is there are less messages sent between the two network entities. The delay itself is slightly lower than the previous scenario.

#### **4.6.3 Secondary User Pre-empted and Handoff**

In the scenario where secondary user is pre-empted due to the arrival of primary user, secondary users are notified by the spectrum broker server of this arrival. There are two separate cases that can happen as described in the previous chapter. If there are no channels available, secondary users must wait until channels become available for them. However, if there are channels available, the spectrum manager can suggest other channels that suit their needs. In each of these cases, a NOTIFY message is sent to the secondary users. Since the message sequencing is the same for both cases, we can assume that the delay is also the same.

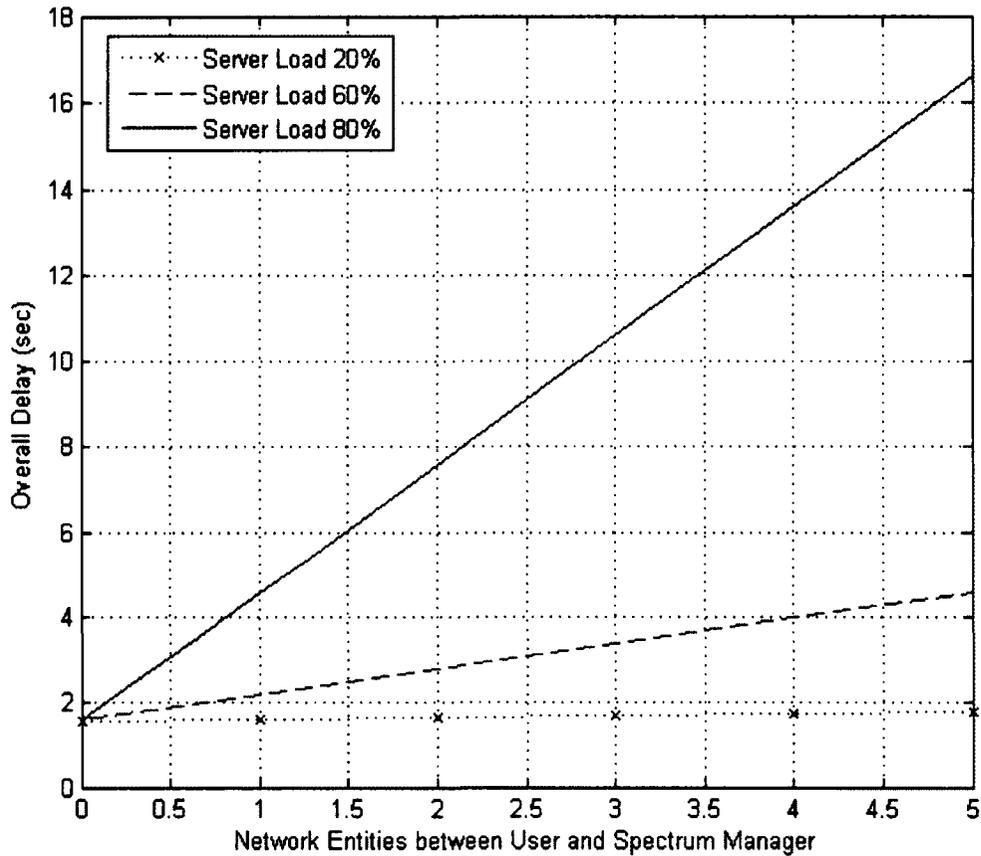


**Figure 4.11: Handoff delay for secondary users**

Fig. 4.11 shows the handoff delay experienced by secondary users when they are preempted by the arrival of a primary user. Results were obtained using the same assumptions as previous simulations. This delay measure the time between primary user send the PUBLISH message to the successful transmission of the NOTIFY message to the secondary user.

Results show that delays follow a similar pattern to previous analysis of messaging delays. The handoff delay is lower given that secondary user has already established a TCP session with the spectrum manager. Although larger frame error rates still

diminishes the feasibility in these scenarios, a larger channel capacity produces delays that are more acceptable.



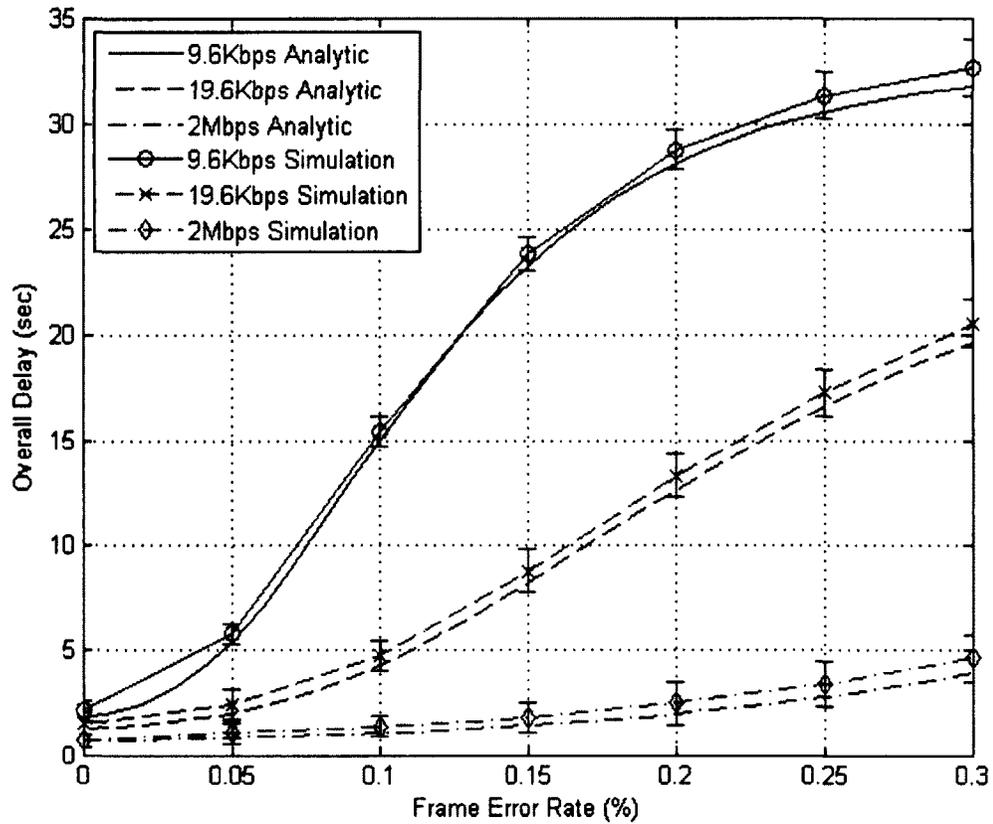
**Figure 4.12: Delay with varying network entities and loads**

