University of Bradford eThesis

Performance Analysis and Modelling of Spectrum Handoff Schemes in Cognitive Radio Networks

Modelling and Analysis of Spectrum Handoff Decision Schemes in Cognitive Radio Networks using the Queuing Theory and Simulation for Licensed and Unlicensed Spectrum Bands

Salah Mohammed Bashir ZAHED

A thesis submitted for the degree of Doctor of Philosophy

School of Electrical Engineering and Computer Science

December 2013

ABSTRACT

Recently, wireless access has become an essential part of modern society. Consequently, the demand for new wireless applications and services, as well as the number of wireless users, are gradually increasing. Given that this amount of expansion is eventually controlled by the available radio frequency spectrum, government regulatory agencies have recently adopted a strict approach to the licensing of limited amounts of spectrum to different entities (e.g., public safety, military, service providers, unlicensed devices, and TV). All of them possess exclusive transmissions to their assigned frequency channels. A new study on spectrum efficiency revealed big geographic and temporal variations in spectrum utilisation, ranging from 15-85% in the bands below 3GHz. These variations were less at frequencies above this figure. Recently, the Cognitive Radio (CR) has risen as an encouraging piece of technology to improve spectrum efficiency and to solve the problem of spectrum scarcity. This is because CR allows the secondary (unlicensed) users to occupy unused licensed spectrum bands temporarily, given that the interference of the primary (licensed) users is prohibited or minimised.

In this thesis, various spectrum handoff management schemes have been proposed in order to improve the performance evaluation for CR networks. The proposed spectrum handoff schemes use the Opportunistic Spectrum Access (OSA) concept to utilise available spectrum bands. The handoff Secondary Users (SUs) have a higher priority to occupy available spectrum channels in the licensed and unlicensed spectrum bands without interfering with the legacy spectrum owner, i.e.

primary users (PUs). However, existing spectrum handoff management schemes in CR networks do not provide high transmission opportunities for handoff secondary users to utilise the available radio spectrum resources. The first part of this thesis addresses the issue of spectrum handoff management in a licensed spectrum band environment. In this case, both reactive and proactive spectrum handoff schemes are proposed. Queuing theory or/and simulation experiments have been used to evaluate the performance of the proposed schemes and compare them with other existing schemes. Handoff delay has mainly been used to investigate the impact of successive handoff operations on the performance of the proposed CR networks. Implemented models have shown an improvement in the adopted performance measures. According to the achieved results, the improvement of the proposed, prioritised handoff schemes in some cases is approximately 75% when compared with existing schemes.

On the other hand, the second part of this research proposed a prioritised spectrum handoff scheme in a heterogeneous spectrum environment, which is composed of a pool of licensed and unlicensed spectrum channels. In general, the availability of substantial numbers of the licensed spectrum channels is the key benefit of using this type of radio spectrum channel. Whereas, accessing with equal rights for all types of users is the main advantage of using unlicensed spectrum channels. In this respect, no transmission interruptions occur once a user obtains a channel. In addition, the proposed schemes use only the unlicensed spectrum channels as their backup channels. This enables the user to resume interrupted transmission in the case of the spectrum handoff operation (mainly; due to the appearance of the primary users), and thus facilitates a SUs communication. The proposed principle is investigated using a retrial queuing theory as well as extensive simulation experiments, and is compared with another non-prioritised scheme which do not give any preference to handoff SUs over new SUs. The results indicate that the proposed model has improved on current average handoff delay.

This thesis contributes to knowledge by further enhancing the efficient utilisation of available radio spectrum resources and therefore subsequently provides an improvement in the spectrum capacity for wireless cognitive radio networks.

Keywords: Cognitive radio networks, Spectrum handoff, Handoff delay, Queuing theory.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my PhD advisors, Professor Irfan Awan, and Doctor Andrea Cullen for their support during my study. I am also very grateful to Professor Khairi Hamdi and my colleague Abdulaziz Aljohani for their scientific advice and knowledge and many insightful discussions and suggestions.

I would like also to thank my friends and colleagues for their support and encouragement when it was most needed.

I dedicate this thesis to

my family, my wife, and my beloved children

for their patience, encouragement, commitment and unconditional

support during throughout this journey.

LIST OF AUTHOR'S PUBLICATIONS

Journals:

- S. Zahed, I. Awan, and A. Cullen, "Analytical modeling for spectrum handoff decision in cognitive radio networks," *Simulation Modelling Practice and Theory,* vol. 38, pp. 98-114, 2013.
- S. Zahed, I. Awan, A. Cullen, A. Aljohani, "Performance Evaluation of Cognitive Radio Networks under Heterogenous Spectrum Bands". Submitted to IEEE Transaction on Parallel and Distributed Systems.

Conferences:

- S. Zahed, I. Awan, and M. Geyong, "Performance Evaluation of Secondary Users in Cognitive Radio Networks," in 27 th UKPEW, Bradford, UK, 2011, pp. 340-343.
- S. Zahed, I. Awan, and M. Geyong, "Prioritized Proactive Scheme for Spectrum Handoff Decision in Cognitive Radio Networks," presented at the 7th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA 2012), University of Victoria, Victoria, Canada, 2012.
- S. Zahed, I. Awan, A. Cullen, "Performance Evaluation of Cognitive Radio Networks under Common Handoff Queue", Seventh International Open Conference HET-NETs 2013 (IFIP), Ilkely, Bradford, UK, 2013. (work in progress).
- S. Zahed, I. Awan, A. Cullen, "Performance Evaluation of Cognitive Radio Networks under Reactive-Decision Spectrum Handoff Scheme", Seventh

International Open Conference HET-NETs 2013 (IFIP), Ilkely, Bradford, UK, 2013. (work in progress).

- Salah Zahed, Irfan Awan, Andrea Cullen, M. Younas, “Performance Evaluation of Cognitive Radio Networks under License and Unlicensed Spectrum Bands”, submitted to the 28th IEEE International Conference on Advanced Information Networking and Applications (AINA-2014) Victoria, Canada, May 13-16, 2014.

TABLE OF CONTENETS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CC	Cognition Cycle
CCC	Common Control Channel
CH	Channel
CR	Cognitive Radio
CRN	Cognitive Radio Network
CSMA	Carrier Sense Multiple Access
DECT	Digital Enhanced Cordless Telecommunication
DSA	Dynamic Spectrum Access
FCC	Federal Communication Commission
FCFS	First Come First Serve
FSA	Fixed Spectrum Access
GSM	Global System for Mobile Communications
HOL	Head of line
IEEE	Institute of Electrical and Electronic Engineers
ISM	Industrial, Scientific and Medical
MAC	Medium Access Control

MATLAB	Matrix Laboratory
NSWH	Non-switching Handoff
OSA	Opportunistic Spectrum Access
Ofcom	Office of Communications
PCA	Priority Contention Access
PRP	Preemptive Resume Priority
PU	Primary User
QoS	Quality of Service
RAH-NEW	Random Handoff-New
RAH-OLD	Random Handoff-Old
REH-NEW	Reactive Handoff-New
REH-OLD	Reactive Handoff-Old
REH-SQ	Reactive Handoff-Shared Queue
RF	Radio Frequency
SDR	Software Defined radio
SPTF	Spectrum Policy Task Force
SWH-OLD	Switching Handoff-Old

SU	Secondary User
SWH-NEW	Switching Handoff-New
TV	Television
UMTS	Universal Mobile Telecommunication System
UNII	Unlicensed National Information Infrastructure
WLAN	Wireless Local Area Network
XG	Next Generation
Wi-Fi	Wireless Fidelity

Chapter 1.

INTRODUCTION

Wireless communication has become extremely important in modern society, with many of our activities implying some kind of wireless access. As a consequence, the number of wireless users, standards, applications and services, are steadily increasing given that most of the limited available radio spectrum resources have been allocated well and can constrain this growth [1]. Also, this amount of expansion is eventually controlled by the available radio frequency spectrum; government regulatory agencies have adopted, until recently, a strict approach to the licensing of limited amounts of spectrum to different entities (e.g., public safety, military, service providers, analogue cellular telephony, and TV), all of them possess exclusive transmissions to their assigned frequency channels. By using this strict approach, the main access method of a radio spectrum resource is based on a fixed spectrum allocation basis, called Fixed Spectrum Access (FSA). The Federal Communication Commission (FCC) published a report prepared by the Spectrum Policy Task Force (SPTF) [2], which proved that most of the dedicated (licensed) spectrum experience low utilisation efficiency, such as those allocated to a TV or analogue cellular telephony, which are underutilised. The study revealed big geographic and temporal changes in spectrum utilisation, ranging from 15-85% in the bands below 3GHz, while ranging lower than this at higher frequencies [3]. These temporally unused spectrums, called spectrum holes or white spaces, are illustrated in Figure 1.1. These spectrum holes are sometimes called “virtual channels”, which are logical channels built over the spectrum holes of the licensed PU channels [4].

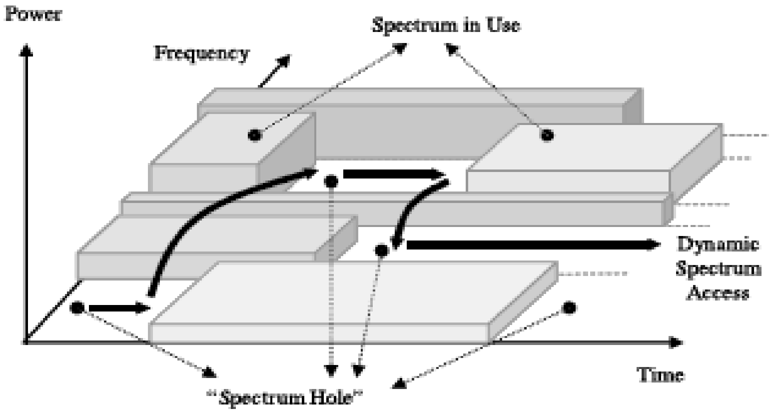

Figure 1.1: Spectrum hole concept [5].