Fig 4.12 shows the overall delay with varying server loads and network entities. We keep a constant frame error rate of 10 percent and channel capacity of 2Mbps. The proportion of SIP messages versus non-SIP messages served by each network entity has been kept constant. Results show that server loads can have a dramatic impact on overall delay. In the scenario with a large number of entities between a user and the spectrum manager, a high load on each server can make this approach unfeasible by user

expectations. However, when server loads are low, the number of network entities between the user and the spectrum manager does not affect the overall delay greatly.

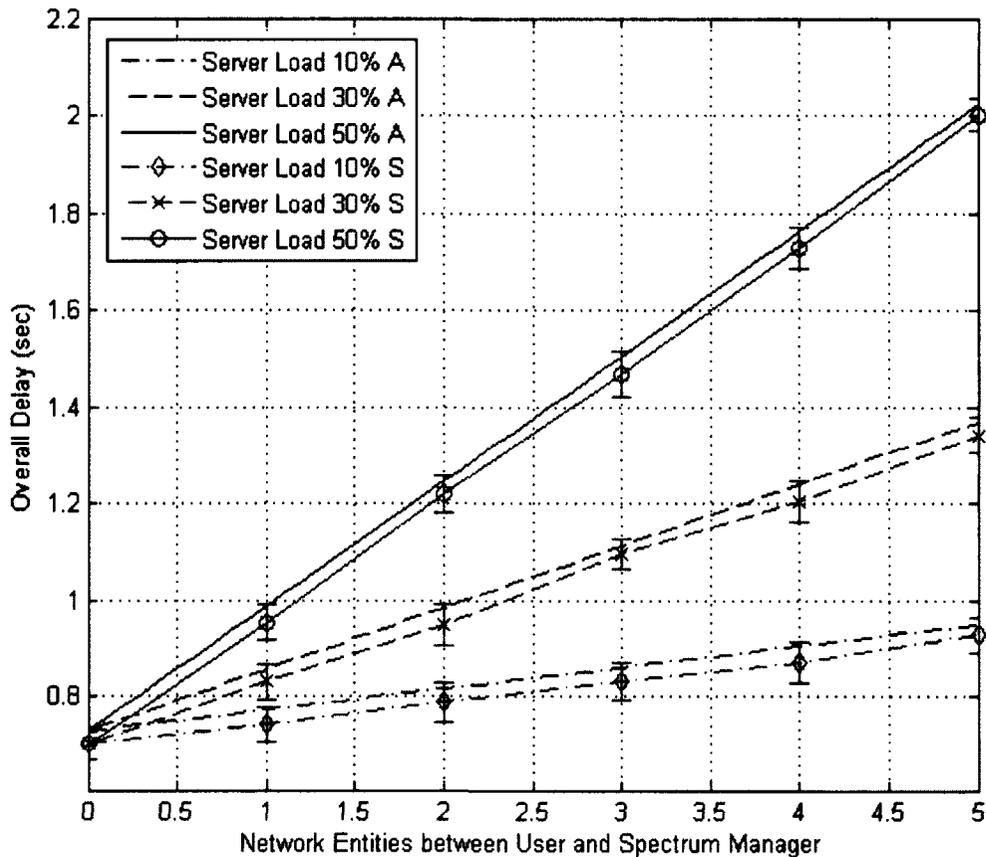
The results obtained from the analytical model will be used as part of our simulation to determine the overall efficiencies of channel usage under different scenarios. The next section will give a description of our simulation model, along with the metric of efficiency we will quantify and compare with existing channel allocation techniques.

The following are simulation results obtained for different scenarios compared with the analytical ones obtained. Each result is obtained using the messaging scenarios outlined in the previous chapter. This changes certain values of parameters which may affect the overall delay. Parameters are the same as those previously mentioned, unless otherwise stated. Error bars show confidence intervals at 95 percent.



**Figure 4.13: Simulation delays for secondary user arrival**

Figure 4.13 shows simulation results compared with the analytical ones for secondary user arrival scenario. Simulation results for secondary user arrival using SUBSCRIBE message follows closely with the predicted analytical results. In this simulation, we make the same assumptions as those previously stated. Spectrum manager loads are 10 percent, two SIP proxies, varying channel capacities and frame error rates.



**Figure 4.14: Simulation Delay with Varying Network Entities and Loads**

Figure 4.14 shows simulation results of overall delay with the analytical findings. In this scenario, we assume an ideal link with no frame errors and no retransmissions needed over 2Mbps link. Here we look at the secondary user arrival scenario as above with changing number of network entities and loads. The number of SIP proxies and their loads are varied in order to see the effects on overall delay. Other assumptions made are the same as previously mentioned.

Results for both simulation show that they follow closely with the predicted analytical values. There is a slight discrepancy between the two values. This is due to a precision error within our simulation software. OPNET allows manipulation of packet delays up to

two decimal places and further ones are truncated. The result is a loss of significant figures in the measurement that carries over in the calculations.

The simulation results themselves shows favorable results under ideal conditions when channel capacity is large and there are no retransmissions. However, given that frame error or lost packets are unavoidable, there will likely be some amount of retransmissions. The following shows results given a percentage of dropped packets.

#### **4.7 Simulation Results for Channel Efficiency**

For our proposed system, the licensed band to be accessed has an operational bandwidth of 60MHz divided into 40 channels of 1.4MHz each for a total bandwidth usage of 56MHz.

Our simulation uses a single cell with one base station serving a number of secondary users under each scenario. A spectrum manager is placed close to the base station to manage channel allocation. Users used channels for intervals following exponential distributions with varying mean times. The arrival rates for users also followed exponential with varying mean times. For this simulation it is assumed that all secondary users are interested in channel use no matter how long they wait. Simulation results are obtained from an average observable steady state environment.

### 4.7.1 Analytical and Simulation Results

Figure 4.15 shows the results from both the analytical analysis and the simulation. Simulation results are similar to the analytical ones. Channel efficiency is slightly lower than those predicated since the overall delays experienced by secondary users have been simulated to be slightly higher. However, the overall curves are the same.

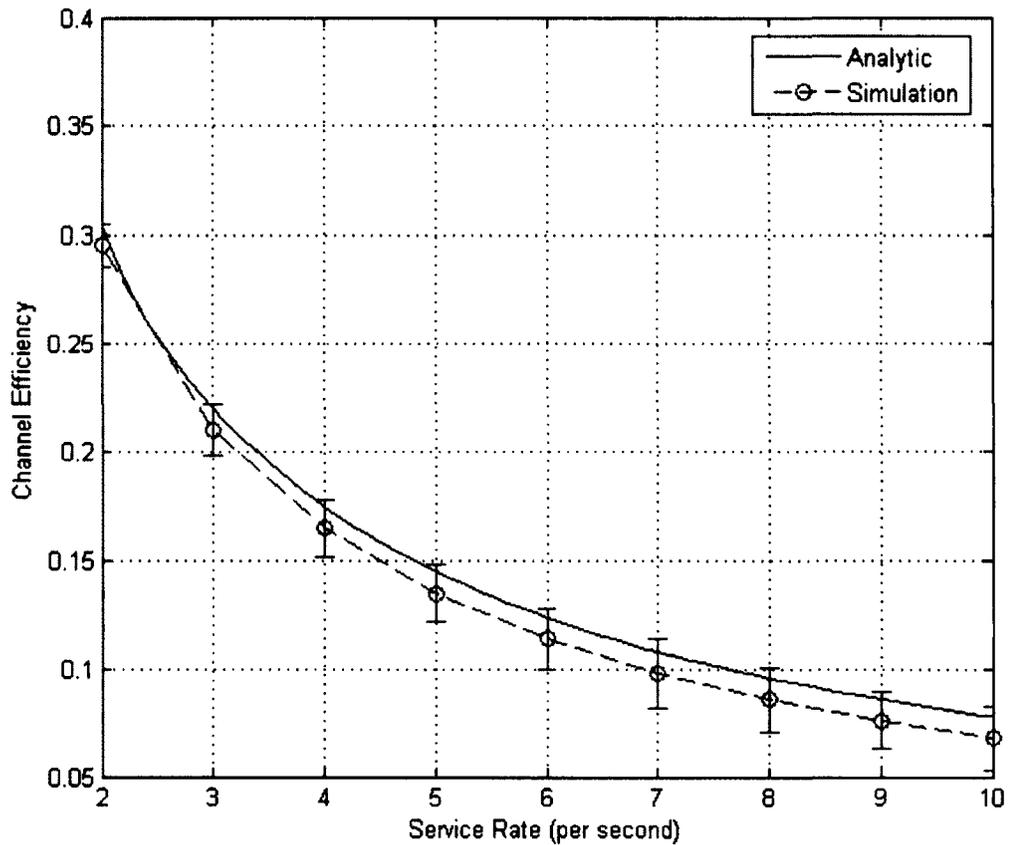


Figure 4.15: Channel Efficiency with Arrival Rate 60/sec

In this scenario, we assume that arrival rates are very high. With 60 users arriving every second, service rates must also be high in order to keep a balance. If they are not, the system would no longer be in equilibrium and arriving users would not be able to get a

channel since waiting times would be infinite. This scenario shows channel utilization percentages in the event that users are arriving fast and leaving fast. They are not using channels for any extended period of time. Since delays occur during the arrivals of new users, the more users there are, the more pronounced the delay becomes. With small service times per user, the time associated with messaging delay for each user may become even higher than the time spent using channels. Figure 4.15 shows that the channel utilization percentage is low when the arrival rates of users are high. With increasing service rates, the channel utilization percentage lowers. This is expected since with less service time dedicated to each user, the ratio of time spent messaging to channel usage is also lower.

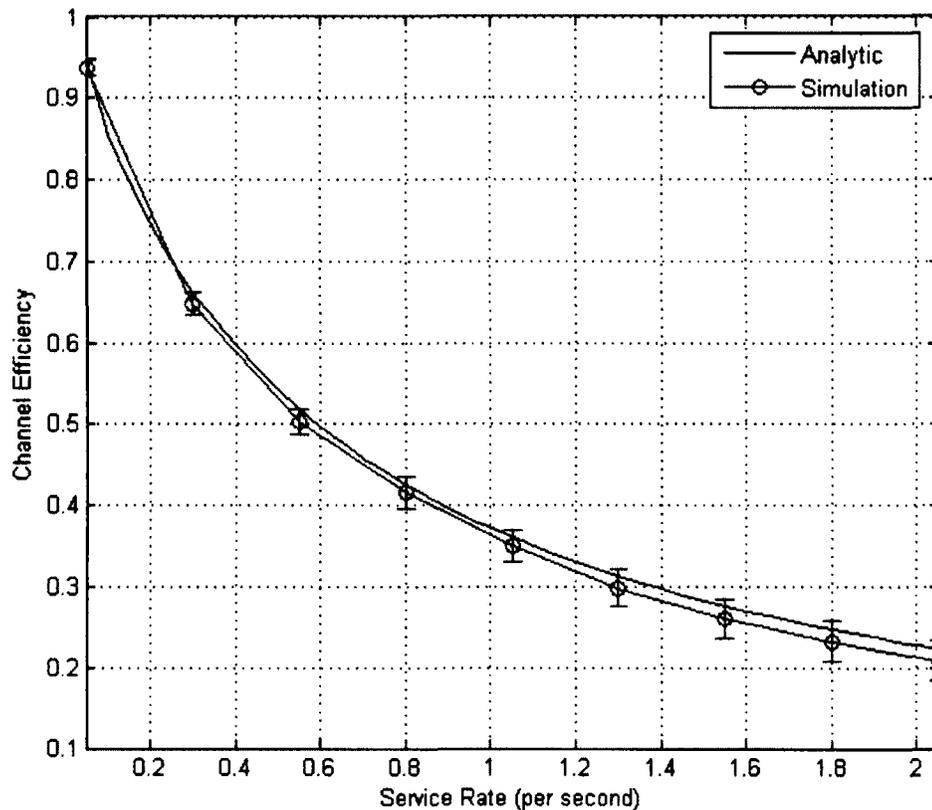


Figure 4.16: Channel Efficiency with Arrival Rates 2/sec

Figure 4.16 shows results from both the analytical and simulations for arrival rates averaging 2 per second. As in the previous scenario, channel efficiency is slightly lower in the simulation than the analysis. Overall, the curves follow the same shape.