Due to this, new regulations are required to utilise the existing spectrum bands more efficiently. There are two solutions for the regulator to overcome this problem. The first solution is to expand or increase the spectrum for the unlicensed bands. However, practically, this solution is not easy to apply and would take a long time for the regulators to implement on the ground. The second solution is to give permission to the unlicensed wireless networks to use the unutilised licensed spectrum bands (white spaces) opportunistically; as a result, the overall utilisation would increase. The latter solution is called Opportunistic Spectrum Access (OSA). To realise the concept of OSA, the state of the art “cognitive radio” [6, 7] is developed as a promising piece of technology which is expected to be the main element of future wireless communication networks. Cognitive radio technology allows the unlicensed (secondary) users to utilise the unused spectrum without any harmful interference to the licensed (primary) users’ transmissions. In the case of the primary users appearing in the bands occupied by the secondary users, the secondary users have to empty the bands and look for new bands to complete the transmissions.

1.1 Motivation

Even though there are several advantages of using cognitive radio networks, such as increasing the spectrum utilisation of wireless networks, there are still some challenges which have not yet been comprehensively discussed:

Firstly, although cognitive radio networks are planned to operate over a heterogeneous spectrum environment, which is comprised of licensed and unlicensed spectrum bands, only some of the research has been carried out in such an environment as merging licensed and unlicensed spectrum bands for transmission increases the spectrum utilisation of cognitive radio wireless networks. Instead, most of the research has been done in only the licensed band of the spectrum.

Secondly, although spectrum utilisation will be improved when considering CRNs, the effect of multiple and sudden appearances of the PUs in the licensed spectrum band have not yet been explored broadly. In order to mitigate and compensate the adverse effects of spectrum handoff, handoff secondary users should be provided with a higher priority over uninterrupted secondary (new) users in utilising the available spectrum band.

The aforementioned challenges motivated us to: develop innovative spectrum handoff schemes in order to improve the performance of the secondary users in terms of their handoff delay, which is one of the main themes of our research as well.

1.2 Research Goals and Objectives

The main goals of this thesis are given below:

- To develop proactive-decision spectrum handoff schemes in CRNs which operate under the licensed spectrum band.
- To develop a reactive-decision spectrum handoff scheme in CRNs that operates under the licensed spectrum band.
- To develop a shared queue spectrum handoff scheme in CRNs that operates under the licensed spectrum band.
- To develop a spectrum handoff scheme in CRNs that operates under a heterogeneous spectrum environment, consisting of the licensed and unlicensed spectrum bands.

These goals are to be evaluated according to the following steps:

- To implement a simulation platform.
- Calculate performance measures for the proposed models.
- Validate simulation results using the results achieved from the analytical model whenever it is provided.

1.3 Contributions

The contributions of this thesis are listed as follows:

- The implementation of novel proactive-decision spectrum handoff schemes (random and switching schemes) to evaluate the performance of the proposed cognitive radio network using the Preemptive Resume Priority (PRP) M/G/1 queuing network model.

- The implementation of a spectrum handoff scheme based on instantaneous queue length in order to investigate the performance of the proposed cognitive radio network using the Preemptive Resume Priority (PRP) M/M/1 queuing network model. The new provided scheme is categorised under a reactive-decision handoff approach.
- The implementation of a reactive-decision spectrum handoff scheme to investigate the performance of cognitive radio networks under common handoff queues using the Preemptive Resume Priority (PRP) M/M/1 queuing network model.
- The development of a novel spectrum handoff scheme so as to evaluate the performance of the proposed cognitive radio network under a heterogeneous spectrum environment with single-licensed and single-unlicensed spectrum channels (N_1=1, N_2=1), and multi-licensed and multi-unlicensed spectrum channels (N_1=n1, N_2=n2). The proposed CRN is modeled using the Preemptive Resume Priority (PRP) M/M/C queuing network model for the licensed band and the M/M/C retrial queuing network model for the unlicensed band.

1.4 Thesis Organisation

This thesis is organised according to the following plane,

Chapter 2 presents background regarding the regulations of the radio spectrum, and fixed and dynamic spectrum access techniques. Also, this chapter provides an overview of the state of the art “cognitive radio”, which includes a software-defined-radio, key challenges of cognitive radio technology and cognitive radio network architecture.

Chapter 3 presents a novel random and switching spectrum handoff decision schemes. The two provided schemes are categorised under the proactive-decision handoff approach. Also, illustrated examples are presented and an extensive and comprehensive comparison study is established between the new and other existing – generated - spectrum handoff schemes, such as the old corresponding schemes. Achieved simulations results are validated using the corresponding analytical models and are analysed by the queuing theory.

Chapter 4 proposes a reactive-decision spectrum handoff scheme. The model assumes a central controller that provides handoff users with the information regarding the states of the secondary users’ queues for each wireless channel, namely, the instantaneous queue length. Then, this information is used as the key factor in determining the spectrum handoff decision. Simulation results obtained from the developed scheme are compared with another existing generated sensing-based scheme as well as with other developed schemes in Chapter 3.

Chapter 5 introduces another reactive-decision spectrum handoff scheme. The developed model provides handoff secondary users from all channels with a shared queue, instead of a separate queue, for each channel. To investigate the performance of the implemented model, a simulation study is conducted to compare the developed scheme and some other spectrum handoff schemes implemented in Chapter 3 and Chapter 4.

Chapter 6 presents a novel prioritised spectrum handoff scheme in a heterogeneous spectrum environment, composed of unlicensed and licensed spectrum bands. Two scenarios can be considered: single-licensed and single-unlicensed spectrum channels (N_1=1, N_2=1), and multi-licensed and multi-unlicensed spectrum channels (N_1=n1, N_2=n2). The proposed model is decomposed into two main parts in order to facilitate the methodical analysis. Simulation results are validated using the queuing theory whenever it is possible.

Chapter 7 completes the thesis and provides conclusions that are related to every chapter. Additionally, this chapter indicates any future work of interest.

1.5 Methodology

The prediction of real world behaviour can be achieved either by simulating the system or by the use of theoretical analysis. The queuing theory, besides simulation experiments, is considered to be from the main standard methodologies in the telecommunications, computer science, and computer engineering. Theoretical analysis approach seems to be cheaper and faster solution, whereas, simulation approach can be used when the mathematical model of the system is very complex. However, simulation approach is an expensive and time consuming solution. In this thesis both approaches are used to analyse the proposed models. For the simulation approach, the discrete event simulator MATrix LABoratory (MATLAB) package tool is used to analyse the delay performance measures, such as the handoff delay, the cumulative handoff delay, the total service time, and several others. Recently, MATLAB has been widely used in simulating and analysing communication networks and it gives good approximations for exponentially distributions traffic. For the theoretical analysis approach, the well-known queuing theory is used to model the corresponding proposed models. The following subsections will provide us with some ideas regarding the theoretical analysis approach, which is relevant to this topic.

1.5.1 Queuing Theory

The queuing theory [8, 9] has been used since the early 1900s to analyse communication systems. Models based-on the queuing theory have achieved relatively high accuracy at a low cost. But, despite the simplicity of its idea (the users arrive, request for service, and wait in a buffer until getting the service, or otherwise leave the system), the subject has a considerable amount of complexity and subtlety [10].

1.5.2 Priority Queues

In queuing systems, some users are provided with special treatment. Systems which provide such treatment are called priority queuing systems. In these systems, there are at least two types of users: low-priority users and high-priority users. The high-priority users are always given a higher priority over the low-priority users in utilising the system resources, such as the spectrum channels. Two categories of priority queues are identified here: Preemptive priority queues and non-preemptive priority queues [11].

1.5.2.1 Preemptive Priority Queues

In preemptive priority queues, newly arriving high-priority users interrupt the transmissions of the low-priority users receiving service, and then start transmitting their data immediately. The interrupted users join the low priority queue. According to the disciplines of this queue, a low-priority user can only transmit its data when there are no high-priority users in the queue at the time [11] .

1.5.2.2 Non-preemptive Priority Queues

In non-preemptive priority queues, the newly arriving users do not interrupt the user in service, regardless of their type. Instead, they wait until the user in the service finishes its transmission. This type of priority queues is sometimes denoted as a head-of-line (HOL) priority queue [12].

1.5.2.3 Retrial Queues

Queuing systems with returning customers are denoted by retrial queues. In many telecommunications, computers, and telephony switching networks, newly arrived users which find the service facility busy enter the orbit (also, called the retrial group), and retries in getting the service after a random amount of time in a random order. The retrial queuing theory is used to model such systems theoretically which have retrial phenomenon [13-15]. The orbit is an invisible buffer which is not possible to be observed [16]. Typical examples of such systems which have retrial phenomenon are random access protocols [17], such as Carrier Sense Multiple

Access (CSMA) which is used mainly in getting access to the unlicensed spectrum bands.

1.5.3 Service Completion

In the preemptive priority queues mentioned above, the completion of the interrupted transmissions can be carried out in various ways; resuming the transmission, repeating the transmission, or sometimes dropping the transmission.

1.5.3.1 Resume Transmission

In this case, interrupted transmissions are not retransmitted again [18] if the interrupted process does not result in any loss. Instead, the interrupted transmission will be retransmitted from the instance of interruption [18].

1.5.3.2 Repeat Transmission

In the repeated transmission service, the interrupted transmission should be totally retransmitted [18].

1.5.3.3 Drop Transmission

In this type of service, interrupted transmissions will be dropped and will not be transmitted any more.

Chapter 2.

RADIO SPECTRUM & COGNITIVE RADIO NETWORKS

2.1 Introduction

The precious radio spectrum and its exploit are firmly controlled by governments in most world countries. Also, regulatory bodies, such as Ofcom in the UK and FCC in the USA, enforce the spectrum allocation paradigm “command-and-control”. Most existing wireless devices and networks follow the legacy fixed spectrum access (FSA) policy in order to use the radio spectrum, in which the spectrum bands are licensed to being devoted (primary) services and users, such as cellular networks, TVs, and vehicular ad hoc networks. However, in these systems, only the primary users can operate in the assigned spectrum. On the other hand, the other systems are not allowed to use this spectrum, even when it is idle.

Although the issue of interference among different devices and networks can be efficiently controlled and coordinated by using FSA, this policy still causes significant spectrum underutilisation, as reported in [19]. The growth of new wireless services and applications along with diverse network architectures (such as heterogeneous networks) has been adversely affected by the inefficiency of the radio spectrum [19]. As a consequence, new flexible and dynamic spectrum regulations should be introduced in order to overcome this issue.

2.2 Dynamic Spectrum Access (DSA)

As mentioned above, the spectrum access policies of today are static. On the other hand, dynamic spectrum access techniques propose that a radio system can adapt to the available spectrum bands dynamically, and use them with limited rights [20]. As a consequence, dynamic spectrum access can overcome the problem of fixed spectrum assignment. The general DSA approaches are explained in Figure 2.1. DSA can be classified into three main models: dynamic exclusive use, open sharing (also denoted by commons), and hierarchical access [21]. The following subsections give an idea for each type of these models.

2.2.1 Dynamic Exclusive Use Model

In such models, an exclusive use of the licensed spectrum is allocated dynamically among possible licensees (users or providers). Radio regulation bodies often govern this kind of DSA models. This model is categorised into the types Spectrum property rights, and Dynamic spectrum allocation;

Spectrum property rights: In this type, spectrum owners are allowed to sell and trade the spectrum and use it through any technology they want.

Dynamic spectrum allocation: This type improves spectrum efficiency through dynamic spectrum assignment for more than just the service. It exploits the spatial and temporal traffic statistics (demands) of different services. In other words, the spectrum is allocated exclusively to specific services at a given time and a given area.

2.2.2 Open Sharing Model (Spectrum Commons)

In this DSA type, the spectrum is open for use and share for different users. The main idea here is that users compete to share the spectrum with equal rights.

2.2.3 Hierarchical Access Model

This type of DSA relies on the hierarchical access structure of the radio spectrum. Two types of users are identified here: secondary users and primary users. Also, the licensed spectrum is open for the secondary users to use. The hierarchical access model is categorised into two types: spectrum underlay and spectrum overlay. The main features which discriminate each type are explained in Section 2.9. The spectrum overlay approach is sometimes called OSA.

The hierarchical access model is the most compatible model, with legacy wireless systems and current policies of spectrum management. It is possible to use this model without any modifications to the spectrum regulations. In this thesis, only the hierarchical access model is considered, in particular, the spectrum overlay approach, in which secondary users utilise spectrum holes in an opportunistic manner.

The main property of OSA is that it opens the licensed spectrum bands for usage by the secondary users, without interfering or negotiating with the primary users. The spectrum allocation paradigm (OSA) is introduced and is motivated by:

- The limitation of the available radio spectrums for transmission,
- The fact that most of the radio spectrum bands are already exclusively assigned,

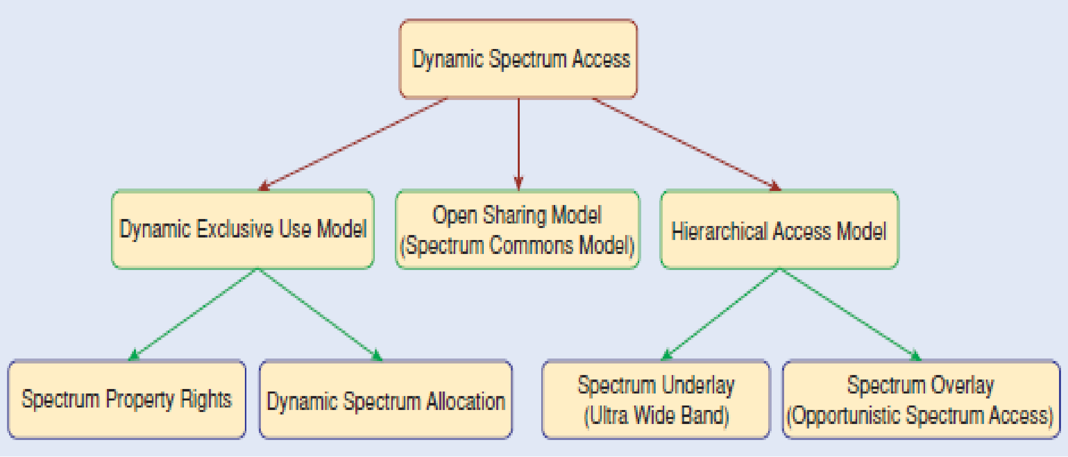

Figure 2.1: Dynamic spectrum access approaches [22].

➢ The spectrum underutilisation.

These three factors give the opportunity for more and more unlicensed (secondary) users to share or reuse the licensed frequency band, without interfering with the transmission of the incumbent licensed (primary) users (e.g., TVs, cellular phones, radars, and microphones). Secondary users (e.g., Wi-Fi, ZigBee, and Bluetooth) occupying the licensed spectrum should vacate the spectrum when required by the spectrum owner [1], [23], [24]. This state of the art trend is called the “Cognitive Radio”,[25], [26] and is defined in the following subsections. However, the cognitive radio requires new wireless technologies, architectures, and algorithms to make it feasible in practice.

As a consequence, properties of the neXt Generation (XG) wireless networks would have to be much more cognitive than recent networks by introducing the learning elements in the networks. This ability will allow the networks to monitor the users' traffic and mobility, and signal the core network to act on adjustment for the utilisation of spectrum and switching components [27].

2.3 Radio Spectrum

The electromagnetic spectrum part, which is used for transmitting data, voice, and video, is called the "radio spectrum". The radio spectrum includes frequencies from 3 kHz to 300 GHz and is systematised by national and international organisations, which are generally denoted as "regulators" [28]. In general, the regulation of radio spectrums can be distinguished into three categories [2], [28]: Licensed spectrums, which is used for exclusive usage and for shared or common usage, and unlicensed spectrums. Further explanations about these types will be provided in the following subsections.

2.3.1 Licensed Spectrum for Exclusive Use

In this case, within a particular geographical area, license users have exclusive and transferable rights to use a particular assigned spectrum band, with elastic use rights that are ruled mainly by technical rules to protect communications from interference. An example of a licensed spectrum for exclusive usage is spectrum bands for the Universal Mobile Telecommunication System (UMTS) in Europe.

2.3.2 Licensed Spectrum for Shared Use

This type is denoted as the "command-and-control" model. This model is the older model, in which licensee users obtain a spectrum band to use under specific conditions. Changing the uses of the spectrum is a deliberative process which needs revision and prospects of the public. Spectrum bands assigned to Digital Enhanced Cordless Telecommunication (DECT) provide us with an example of this type of licensed spectrum sharing.

2.3.3 Unlicensed Bands

To be precise, there are two main unlicensed spectrum bands which are the industrial, scientific and medical (ISM) bands [29], and the Unlicensed National Information Infrastructure (U-NII) band [30]. The allocated radio spectrum bands for unlicensed usage are small compared with the spectrum bands allocated to the licensed usage. Spectrum access to unlicensed spectrum bands is open and all users have equal priority to utilise free spectrum bands. To avoid interference amongst the spectrum, users have to follow certain technical rules (called etiquettes) before starting communication process.

2.4 Cognitive Radio: Overview

The exploration of the cognitive radio trend involves an interdisciplinary effort from different technical areas, including: communications engineering, the networking, spectrum policies, adaptive systems and learning, signal processing, the information theory, the game /cooperative theory, economics and social sciences.

Although there have been many advances in the cognitive radio trend with respect to enabling DSA networks, further study is required to make DSA a feasible spectrum allocation paradigm in CRNs. Major existing studies on cognitive radios cover:

- Spectrum policy alternatives and system models
- Spectrum sensing algorithms
- Cognitive radio architecture and software abstractions
- Cognitive algorithms for adaptation and resource management
- DSA technology and algorithms
- Protocol architectures for CRNs
- Network security for CRNs
- Cooperative wireless communications
- Cognitive medium access control (MAC)
- Cognitive networking and the Internet
- Game theory for cognitive radio networks [31]
- Physical layer aspects

Some of research focuses on the implementation of CRs which require no changes in existing networks (infrastructures) [32]. Other research suggests more than a single radio for each secondary user, or they assume secondary users have a wide band spectrum sensing [33], [34]. The following subsections introduce some basic and vital aspects and definitions related to the cognitive radio.

2.5 Software-Defined Radio (SDR)

Fortunately, the rapid evolution of microelectronics has enabled the development of software-defined-radio (SDR) technology, where the baseband digital processing is completely achieved in software. This technology allows to adapt to operating parameters (e.g., transmit power, carrier frequency, coding scheme and modulation strategy), based on learning from previous events and current inputs to the system. This will give opportunities to unlicensed users to access the licensed spectrum without interfering with the primary users [6].

2.6 Cognitive Radio (CR)

The cognitive radio built on SDR is introduced as the enabling technology for dynamic spectrum access. Cognitive radio was first introduced by Joseph Mitola in 1999 [25], [26]. It is defined as “an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment, and adapt its internal states to statistical variations in the incoming RF stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier-frequency, and modulation strategy) in real-time” [6]. Cognitive radio will provide us with “highly reliable communications whenever and wherever needed” [6] with “efficient utilisation of the radio spectrum” [6] via heterogeneous wireless architectures. This can be done by letting secondary users (SUs) temporarily utilise the primary users’ (PUs) unused licensed spectrum bands [35-38].