With lower arrival rates of 2 users per second, the service rates are allowed to be smaller. With a larger amount of time allowed for channel usage, the channel utilization percentages are much higher. In this scenario, the service times can be longer and so the ratio of time spent using channels versus the total time including delays is larger. As the service rates increase, the amount of time spent using channels decreases and so the channel utilization percentage also decreases.

Primary users can not use spectrum immediately if they choose to use a channel that is currently occupied by a secondary user. In the scenario that was mentioned previously in section 4.2.3, an amount of time is lost due to message signaling from the spectrum manager to the secondary user to vacate the channel. The amount of time this consumes is the delay in the handoff scenario. Figure 4.11 shows the possible amount of time due to the handoff delay. The probability that an arriving primary user pre-empts an incumbent secondary user depends on several factors including the total number of available channels, and how heavily they are being used. Section 4.2.3 gives some estimates on the likelihood of this scenario. Under the best conditions it may be in the range of several hundred milliseconds.

If we assume that the worst case scenario is the secondary user occupying the channel does not receive the NOTIFY message due to transmission failure, they will be forced to

time out before the primary user can use it. This may take over 10 seconds. The probability of the worst case happening also depends on the handoff scenario occurring as well as assuming all retransmission of the message fails. With frame error rates,  $p$ , at 30% and assuming a single frame for messages a corresponding retransmission rate,  $q$ , is 51%. The probability of all seven message attempts failing,  $P_f$ , would be 0.897%. If we assume payload sizes are large and transmission speeds are low, the number of frames could be as high as 37. Under these conditions, the probability of seven message attempts failing can be as high as 99% and a timeout would occur after 10 seconds.

Figures 4.15 and 4.16 do not take into account the amount of time lost due to the above two scenarios. Therefore, it is more optimistic than what would normally be expected. However, the probability of these scenarios occurring, specifically failing to receive messages, was determined in section 4.2.3 to be quite low. The curves in these two figures for channel utilization are average times, so considering these two scenarios would place results lower than what was obtained, but not by much.

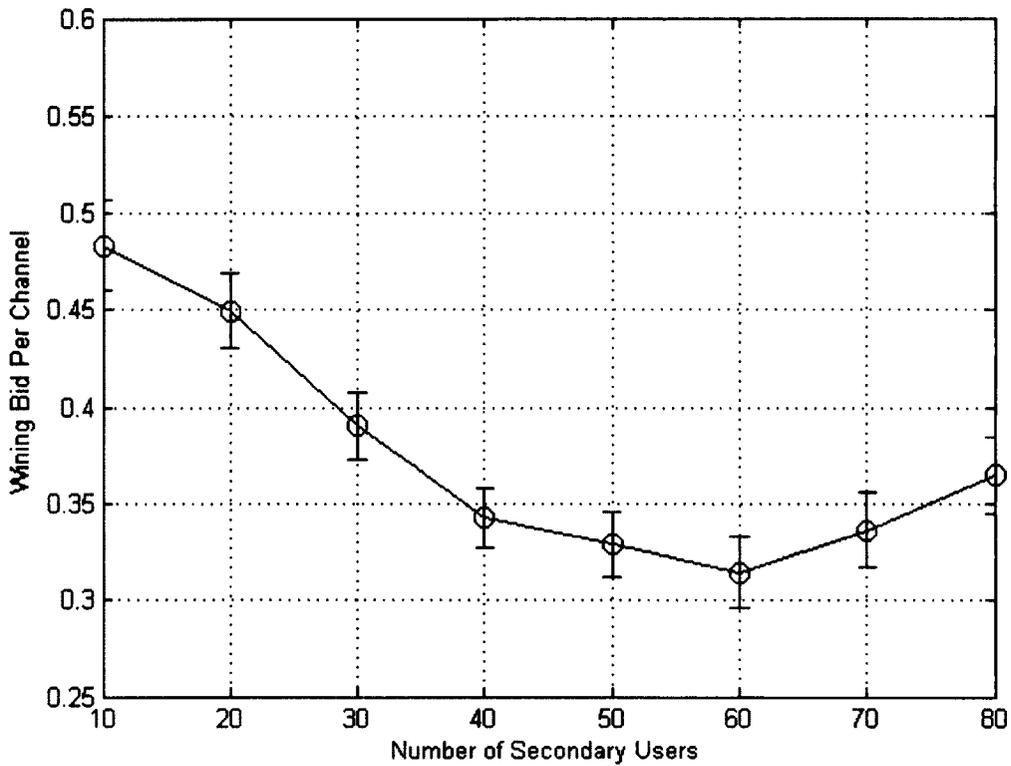
#### **4.8 Simulation Results for Auctioning Mechanism**

The following results show how mean bids and overall gain change as the average number of users increases. In this simulation a set number of secondary users compete with each other over a finite number of available channels. Secondary users are in either active or inactive states. These users follow exponential distribution for both arrival rates and spectrum usage duration. The following is a table of parameters used for this simulation.