2.7 Cognitive Cycle (CC)

In cognitive radio networks, secondary users should perform the four spectrum-awareness operations which form the cognitive cycle: spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility. These operations (sometimes called steps) are illustrated in

Figure 2.2, and are discussed in detail in the Section 2.9. In general, secondary users should sense the surroundings whenever they have data to transmit, and characterise available (empty) spectrum bands in terms of radio environment local observations (i.e., interference, received signal strength) and the primary users' statistical behaviors. To characterise spectrum bands in terms of the primary users' statistical behaviors, prior information regarding the activity of the primary users is needed, such as their arrival and service processes. This can be achieved by long-term observations of the licensed spectrum band. This can help in the decision making step in order to select the best spectrum bands which satisfy the secondary users' QoS. Moreover, multiple secondary users should share available spectrum bands in order to avoid collisions. In case the primary users return to their licensed bands, secondary users occupying those bands should perform spectrum mobility and empty the occupied bands immediately, and then search for other unused bands.

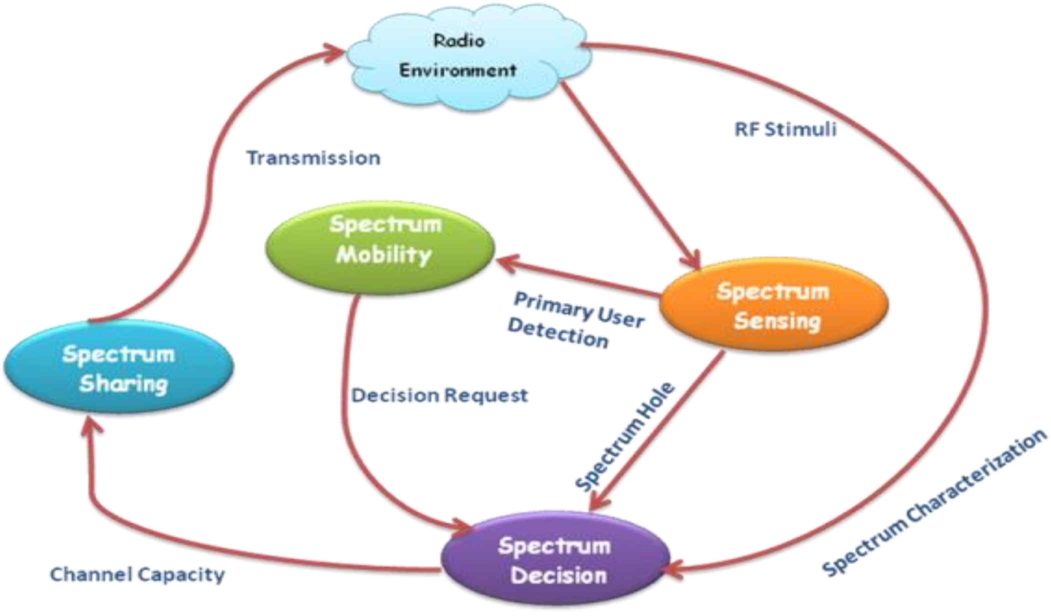

Figure 2.2: Cognitive cycle [39].

2.8 Main Challenges of Cognitive Radio Technology

The two main aspects of cognitive radio are:

- Cognitive capability: Because of the high variation in the available spectrum and the diversity of applications that can be used by CR technology, CR users must be able to:
    - Identify and detect which parts of the spectrum are available at a specific time and location through real time interaction with the surrounding environment (spectrum wideband sensing)
    - Choose the appropriate band for transmission (spectrum decision)
    - Share access processing with other CR users (spectrum sharing)
    - Vacate the spectrum when a primary user appears in the selected band (spectrum mobility)

These network capabilities can be realised by spectrum management functions (see Section 2.9).

- Re-configurability**:** CR users must have the following abilities:
    - They must be able to operate on different systems with different protocols in a wideband frequency range supported by its hardware design. Using this capability, the best networks and channels can be selected and then CR users are able to reconfigure themselves to be compatible with the new environment.

  - They must be able to control the transmission power, for example, transmitting power control can be used to control the power level dynamically, which reduces the emitted power to allow greater sharing of the spectrum when a higher power level is not necessary.
  - They must be able to adapt to the modulation scheme in order to improve spectrum access process.
  - They must be able to frequency agility, which means the ability of the CR user can change its transmission frequency.

2.9 **Cognitive Radio Network Architecture**

The heterogeneity of cognitive radio networks provides us with high reliable wireless communication whenever and wherever needed with efficient utilisation of the radio spectrum. The main features of the CR network architecture are illustrated in Figure 2.3 and can be explained as shown below.

- Network Construction: Two types of networks are identified. The CR network and the primary network.
  - The primary network (also called the licensed network) is an existing network in which the primary users have a certain licensed spectrum band to use. If the primary network has an infrastructure, then the primary users' transmissions are controlled by the primary Base Station (BS) within the transmission range of the same base station. As primary users have the

highest priority in spectrum access, the transmissions of unlicensed users will not disturb the transmission of the primary users.

  - The CR network can also be referred to as the unlicensed network, the secondary network, or the dynamic spectrum access network. The CR network does not have the authority to use the intended spectrum band. As a consequence, in order to share the unused licensed spectrum bands without interfering with the primary users’ transmissions, additional functionalities are needed for CR users. Also, secondary networks can have infrastructure (such as base station in a cellular network or an access point in a wireless local area network (WLAN) that support single hop communications to CR users or infrastructureless (Ad hoc) networks as shown in Figure 2.3. Moreover, secondary networks may be provided with spectrum brokers to distribute the spectrum between other CR networks [40]. The brokers may lease the spectrum according to some conditions such as the maximum transmission power and the spatial region where the spectrum will be used [41].

- The Heterogeneous Spectrum Environment: CR users have the ability to access both licensed and unlicensed bands of the spectrum [36, 42, 43]. This can be achieved by using wideband access technology. As a consequence, two types of spectrum access are identified here: licensed band access and unlicensed band access. The primary network is principally operated using the licensed band, therefore CR networks should identify the unused spectrum bands in such cases. But, if primary users return back to their spectrum, CR users should immediately empty the spectrum band and resume their transmission in another spectrum band or at another opportunity. On the other hand, CR users have the same rights to use the unlicensed bands in the absence of the primary users. However, spectrum sharing techniques are required for the secondary users to compete for the unlicensed spectrum bands.

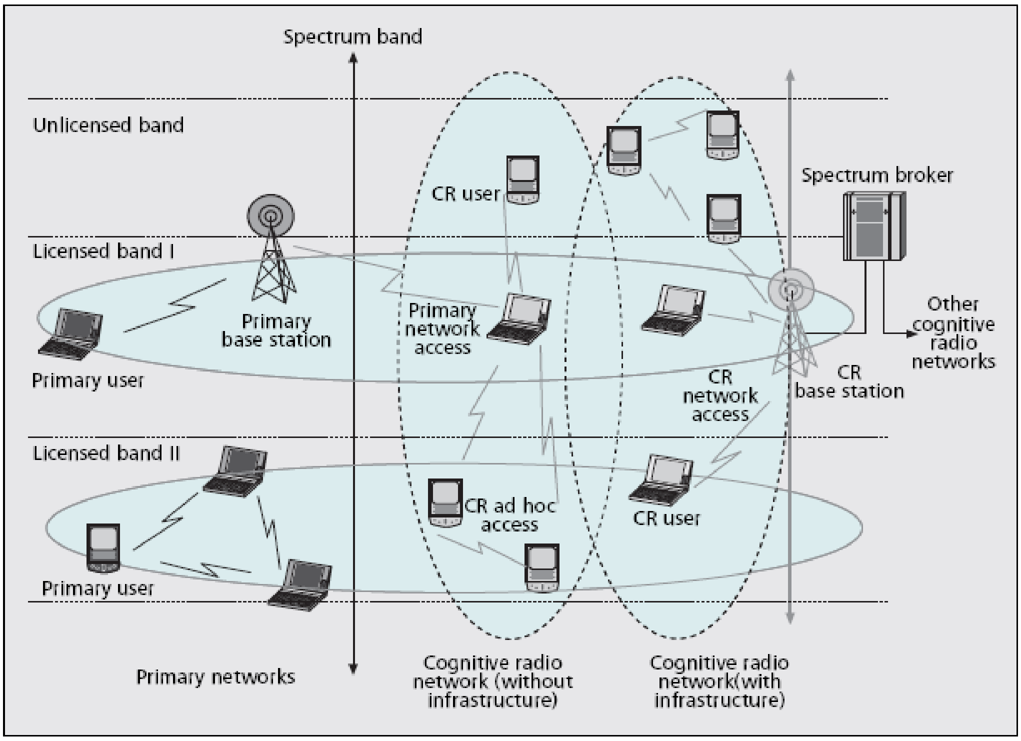

Figure 2.3: Cognitive radio network architecture [5].

- The Heterogeneous Network Access: CR users can access the available spectrum bands by three different ways, as shown in Figure 2.3. The first access method is CR network access, in which CR users communicate with each other through their base station using both licensed and unlicensed spectrum bands. In the second access method, which is CR ad hoc access, the communication between CR users is done in an ad hoc manner using both licensed and unlicensed spectrum bands. Finally, the last access method is primary network access; in this type, CR users can access the primary base station through the licensed spectrum band.

➢ Spectrum Management Outlines: the coexistence of CR networks (CRNs) with primary networks as well as a variety of QoS requirements which can create big challenges. Therefore, sophisticated spectrum management functions are required for CRNs besides some design issues, such as CRNs should avoid interference with primary networks. Also, CRNs should support QoS-aware communication in order to select the best spectrum band, taking into consideration the heterogeneous and dynamic spectrum environment. Another issue is that spectrum mobility in CRNs should be performed with seamless communication. In order to address such challenging issues, the main features of spectrum management in CRNs should cover the four steps [44], shown in Figure 2.2.