**Table 4.3: Auctioning Simulation Parameters**

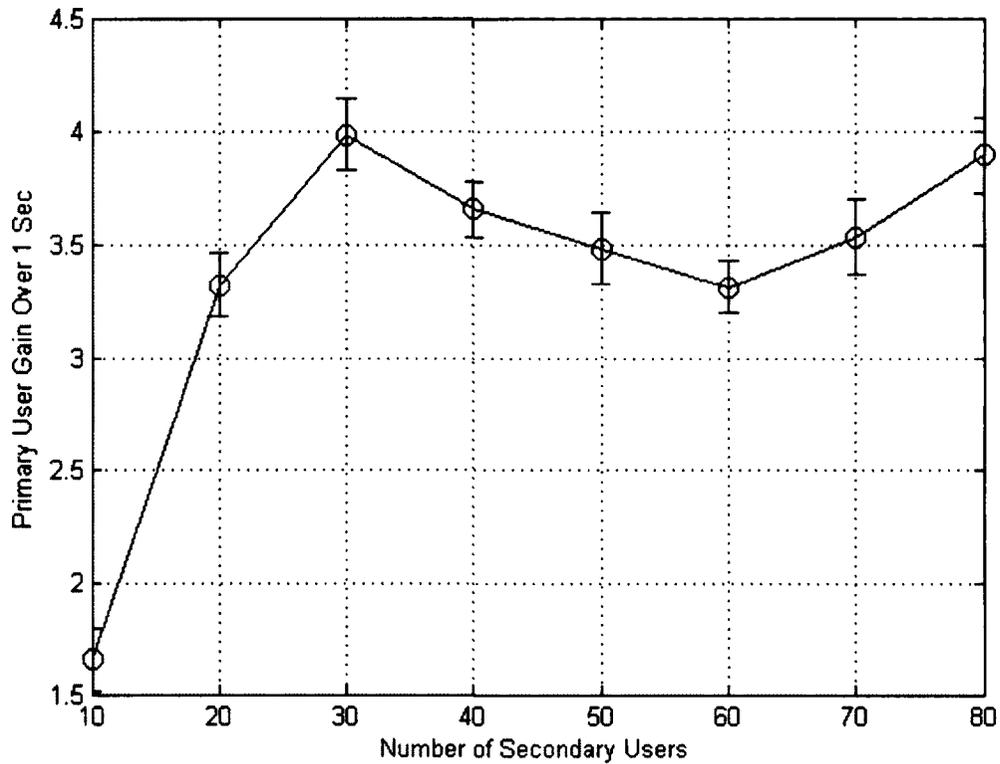
<b>Parameters</b>	<b>Settings</b>
<b>Number of Secondary Users</b>	varies
<b>Mean Arrival Rate</b>	1 hour

Secondary users bid uniform distribution on a scale of 0 to 1 (0 lowest, 1 highest) for channels. Winning bidders are selected using the second-place sealed bid auctioning scheme as described in section 4.3 of the previous chapter. We use the second-place sealed bid auctioning scheme since it gives an incentive to not bid lower in order to receive a lower winning bid. The dominate strategy for users in this auctioning scheme is to bid truthfully on how much they value the channel [56].



**Figure 4.17: Average Winning Bid**

Figure 4.17 show the average winning bid as the potential number of secondary users increases. Results show that as when numbers of users are low, the winning bid is close to the 0.5 point. However, as more users are competing, the average winning bid decreases. This is due to the fact that without competition, users pay their bidding price which averages to 0.5. When additional users are present, the second-place sealed bid auctioning scheme is more likely to pick a second-place price that is lower than average resulting in a decrease in the average winning bid. When the number of users starts exceeding the total number of channels, competition for channels increases and average winning bids start to increase again.



**Figure 4.18: Revenue Gain over 1 second**

Figure 4.18 shows the average gain for primary users every second. The gain for primary users is calculated as the amount of revenue generated per second assuming the bid price is paid per second for each channel in use. The results echo the ones in the average winning bids. The amount of revenue generated increases until the winning price paid per channel starts decreasing. It starts to increase again as more users increase the competition when channels become scarce.

The auctioning mechanism used in this simulation is one example used. There are other auctioning methods such as first-place sealed bid that may yield better results in certain situations.

## 4.9 Comparison of Results

Results from the simulations show that the amount of time taken up by the delay can be quite significant in some cases. Circumstances that contribute to the overall amount of delay include the air link capacity between users and the base station part of the access network, the condition of the air link may contribute to frame error and retransmissions, the amount of proxy SIP servers messages must traverse to reach the spectrum manager, and the loads these network entities have. Other factors may also contribute to the amount of delay, but these are the most significant ones. Channel efficiency largely depends on this delay as well as the average amount of time users use channels.

As discussed in section 3.4, traditional cognitive radios have a trade-off between detection probabilities and amount of time spent sensing. In [9], there is shown a good-point of channel efficiency given the number of detection symbols as well as the probability of detection desired. Results show that better channel efficiency can be obtained with lower probabilities of detection up to a maximum of 65% with a probability of detection of 90%. When higher levels of detection are wanted, at 99.9% probability of detection, efficiency drops dramatically to a maximum of 29%.

Our method has clear advantages in terms of channel efficiency under certain circumstances. One advantage is that we operate at 100% detection probability given that all users report their usage, so there is no need for detection and no chance of collisions between users. However, under certain conditions, our method has disadvantages. When arrival rates of users are high as shown in figure 4.16, the efficiency can be low. With high arrival rates, the service times need to be quick. This contributes to the overall low

efficiency since the ratio of time spent using the channel to the delays is much lower. Under better circumstances, when arrival rates are lower, and users spend more time using channels, the overall efficiency is much greater as shown in figure 4.17. In such a scenario, the time spent obtaining permission to use a channel from the spectrum manager is worth it.

There have been various attempts to improve efficiency and probability of detection discussed in chapter 2. One prominent method is the use of cooperative sensing where multiple users pool their resources by sensing smaller portions of the spectrum. In [37], sensing times are given with respect to the number of cooperative users and probability of detection at 99%. Their results show that sensing times can be much lower, on the scale of milliseconds. However, results only show the delay due to the sensing part. Other delays such as information dissemination, heavier processing loads, and competition for available channels are not addressed. Analysis here is assumed to be under ideal channel conditions with error free communication between users. These factors would contribute to the total time needed to access channels.

Our analysis shows all potential delays including channel error causing retransmission. Delays are heavily influenced by the frame error rate and channel capacity as shown in figure 4.13. Under unfavourable conditions, delay can be as high as 30 seconds which may make our method unfeasible. However, when frame errors are low, or non-existent, the delay can be a very reasonable value that is comparable with those of cooperative sensing under similar circumstances.

Our result shows several scenarios in which delay can be overall large. In particular, when frame error rate in the air link is high, and when loads on SIP proxy servers are high. However, under ideal circumstances, the delay has shown to be comparable to other more traditional cognitive radio methods.

## **Chapter 5**

### **Conclusion and Future Work**

The ongoing trend of more wireless devices is increasing rapidly. Users are expecting their wireless devices to provide more in terms of applications and services as well as convenience of always being connected. The wireless spectrum is limited in its capacity and must be used more efficiently if more wireless devices are to be made. Cognitive radios is one of the efforts put forth to make devices smarter and more adaptable in order to meet the coming needs of efficient wireless spectrum usage. However, cognitive radios are still in development with many issues yet to be solved before their deployment and use can be further realized. The need for spectrum sensing in order to detect unused bands has been a problem that has yet to be addressed in an elegant solution. Also, the coordination among these cognitive radio users and incumbent users has led to additional complexities in order to manage their spectrum usage.

In this thesis, we have provided a solution of a centralized cognitive radio network that utilizes the primary user's shared spectrum usage information. This allows us to eliminate the need for spectrum scanning and avoid usage collisions. In addition we explore the use of higher layer protocols, namely SIP, to facilitate communication and channel distribution between licensed users and cognitive radio users. We present our contributions through various event scenarios using messaging diagrams of SIP messages, analytical models of the messaging delays, along with simulations using OPNET to confirm our analysis.

We have demonstrated through an example using 3GPP's Long Term Evolution network, that SIP messaging can be used as an overlay to existing network architecture in order to facilitate communication between different users for the purposes of spectrum management and channel distribution in a cognitive radio network. Communication using SIP messages also allows us the exchange of information for the purposes of billing. We have proposed an auctioning mechanism in order to provide an incentive for primary users to share their spectrum usage information. Results show the amount and gain under different scenarios.

In our simulations, we devise and implement a model to simulate the behaviour of SIP messaging over network architecture with the same parameters as our analysis in order to observe the effects of delay on channel efficiency. During the simulation, various parameters are changed in order to simulate different scenarios. The results collected from the simulation follows closely with our predicted analysis. They show that under certain scenarios, the amount of delay experienced is small enough that it does not contribute highly to spectrum downtime.

The benefits derived from our solution solves several open issues related to cognitive radios including reliability of spectrum sensing, sharing spectrum usage information, and reducing device complexity. Our method eliminates the need for spectrum sensing and in turn removes the possibility of false or missed detection of primary users. We also centralize spectrum management duties in order to remove the need for more complex computations and devices associated with cooperative sensing.

Future work can be undertaken following several directions. In our model, we have taken liberties in some of our assumptions in order to simplify the analysis. A more accurate model of this system can be made. Users are assumed to wait as long as needed to obtain channels. A more accurate model may have users renege after a certain period of time to better model human behaviour. This may change the overall efficiency of channel usage as well as users overall satisfaction with the system. We have also assumed that service rates for all entities are the same to simplify analysis. This assumption may not always be true. Propagation delays of messages are set to be constant, but may not always be the same. This may change the overall delay and efficiency depending on how large or small it is.

Finally, the current model only takes into consideration one cell. Another future direction is to look at scenarios with multiple adjacent cells. Handoff scenarios when users move from one cell to the next as well as spectrum management across multiple cells have yet to be explored.

## References

- [1] Jianfeng Wang; Ghosh, M.; Challapali, K.; , "Emerging cognitive radio applications: A survey," *Communications Magazine, IEEE* , vol.49, no.3, pp.74-81, March 2011
- [2] Mitola, J., III; Maguire, G.Q., Jr.; , "Cognitive radio: making software radios more personal," *Personal Communications, IEEE* , vol.6, no.4, pp.13-18, Aug 1999
- [3] Shunqing Zhang; Tianyu Wu; Lau, V.K.N.; , "A low-overhead energy detection based cooperative sensing protocol for cognitive radio systems," *Wireless Communications, IEEE Transactions on* , vol.8, no.11, pp.5575-5581, November 2009
- [4] Yi Zheng; Xianzhong Xie; Lili Yang; , "Cooperative spectrum sensing based on blind source separation for cognitive radio," *Future Information Networks, 2009. ICFIN 2009. First International Conference on* , vol., no., pp.398-402, 14-17 Oct. 2009
- [5] Yunxue Liu; Dongfeng Yuan; Mingyan Jiang; Yanbo Ma; Jiali Xu; , "Novel study on Cooperative Spectrum Sensing for Cognitive Radio," *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM '08. 4th International Conference on* , vol., no., pp.1-5, 12-14 Oct. 2008
- [6] Tao Jiang; Grace, D.; Yiming Liu; , "Cognitive radio spectrum sharing schemes with reduced spectrum sensing requirements," *Cognitive Radio and Software Defined Radios: Technologies and Techniques, 2008 IET Seminar on* , vol., no., pp.1-5, 18-18 Sept. 2008
- [7] Zhi Quan; Shuguang Cui; Sayed, A.H.; Poor, H.V.; , "Wideband Spectrum Sensing in Cognitive Radio Networks," *Communications, 2008. ICC '08. IEEE International Conference on* , vol., no., pp.901-906, 19-23 May 2008
- [8] Peng Wang; Limin Xiao; Shidong Zhou; Jing Wang; , "Optimization of Detection Time for Channel Efficiency in Cognitive Radio Systems," *Wireless Communications and Networking Conference, 2007. WCNC 2007. IEEE* , vol., no., pp.111-115, 11-15 March 2007