- The first step is spectrum sensing in which secondary users should observe the available spectrum bands and identify spectrum holes by performing wideband spectrum sensing. Also, characteristics and statistics of available spectrum bands should be carefully monitored and captured in this step. Moreover, spectrum sensing should be performed periodically to detect primary users and protect them from any interference.

- The second step is spectrum decision. After identifying the available spectrum bands, CR users need to decide which band to choose according to the targeted QoS. Characterisation of the available spectrum bands according to the primary users' activity and radio environment is an essential issue [44]. Also, prior information regarding the users' activity is needed here which can be provided and updated periodically by a centralised database [45], such as channels and queues status. When a CR user desires to transmit data, it request needed information regarding available channels' characteristics from the base station, based on those information it selects appropriate channel for transmission.

- The third step is spectrum sharing. It is possible for multiple CR users to try to access the same spectrum bands at the same time, which would cause disturbances and collisions to their transmissions. Therefore, spectrum sharing should coordinate the transmissions to limit collisions between multiple CR users, and provide the ability to share the unused spectrum opportunistically between multiple CR users without interfering with the primary network. An important issue should be considered in this step, which is transmitter-receiver handshaking; when a spectrum band is determined for transmission, the transmitter should inform the receiver regarding the selected spectrum band by exchanging RTS/CTS messages through a dedicated control channel. First the transmitter sends an RTS message to the receiver. The RTS message containing the determined band for transmission. When the receiver respond by the CTS message, then the transmission will be continued using the determined spectrum band.

Spectrum sharing can be achieved horizontally or vertically. If all users have the same priority to access the spectrum band (for example, accessing the unlicensed spectrum bands such as ISM bands), then this type of spectrum sharing is called “horizontal spectrum sharing”. On the other hand, if one user (i.e., the spectrum owner) has a higher authority to use the spectrum band (for example, accessing the licensed spectrum bands such as TV broadcasting and Global System for Mobile Communication (GSM) bands), then this type of spectrum sharing is called “vertical spectrum sharing”.

Moreover, spectrums can be shared on the basis of time domain, space domain, frequency domain, power domain, or a mixture of these domains. Transmission opportunities are obtained from the time domain, space domain, frequency domain, or a mixture of these domains, which are denoted as spectrum holes, or white spaces, whereas they are denoted as gray spaces in the power domain [6].

Also, spectrum sharing can be classified in terms of access technology into two main classes: overlay spectrum sharing, and underlay spectrum sharing. In the former type, CR users can only utilise spectrum bands that have not been used by the primary users, as a consequence, this can minimise the interference to the primary users. Many techniques can be used to realise this approach, such as Time Division Multiple Access (TDMA) and Frequency Division Multiple Access (FDMA). On the other hand, the other type of spectrum sharing permits CR users to transmit with a high power in the time domain, space domain, frequency domain, or any mixture of these domains, but without exceeding the interference tolerance limit to the primary users. In this case, primary users consider the transmission of CR users as a noise. Underlay spectrum sharing can utilise the spectrum better but increase in complexity. In this thesis, we only consider the overlay spectrum sharing class to utilise the available spectrum bands opportunistically.

Furthermore, spectrum sharing can be classified according to the access mode, hence two access mode types are identified: contention-free and contention-based. In the first type, the access to the spectrum is based on the time slots which construct the frames, whereas in the second type, CR users compete randomly to access the spectrum band. The carrier sense multiple access (CSMA) mechanism is an example of this type of access mode.

Finally, spectrum sharing can be classified from the point of architecture, in this case two types are considered: centralised and distributed architecture. In the centralised architecture, a centralised entity (i.e., the base station) controls spectrum allocation (selecting a proper band based on the required QoS) and spectrum access. On the other hand, in the distributed architecture, each CR user performs spectrum allocation and accesses the spectrum individually. This solution can be employed where infrastructure construction is not desirable.

- The fourth and last step is spectrum mobility. With respect to spectrum mobility, if a primary user is detected in the specific spectrum band in use, CR users should vacate the band immediately and resume their transmission in another unoccupied band. The management of spectrum mobility ensures fast and smooth spectrum transition in the handoff process. This can lead to minimum QoS degradation for CR users. Therefore, spectrum mobility requires a handoff mechanism to detect the primary users and to switch to another spectrum band with minimum service degradation. Another possibility is to wait until primary users finish their transmission and then to continue transmission in the same band. Spectrum handoff (mobility) has received less attention from the research community in comparison to the other functionalities [36, 46]. In general, spectrum handoff approaches can be classified, from the point of view of decision timing for choosing the goal channel for next spectrum handoffs, into two major groups: proactive-decision spectrum handoffs and reactive-decision spectrum handoffs [47, 48], as shown in Figure 2.4. The next chapters will discuss these approaches in detail and propose some spectrum handoff schemes in cognitive radio networks in order to address the issues of spectrum mobility.

2.10 **Related Work**

Spectrum handoff happens when a PU appears in a spectrum band which is already occupied by an SU. Little work has been done on spectrum handoffs in comparison to the functionalities of cognitive radio networks, such as spectrum sensing, spectrum management, and spectrum sharing [36, 46].

The main properties of the proactive and the reactive spectrum handoff approaches mentioned above are:

- In the proactive-decision spectrum handoffs type, such as [49-56], SUs prepare goal channels for spectrum handoffs before they start their transmission, after periodically monitoring all the channels to collect information about channel statistics and decide the appropriate set of goal channels for future spectrum handoffs, based on long-term observation results.

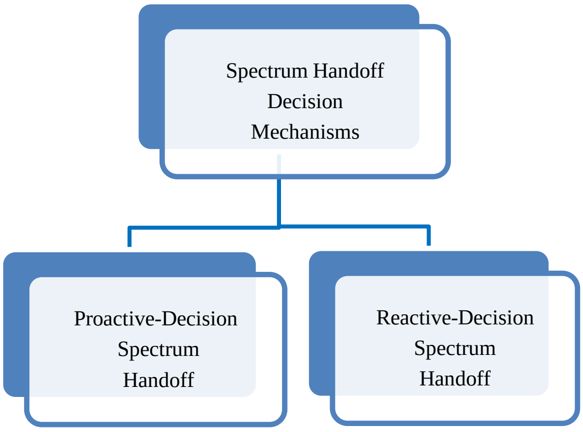

Figure 2.4: Spectrum handoff decision in CRNs.

- For reactive-decision spectrum handoffs [35, 57-61], the interrupted secondary user searches for goal channels in an 'on-demand' way (mostly by instantaneous wideband sensing) after the interruption process occurs [61, 62]. Following this search, the interrupted transmission can be resumed on one of the goal channels.

However, each type has its advantages and disadvantages. For example, proactive-decision spectrum handoff schemes do not waste time on sensing as instantaneous wideband sensing is not required in this type, thus results decreased in the total service time [63] and of course handoff delay. However, the problem here is that the pre-selected goal channel(s) may no longer be accessible when interruption events occur. On the other hand, even though reactive-decision handoff schemes waste time on sensing for free channels, sensed results are more accurate and reliable. The problem here, this type needs handshaking time to reach agreement on the target channel between sender and receiver. In [48], a comparative study between the two types is provided.

In the following paragraphs, we will give an idea regarding the existing work in spectrum handoff with the concept of OSA. In such models, secondary users access the available wireless channels opportunistically. In general, when primary users appear in the licensed channels, the spectrum handoff procedure is initiated and the on-going transmission of the secondary users is paused until operating or when another channel becomes available for resuming the interrupted transmission, otherwise, interrupted transmissions will be queued or dropped.

Available spectrum bands for data transmissions compose of licensed spectrum bands and unlicensed spectrum bands as shown in Figure 2.5. According to [36, 42, 43], secondary users in future networks can access and operate on both the spectrum bands. However, most of the existing spectrum handoff schemes with OSA do not investigate into the performance of secondary users in a heterogeneous spectrum environment. This means that the existing models do not take into consideration that secondary users can operate on both of the spectrum band environments. In other words, they ignore the possibility of unlicensed bands to become available after some time and hence can be used for transmission.

In this research, an extensive study has been conducted to explore the existing spectrum handoff models with OSA from the point of view of the following features:

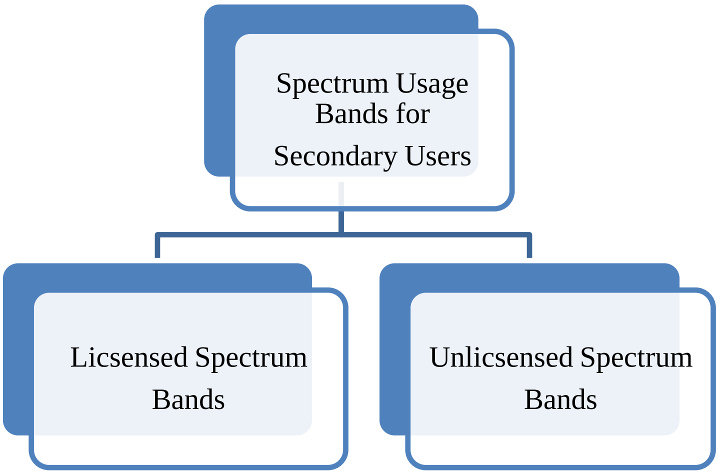

Figure 2.5: Spectrum handoff bands in CRNs.

- The spectrum handoff process: In the context of our research, handoff was mainly triggered by the arrival of the PUs which found their licensed bands occupied by unlicensed users [64] (which is different from handoff in [65]} which assumes handoff can happen before PUs arrival). In this case, the unlicensed users should immediately initiate the spectrum handoff procedure and empty the spectrum band for the licensed user and move to another idle band.
- Spectrum handoff decision timing: Selecting a target channel for spectrum handoff can be executed either proactively or reactively according to the adopted handoff scheme.
- The priority of the handoff secondary users: SUs can be served according to the arrival time, i.e. FCFS basis. On the other hand, it is possible to provide interrupted secondary users with a higher priority over new secondary users to utilise available spectrum channels in order to mitigate the adverse effect of the interruption/handoff process, and meanwhile improve the quality of service.

- Operating the spectrum environment: As it is known, available spectrum bands for wireless communications consist of licensed spectrum bands and unlicensed spectrum bands. As mentioned earlier, XG wireless networks should allow the unlicensed users to utilise both types of spectrum bands.

In this section, an overview regarding existing spectrum handoff models in cognitive radio networks has been provided from the above mentioned points of view. Our extensive conducted survey classifies the existing spectrum handoff models based on the spectrum operating environment. Therefore, two main groups are identified: Licensed spectrum band models and Heterogeneous spectrum band models.

Licensed spectrum band models: In this type of operating spectrum environments, only licensed channels can be used by either the primary users as operating channels, or by the secondary users as operating channels or as target channels in the case of spectrum handoff. Some examples of this type are [59], [52], [64], [58], [66], [53] and its extension [56], [67], [61, 62], [68], [69], [70], [71], [72], and [47].

A reactive-decision spectrum handoff scheme was presented in [59]. The proposed scheme selects the target channel for spectrum handoff based on on-demand wideband sensing after primary users return to their licensed bands. The characteristics of the spectrum usage behaviour between the primary users and the secondary users are modelled using a preemptive resume priority (PRP) M/G/1 queuing network model. The channel utilisation and transmission latency of SU transmissions are adopted as the performance metrics in this work. However, the presented model only operates over the licensed band.

In [52] and its extensions [64], a proposed proactive-decision handoff scheme assumes that each SU should prepare a list of target channels for spectrum handoff in case of the appearance of the primary users. Since the list of the target channels are known for any communicating secondary users, only the switching delay contributes to the handoff delay. The proposed scheme is modelled using a preemptive resume priority (PRP) M/G/1 queuing network model. The presented results show that this scheme can decrease the total service time comparing with the random channel selection scheme. However, in this work the existence of unlicensed channels is completely ignored and the operation channels were only selected from the licensed band. Moreover, in [66], a proactive-decision spectrum handoff scheme was presented in which guard channels were used as backup channels in the case of the appearance of the primary users. According to the FCC (Federal Communication Commission) new released rules, TV guard channels can be used for communication. In addition, cognitive radio users can be equipped with a TV channel database. These changes can eliminate the use of spectrum sensing in searching for idle channels. The suggested scheme is examined using a preemptive resume priority (PRP) M/G/1 queuing network model. Achieved results proved that the proposed proactive-decision scheme outperforms the existing random scheme in terms of handoff delay and total service time. The improvement in the total service time can be more than 20%.

The work in [67] presented a mathematical analysis for the suggested proactive-decision spectrum handoff model. The performance of the SUs is examined in terms of the link maintenance probability, the switching delay, the expected number of spectrum handoffs and the non-completion probability. However, the model considers the spectrum handoff only in the licensed bands.

In [61, 62] a new link maintenance model based on the reactive-decision spectrum handoff was proposed for CR systems. Interrupted secondary users perform spectrum sensing, spectrum selection and negotiate with the receiver to reach consensus on the target channel for spectrum handoff. In addition, a detailed performance analysis for the models was not presented. ‘Goodput’ is used as the performance measures to assess the proposed models.

In [68], a proactive spectrum handoff model was proposed. This model supposes that secondary users are synchronised to hop over channels by using an identical hopping order. When two communicating secondary users establish a link, they stop temporally channel hopping and stay at the same channel until completing their transmission. When primary users return back to their channels, both of the communicating secondary users hop to the next channel and complete the interrupted transmissions. Simulation results proved that the implemented model outperforms some conventional spectrum handoff schemes in terms of the throughput, the average service time, the number of collisions, and the packet delivery rate.

In [69], a proactive-decision spectrum handoff scheme was revealed. In this scheme, both of the secondary users (the transmitter and the receiver) arrange the available common target channels for spectrum handoff, based on longer idle time duration. If the primary user returns to its channel, the communicating secondary users will try to ‘handshake’ with the channel which has the longest idle time duration. If the handshake fails, they will try this with the next channel on the list, and so on. If all handshake trials fail, then the spectrum handoff process also fails. As a consequence, this will minimise the probability of spectrum handoff failure. Simulation results show that the proposed scheme outperforms the random channel selection scheme in terms of the average number of handshakes trials until a successful trial.

In [70], a new reactive-decision spectrum handoff scheme was developed using fuzzy logic and neural networks in order to achieve efficient spectrum decision. The implemented scheme focuses on moving secondary users. Spectrum handoff was mainly due to the users’ mobility and the appearance of the primary users in the operating channels. By using the aforementioned approach, the accuracy in the decision making for spectrum handoff situations was improved to 100%.

Besides the work presented in [70], [71] also proposed a spectrum handoff scheme that was based-on fuzzy logic. The proposed scheme can operate on both overlay and underlay access technologies. To avoid interference with the primary users, the transmission power of the secondary users was selected optimally and handoff decision making was incorporated with the knowledge of the transmission power, the data rate, and the average idle period. The scheme selects the channel with the highest expected idle period. Furthermore, the handoff process was mainly initiated by the appearance of the primary user, the unacceptable interference with the primary user transmissions, and when degrading the QoS of the secondary user. The scheme assumes information regarding the primary user activity (i.e., expected idle period) will be offered by, what is called, the “spectrum server” or by calculations achieved by the secondary user itself. Obtained simulation results showed that the suggested scheme can reduce the number of handoffs compared with the random selection scheme.

The work presented in [72] suggests an optimal handoff scheme. The proposed work also considers the moving secondary users. To reduce time to search for an idle channel in the case of the appearance of the primary user, the nearby secondary users will search for idle channels in a distributed, cooperative manner, and send searching results regarding primary user activity to the attended secondary user through the dedicated Common Control Channel (CCC). Traditionally, the secondary users decide that the target channel is based-on the following three factors:

- The mean idle time
- The probability of the channel idle
- Or, the expected transmission time

However, the proposed optimal handoff scheme combined the effect of all the individuals’ aforementioned factors. The provided simulation results show that the proposed scheme outperforms the conventional schemes, and can considerably reduce the switching handoff times.

In [47], the two main spectrum handoff schemes, reactive and proactive based on spectrum sensing, are optimised and compared. The scheme adopted the estimated spectrum handoff latency as a performance measure, a Preemptive Resume Priority (PRP) M/G/1 queuing model was provided to investigate which scheme is better under different traffic circumstances and sensing times.

However, in our previous work [53] and its extension [56], the proposed proactive-decision spectrum handoff schemes provided the handoff secondary users with a higher priority over the new secondary users to utilise the available spectrum channels, but the target channel for spectrum handoff is only chosen from the licensed spectrum channels. The former work is conducted through extensive simulation experiments and is modelled using a PRP M/M/1 queuing network model. Whereas, the latter work is provided with the analytical model which is modelled using a PRP M/G/1 queuing network model. Regardless of the operating spectrum environment, the achieved results show remarkable improvements to the associated delay performance measures, compared with existing schemes.

Additionally, in our work [58], a reactive spectrum handoff decision scheme is proposed in which interrupted secondary users have a higher priority over uninterrupted users to utilise the idle channel. In this scheme, interrupted secondary users do not use the sensing-based approach to search for backup channels in the case of the appearance of primary users in the operating channel. In other words, the sensing process is avoided in the new scheme in order to eliminate sensing delay from being included in the overall handoff delay. Instead, information regarding the instantaneous secondary users' queue lengths can be provided by a centralised database to help interrupted users to choose the channel with the shortest queue length at the moment of interruption, in order to resume unfinished transmissions.

When an interruption event occurs, interrupted user request information regarding channels states from the base station, after receiving the information requested the user will chose the channel with the shortest queue. Furthermore, a control channel is used to exchange handshaking RTS/CTS messages in order to reach consensus on the target channel between the communication SUs as explained in Section 2.9. Moreover, secondary users might choose to stay at the operating channel when the number of secondary users in all the channels is equal. The proposed scheme is modelled using a PRP M/M/1 queuing network model.

However, all of the above mentioned licensed spectrum band models do not consider the unlicensed part of the spectrum band in the communication process. Therefore, they neglect the effect of the unlicensed spectrum bands on the behaviour of the secondary users, which can reduce the number of spectrum handoffs and, consequently, reduce the associated delays.

Heterogeneous spectrum band models: In these models, licensed users can utilise only the licensed spectrum channels, whereas the unlicensed users can utilise both the unlicensed channels and, opportunistically, the vacant licensed channels temporally from licensed users. Some examples of this type are presented in [73, 74], and in their extensions [75] and [76], [77], and [78].

In [73, 74], and their extension [75], both licensed and unlicensed spectrum bands are used to improve the performance of the SUs. This has been done by allocating the licensed channels for initial and handoff usage, and the unlicensed channels as target channels for handoff. The proposed models in [73, 74] are analysed using the Markov chain technique and do not discuss the spectrum handoff process broadly. Instead, the system throughput, the blocking probability and the dropping probability of the secondary users were analysed. Whereas, [75] presents the spectrum handoff process comprehensively using a mathematical model to assess the performance of the proposed proactive-decision spectrum handoff model, in terms of the expected number of spectrum handoffs and the link maintenance probability.

In [76], a mixed environment of licensed and unlicensed spectrum channels were used. If the arrived secondary user (SU) senses that all the channels are busy, the secondary user will join the retrial group and retry for service later, or it will leave the system for good. The Markov chain is used to investigate the performance of the secondary users and matrix analysis is used to derive the steady state probabilities of the system. Also, this model is an extension of the work presented in [74]. However, the spectrum handoff process is not discussed extensively in this model. Instead, the model investigated the performance of the SUs thoroughly in terms of the loss of probability and throughput.