- [9] Fehske, A.; Gaeddert, J.; Reed, J.H.; , "A new approach to signal classification using spectral correlation and neural networks," *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on* , vol., no., pp.144-150, 8-11 Nov. 2005
- [10] Biswas, A.R.; Aysal, T.C.; Kandeepan, S.; Kliazovich, D.; Piesiewicz, R.; , "Cooperative Shared Spectrum Sensing for Dynamic Cognitive Radio Networks," *Communications, 2009. ICC '09. IEEE International Conference on* , vol., no., pp.1-5, 14-18 June 2009
- [11] Perez-Romero, J.; Salient, O.; Agusti, R.; Giupponi, L.; , "A Novel On-Demand Cognitive Pilot Channel Enabling Dynamic Spectrum Allocation," *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on* , vol., no., pp.46-54, 17-20 April 2007
- [12] Beibei Wang; Liu, K.J.R.; , "Advances in Cognitive Radio Networks: A Survey," *Selected Topics in Signal Processing, IEEE Journal of* , vol.5, no.1, pp.5-23, Feb. 2011
- [13] Roy, S.D.; Mondal, S.; Kundu, S.; , "A new algorithm for admission control of secondary users in CDMA based Cognitive Radio Network," *Computer and Communication Technology (ICCCT), 2010 International Conference on* , vol., no., pp.35-39, 17-19 Sept. 2010
- [14] Raman, C.; Yates, R.D.; Mandayam, N.B.; , "Scheduling variable rate links via a spectrum server," *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on* , vol., no., pp.110-118, 8-11 Nov. 2005
- [15] Qing Zhao; Sadler, B.M.; , "A Survey of Dynamic Spectrum Access," *Signal Processing Magazine, IEEE* , vol.24, no.3, pp.79-89, May 2007
- [16] Salami, G.; Durowoju, O.; Attar, A.; Holland, O.; Tafazolli, R.; Aghvami, H.; , "A Comparison Between the Centralized and Distributed Approaches for Spectrum Management," *Communications Surveys & Tutorials, IEEE* , vol.13, no.2, pp.274-290, Second Quarter 2011

- [17] Beibei Wang; Liu, K.; Clancy, T.C.; , "Evolutionary Game Framework for Behavior Dynamics in Cooperative Spectrum Sensing," *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE* , vol., no., pp.1-5, Nov. 30 2008-Dec. 4 2008
- [18] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation dynamic spectrum access cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, pp. 2127–2159, May 2006.
- [19] Yucek, T.; Arslan, H.; , "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys & Tutorials, IEEE* , vol.11, no.1, pp.116-130, First Quarter 2009
- [20] W. D. Horne, "Adaptive spectrum access: Using the full spectrum space," in Proc. Annu. Telecomm. Policy Res. Conf., Arlington, VA, Oct. 2003
- [21] J. Lehtomaki, J. Vartiainen, M. Juntti, and H. Saarnisaari, "Spectrum Sensing with forward methods," in Proc. IEEE Military Commun. Conf., Washington, D.C., Oct. 2006, pp. 1–7
- [22] Kyungtae Kim; Yan Xin; Rangarajan, S.; , "Energy Detection Based Spectrum Sensing for Cognitive Radio: An Experimental Study," *GLOBECOM 2010, 2010 IEEE Global Telecommunications Conference* , vol., no., pp.1-5, 6-10 Dec. 2010
- [23] Jian Chen; Gibson, A.; Zafar, J.; , "Cyclostationary spectrum detection in cognitive radios," *Cognitive Radio and Software Defined Radios: Technologies and Techniques*, 2008 IET Seminar on , vol., no., pp.1-5, 18-18 Sept. 2008
- [24] Maeda, K.; Benjebbour, A.; Asai, T.; Furuno, T.; Ohya, T.; , "Recognition Among OFDM-Based Systems Utilizing Cyclostationarity-Inducing Transmission," *New Frontiers in Dynamic Spectrum Access Networks, 2007. DySPAN 2007. 2nd IEEE International Symposium on* , vol., no., pp.516-523, 17-20 April 2007
- [25] Proakis, J.; , *Digital Communications*. New York: McGraw-Hill, 2001

- [26] A. Sahai and D. Cabric, "A tutorial on spectrum sensing: Fundamental limits and practical challenges," in *Proc. IEEE Symp. New Frontiers Dynamic Spectrum Access Netw. (DySPAN)*, Baltimore, MD, Nov. 2005
- [27] Cabric, D.; Mishra, S.M.; Brodersen, R.W.; , "Implementation issues in spectrum sensing for cognitive radios," *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on* , vol.1, no., pp. 772- 776 Vol.1, 7-10 Nov. 2004
- [28] Ganesan, G.; Ye Li; , "Cooperative Spectrum Sensing in Cognitive Radio, Part I: Two User Networks," *Wireless Communications, IEEE Transactions on* , vol.6, no.6, pp.2204-2213, June 2007
- [29] Peh, E.; Ying-Chang Liang; , "Optimization for Cooperative Sensing in Cognitive Radio Networks," *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE* , vol., no., pp.27-32, 11-15 March 2007
- [30] Chair, Z.; Varshney, P.K.; , "Optimal Data Fusion in Multiple Sensor Detection Systems," *Aerospace and Electronic Systems, IEEE Transactions on* , vol.AES-22, no.1, pp.98-101, Jan. 1986
- [31] Jiang, T.; Qu, D.; , "On minimum sensing error with spectrum sensing using counting rule in cognitive radio networks," in *Proc. 4<sup>th</sup> Annual Int. Conf. Wireless Internet (WICON'08)*, Brussels, Belgium, pp.1-9, 2008
- [32] Wei Zhang; Mallik, R.K.; Ben Letaief, K.; , "Cooperative Spectrum Sensing Optimization in Cognitive Radio Networks," *Communications, 2008. ICC '08. IEEE International Conference on* , vol., no., pp.3411-3415, 19-23 May 2008
- [33] Feng Gao; Wei Yuan; Wei Liu; Wenqing Cheng; Shu Wang; , "Pipelined Cooperative Spectrum Sensing in Cognitive Radio Networks," *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE* , vol., no., pp.1-5, 5-8 April 2009
- [34] Hwang, S.-H.; Baek, J.-H.; Dobre, O.A.; , "Spectrum sensing using multiple antenna-aided energy detectors for cognitive radio," *Electrical and Computer Engineering, 2009. CCECE '09. Canadian Conference on* , vol., no., pp.209-212, 3-6 May 2009

- [35] Subramani, S.; Armour, S.; Kaleshi, D.; Zhong Fan; , "Spectrum Scanning and Reserve Channel Methods for Link Maintenance in Cognitive Radio Systems," *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE* , vol., no., pp.1944-1948, 11-14 May 2008
- [36] Weifang Wang; , "Spectrum Sensing for Cognitive Radio," *Intelligent Information Technology Application Workshops, 2009. IITAW '09. Third International Symposium on* , vol., no., pp.410-412, 21-22 Nov. 2009
- [37] Ghasemi, A.; Sousa, E.S.; , "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *Communications Magazine, IEEE* , vol.46, no.4, pp.32-39, April 2008
- [38] Geirhofer, S.; Lang Tong; Sadler, B.M.; , "Cognitive Medium Access: Constraining Interference Based on Experimental Models," *Selected Areas in Communications, IEEE Journal on* , vol.26, no.1, pp.95-105, Jan. 2008
- [39] Juncheng Jia; Qian Zhang; Xuemin Shen; , "HC-MAC: A Hardware-Constrained Cognitive MAC for Efficient Spectrum Management," *Selected Areas in Communications, IEEE Journal on* , vol.26, no.1, pp.106-117, Jan. 2008
- [40] Xiaorong Zhu; Lianfeng Shen; Yum, T.-S.P.; , "Analysis of Cognitive Radio Spectrum Access with Optimal Channel Reservation," *Communications Letters, IEEE* , vol.11, no.4, pp.304-306, April 2007
- [41] Wei Ren; Qing Zhao; Swami, A.; , "Power control in cognitive radio networks: how to cross a multi-lane highway," *Selected Areas in Communications, IEEE Journal on* , vol.27, no.7, pp.1283-1296, September 2009
- [42] Yan Chen; Guanding Yu; Zhaoyang Zhang; Hsiao-hwa Chen; Peiliang Qiu; , "On cognitive radio networks with opportunistic power control strategies in fading channels," *Wireless Communications, IEEE Transactions on* , vol.7, no.7, pp.2752-2761, July 2008
- [43] Fangmin Xu; Luyong Zhang; Zheng Zhou; Yabin Ye; , "Architecture for Next-Generation Reconfigurable Wireless Networks using Cognitive

Radio," *Cognitive Radio Oriented Wireless Networks and Communications, 2008. CrownCom 2008. 3rd International Conference on* , vol., no., pp.1-5, 15-17 May 2008

- [44] Handley, M.; Schulzrinne, H.; Schooler, E.; Rosenberg, J.D.; , "SIP: Session Initiation Protocol," *IETF RFT 2543*, March 1999
- [45] Bo Rong; Yi Qian; , "An enhanced SIP proxy server for wireless VoIP in wireless mesh networks," *Communications Magazine, IEEE* , vol.46, no.1, pp.108-113, January 2008
- [46] Cuevas, A.; Moreno, J.I.; Vidales, P.; Einsiedler, H.; , "The IMS service platform: a solution for next-generation network operators to be more than bit pipes," *Communications Magazine, IEEE* , vol.44, no.8, pp.75-81, Aug. 2006
- [47] Rosenberg, J.; Schulzrinne, H.; Camarillo, G.; Johnson, A.; Peterson, J.; Spar, R.; Handley, M.; Schooler, E.; , "The session initiation protocol (SIP)," *IETF RFC 3261*, 2001
- [48] Ohta, M.; , "Performance comparisons of transport protocols for session initiation protocol signaling," *Telecommunication Networking Workshop on QoS in Multiservice IP Networks, 2008. IT-NEWS 2008. 4th International* , vol., no., pp.148-153, 13-15 Feb. 2008
- [49] Roach, A.; , "Session Initiation Protocol (SIP) – Specific Event Notification," *IETF RFC 3265*, June. 2002
- [50] Hung-Bin Chang; Kwang-Cheng Chen; Prasad, N.R.; Chih-Wei Su; , "Auction Based Spectrum Management of Cognitive Radio Networks," *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th* , vol., no., pp.1-5, 26-29 April 2009
- [51] Huang, J.W.; Krishnamurthy, V.; , "Application of Mechanism Design in Opportunistic Scheduling under Cognitive Radio Systems," *Advances in Mesh Networks, 2009. MESH 2009. Second International Conference on* , vol., no., pp.50-54, 18-23 June 2009
- [52] Nossenson,R.;"Long-term evolution network architecture," *Microwaves, Communications, Antennas and Electronics Systems, 2009. COMCAS*

2009. *IEEE International Conference on* , vol., no., pp.1-4, 9-11 Nov. 2009

- [53] Doppler, K.; Rinne, M.; Wijting, C.; Ribeiro, C.; Hugl, K.; , "Device-to-device communication as an underlay to LTE-advanced networks," *Communications Magazine, IEEE* , vol.47, no.12, pp.42-49, Dec. 2009
- [54] Fathi, H.; Chakraborty, S.; Prasad, R.; , "Optimization of VoIP session setup delay over wireless links using SIP," *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE* , vol.6, no., pp. 4092- 4096 Vol.6, 29 Nov.-3 Dec. 2004
- [55] Banks, J.; Carson, J.S.; Nelson, B.; Nicol, D.M.; , *Discrete-Event System Simulation 3<sup>rd</sup> Ed.* Upper Saddle River, NJ: Prentice Hall, 2001, pp. 231-235
- [56] Bin Chen; He-Kun Wu; Anh Tuan Hoang; Ying-Chang Liang; , "Optimizing the second-price auction algorithm in a dynamic cognitive radio network," *Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on* , vol., no., pp.1514-1518, 19-21 Nov. 2008
- [57] Holma, H.; Toskala, A.; , *LTE for UMTS: Evolution to LTE-Advanced 2<sup>nd</sup> Ed.* West Sussex, UK: John Wiley & Sons, Ltd, 2011, pp. 289
- [58] Dimou, K.; Min Wang; Yu Yang; Kazmi, M.; Larmo, A.; Pettersson, J.; Muller, W.; Timner, Y.; , "Handover within 3GPP LTE: Design Principles and Performance," *Vehicular Technology Conference Fall (VTC 2009-Fall), 2009 IEEE 70th* , vol., no., pp.1-5, 20-23 Sept. 2009