In [77], both the DSA with and without the queuing mechanism for the new/originating SUs are proposed and taken into consideration the random access environment. In this work, a Markov Chain technique is proposed to predict the performance of the suggested model in a mixed spectrum environment consisting of licensed and unlicensed spectrum bands. The proposed work considerers the interrupted probability, blocking probability and forced termination probability as the main performance measures to evaluate the work. However, the queuing mechanism is not used with interrupted secondary users and the spectrum handoff operation is not taken into consideration in this work.

In [78], only the simulation model for fair and efficient MAC protocols and for QoS provisioning was suggested. In this model, the performance of a CR user working with the primary users on a heterogeneous spectrum environment of licensed and unlicensed spectrum channels is investigated in terms of the normalised throughput, the access delay and the dropping rate.

However, to the best of the author’s knowledge, existing work in the licensed spectrum band and in the heterogeneous spectrum environment do not differentiate between handoff and new/originating secondary users as no priority has been given to handoff users in getting services, which leads to increase in the handoff delay and total service time of interrupted users. Furthermore, most previous work did not broadly discuss the issue of spectrum handoff operations in terms of the spectrum handoff delay; instead, previous work focuses on other performance measures such as system throughput, the blocking probability and the dropping probability of the SUs.

Chapter 3.

SPECTRUM HANDOFF MODELLING UNDER LICENCED SPECTRUM ENVIRONMENT

3.1 **Introduction**

Spectrum handoff plays a critical role in cognitive radio networks as it provides a reliable transmission for interrupted secondary users when the primary users return to their spectrum, and also helps the secondary users to resume their unfinished transmissions, either at the same channel or at another vacant channel. This can guarantee smooth and fast switching which leads to minimised performance degradation during a spectrum handoff [5].

Repeated spectrum handoffs can negatively affect the QoS for interrupted users by increasing the handoff delay and the total service time. Giving priority to handoff (interrupted) users over new (uninterrupted) users can show significant performance improvements. In some traditional wireless systems, [79-83] on-going (handoff) calls are assigned or given priority over originating (new) calls, since it is much less desirable and less tolerable to force the termination of calls in progress than to block calls which are yet to be connected. In this thesis, we borrow liberally from traditional wireless systems to argue that there should be a high priority level assigned to handoff (interrupted) users over new (uninterrupted) users. By giving such priority to the handoff users, this means the handoff users will be served before any uninterrupted users. For example, in real time applications, there is an acceptable limit on the handoff delay and any transmission exceeding this limit will be dropped which can degrade the QoS. But, by adopting our prioritising principal, this can compensate the handoff delay resulting from multiple interruptions and consequently reduced the handoff delay which leads to improvement on the QoS.

3.1.1 Switching and Non-switching Spectrum Handoff

In the context of cognitive radio networks, a handoff means the transition of a spectrum from a low priority user (secondary user) to the spectrum's owner (primary user). However, in cognitive radio networks, the term "handoff" does not necessary indicate spectrum switching. Hence, two types of spectrum handoffs are identified in CRNs, namely: switching spectrum handoff, and non-switching spectrum handoff, as illustrated in Figure 3.1.

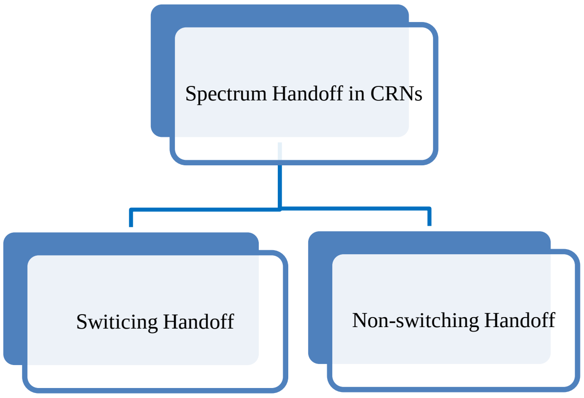

Figure 3.1: Handoff process in CRNs.

This chapter proposes and implements queuing network models to investigate the effects of repeated spectrum handoff delays on the total service time in cognitive radio networks. This work is an extension of our previous works [53]. The implemented schemes are validated against a simulation and are compared with existing handoff schemes through extensive simulation experiments.

From the author's knowledge, existing work does not give priority to interrupted secondary users over uninterrupted ones, with respect to transmission resumption in the new channel. However, existing work gives such priorities in cases where interrupted users choose to wait (stay) at the operating channel to resume their unfinished transmissions. This means that interrupted users who decide to change their operating channel will have to wait in a queue until all primary users receive their services. Furthermore, interrupted secondary users are subjected to join the tail of the secondary users' queue of the new channel; of course this will incur extra delay to interrupted users and will increase their handoff delay and total service time. However, by giving higher priority to interrupted secondary users in order to utilise the idle channels, the handoff delay and the total service time can be reduced. This, in turn, will improve the QoS of the interrupted secondary users.

3.2 Spectrum Decision Handoff Modelling

In cognitive radio networks, the spectrum mobility function aims to help the secondary users select the best channel(s) to send and receive their data in the case of spectrum handoff. Up till now, there has been limited attention given to the performance analysis of spectrum mobility in CR networks using analytical models. This is in light of the importance of analytical modelling for performance analysis, and its ability to provide a useful interpretation of the process of spectrum mobility.

There have been several earlier studies conducted to evaluate the total service time of SU transmissions in cognitive radio networks. In [52] and its extensions [64] and [84], a preemptive resume priority (PRP) M/G/1 queuing model is proposed for proactive-decision spectrum handoff. In [84-86], the interrupted communication of secondary users on a particular channel is resumed on the same channel when the channel becomes idle. Conversely, although interrupted secondary users in all existing models that allow interrupted users to change their operating channel after an interruption event occurs, such as [66], [64], no priority is considered for interrupted users over the existing secondary users to resume unfinished transmissions in the new target channel. In this case, the most common disadvantage is the delay resulting from the repeated spectrum handoffs.

3.3 Proposed Spectrum Handoff Model

3.3.1 System Model

In this model, we adopted a cognitive radio network with a time-slotted system as shown in Figure 3.2. Each user divides its data into equal-sized time slots. Each time-slot is divided into two parts; the first part is for spectrum sensing and the second is for data transmission. When a secondary user starts transmitting in a channel, this channel must be sensed (monitored) periodically in every time slot. If a SU senses, during the first part of the slot, that the present channel is idle, transmission will be commenced in the second part of the slot. However, in general, if the current channel is busy, a spectrum handoff procedure must be performed to help the interrupted user to resume the unfinished transmissions in an appropriate channel.

The proposed system model suggests that primary users and secondary users will compete for utilising spectrum channels using access points (APs) for both uplink and downlink transmissions, as illustrated in Figure 3.3. The model consists of two independent wireless channels, as shown in Figure 3.4. Each wireless channel is comprised of three priority queues: low-priority queue (mainly for secondary users' transmissions), high-priority queue (for primary users' transmissions), and handoff (HD) queue (for interrupted users' transmissions).

The queues are modelled with infinite length for simplicity. In addition, a Preemptive Resume Priority (PRP) M/G/1 queuing model is proposed to manage the spectrum handoff procedure.

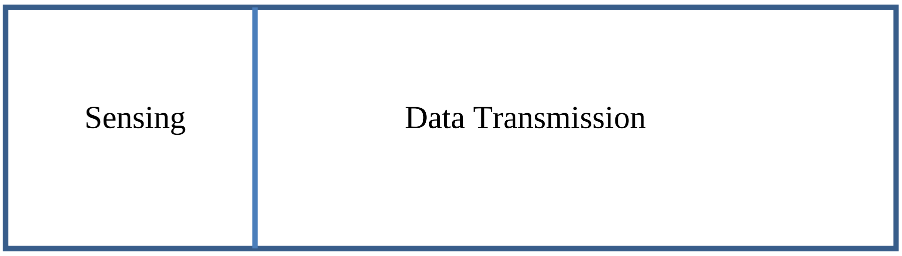

Figure 3.2: Time slot structure of the secondary networks.

PUs will preempt the transmission of secondary users at the moment of their arrival if they find that their channels are being used by secondary users. The interrupted SU will pause its transmission in the operating spectrum immediately and continues in another available spectrum channel. The preemption priority queue characterises the inherent traffic structure in CRNs as it gives a right to spectrum owners (PUs) to interrupt the secondary users' transmissions at the time of their arrival. Based on this model, we proposed prioritised proactive spectrum handoff decision schemes. This scheme gives higher priority to interrupted secondary users to utilise idle channels over existing uninterrupted SUs, which improves the handoff delay and the total service times.

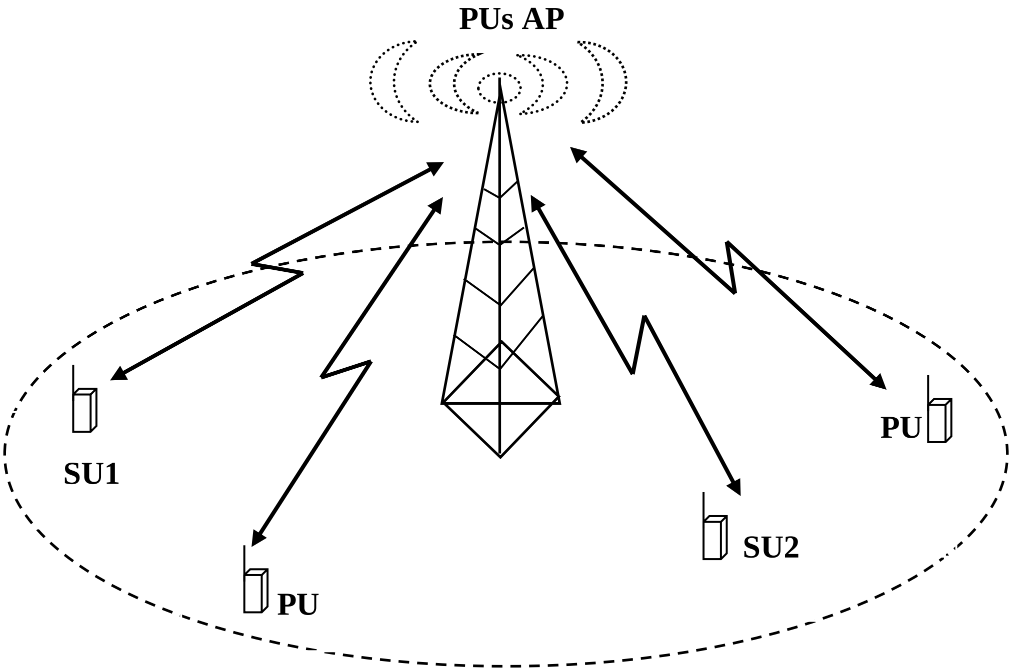

Figure 3.3: Proposed CR network scenario.

Some preliminary properties for the PRP M/G/1 queuing network model are listed below:

- The low-priority (secondary) users and the high-priority (primary) users will arrive at their default channel, say *k*, according to the Poisson processes with mean rates of $\lambda_s^{(k)}$ and $\lambda_p^{(k)}$, respectively. If the channel is busy, then the arrived users will wait in their corresponding queues until it becomes idle. In addition, their service times are generally distributed with mean of $E\left[X_s^{(k)}\right]$ and $E\left[X_p^{(k)}\right]$, respectively.

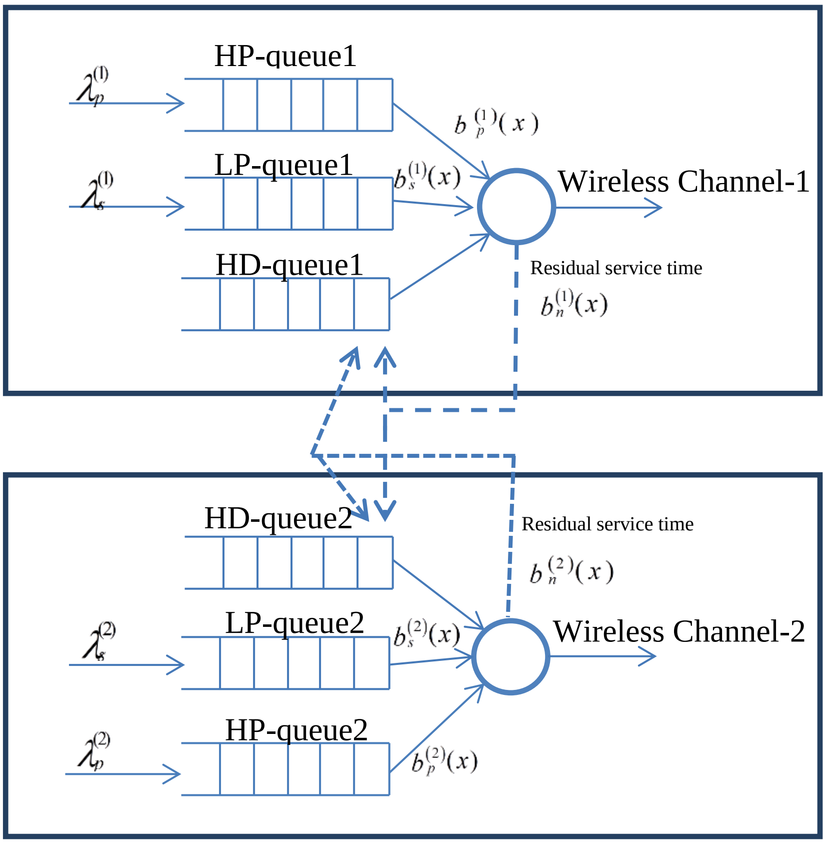

Figure 3.4: Proposed (PRP) M/G/1 queuing network model where n denotes n^{th} interruption.

- Primary users have a higher priority over secondary users to utilise the two wireless channels. Therefore, primary users will interrupt (preempt) the secondary users' transmission when they arrive and find their default channels being used by secondary users. Within the primary users' class, PUs will compete to utilise the default frequency channels on the basis of the first-come-first-served (FCFS) scheduling algorithm at each channel. Handoff secondary users will be served after all primary users, and before any other uninterrupted secondary users already waiting in the low-priority queue of the channel.

- Interrupted users will arrive at their target channel, say *k*, according to the Poisson process, with a mean rate of $\lambda_i^{(k)}$ and an effective transmission time with a mean $E\left[X_i^{(k)}\right]$, where $i \geq 1$.
- Interrupted users will stay or change their operating channel depending on the adopted spectrum handoff scheme.
- Interrupted secondary users will be put into the handoff-priority queue in the target channel and will be served on a FCFS basis before any uninterrupted secondary users.

3.3.2 Examples for Various Spectrum Handoff Decision Schemes

To investigate the performance of the proposed proactive spectrum handoff decision schemes, we presented other various existing schemes and compared them.

In this section, the effect of multiple spectrum handoffs in Total Service Time (TST) and Handoff Delay (HD) will be explained through the timeline illustrated in the next figures. TST is defined as the period from the moment of launching the transmission up to the point of completing the transmission [52, 64, 84]. Whereas the HD is defined as the period from the point of pausing the transmission until the moment of resuming the transmission [47]. In the case of an interruption, the spectrum handoff procedure will be initiated immediately. However, the interrupted user will choose the target channel according to one of the following handoff schemes to resume unfinished transmissions:

3.3.2.1 Non-switching-handoff Spectrum Decision Scheme

In this type of scheme (also known as a Non-handoff scheme), the interrupted secondary user will stop transmitting, stay in the original channel at the interrupted secondary user's queue, and wait until the primary user finishes transmitting all of its data [64]. This technique is like the non-hopping approach of IEEE 802.22 [57, 59].

Figure 3.5 illustrates spectrum handoffs which are described in detail as follows:

- Initially, a secondary user SU1 starts to transmit its data on its default channel CH_n to SU2, where n=1, 2 ...
- After this, when a primary user arrives at the same channel, because it is its default channel, it will interrupt the transmission of SU1, and then the non-switching-handoff procedure will be initiated. SU1will stay at the same channel and join the head-of-line (HOL) of the SUs' queue to resume unfinished transmissions after the primary user finishes its transmission. However, some other primary users may arrive during the period of the waiting time of SU1. Upon completion of the primary users' transmission, SU1 will immediately resume its unfinished transmission. Hence, handoff delay here is the waiting time for interrupted users, which is exactly equal to the busy period (Y_p) resulting from the primary users in the same channel.
- After each interruption, the same handoff procedure will be repeated until SU1 finishes its data transmission.

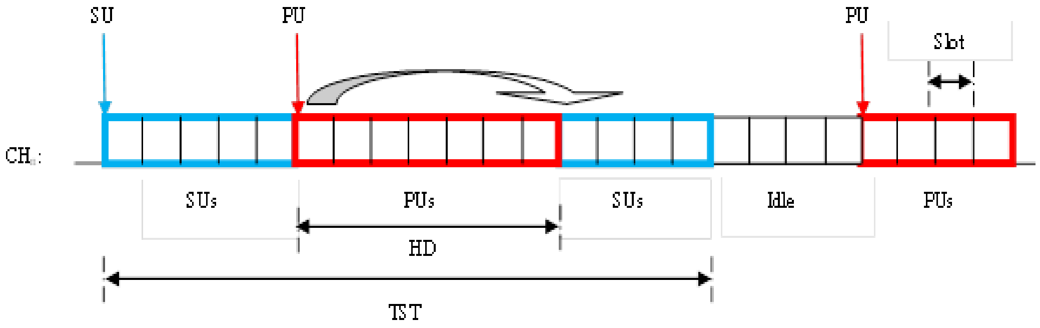

Figure 3.5: Non-switching handoff process.

3.3.2.2 Switching-handoff Spectrum Decision Scheme

In this switching scheme (also known as handoff scheme), whenever the secondary user faces interruption, it will have to change its operating channel until it has finished transmitting all of its data [64]. Clearly, successive channel switching would increase the average handoff delay for the interrupted users, which could degrade the quality of service of secondary users. This is because handoff users might find the others channels busy and they have to wait at their queues before they can resume their service again, consequently, this will increase the handoff delay as well. The switching-handoff procedure described in this process is shown in Figure 3.6:

- In the beginning, SU1 initiates data transmission in its default channel CH1 to SU2.
- When the first interruption occurs, the switching-handoff procedure will be initiated to change the operating channel to CH2. Since, CH2 is idle, SU1 will resume transmission immediately. In this case, the handoff delay is just the channel switching delay (T_{sw}).

- In the second interruption, the target channel CH1 is busy. As discussed earlier, by applying our prioritised principle, SU1 can only get service if all the other primary users in the primary users' queue of CH1 and any other previously interrupted users have been served. However, the old switching-handoff scheme [52, 64] does not differentiate between interrupted and uninterrupted secondary users, as it serves them in a FCFS fashion which increases the handoff delay and the total service time of the interrupted users. In both models, the handoff delay is the sum of the switching delay (T_{sw}) and the waiting delay (W_s).
- This procedure will be continued until SU1 finishes sending its data. Here the operating channel will be changed continuously between CH_1 and CH_2, which significantly increases the handoff delay and the total service time of the interrupted secondary users; especially at high PUs arrival rates. This effect can be minimised by giving a higher priority to the interrupted users to finish their transmission before any of the other uninterrupted secondary users in the new channel. By introducing this prioritised principle, perhaps unsurprisingly, the QoS of the interrupted users can be improved.

3.3.2.3 Random-handoff Spectrum Decision Scheme

In this scheme, the spectrum handoff procedure will randomly select a target channel for the spectrum handoff from available channels [52]. Since two wireless channels are assumed in this work, there is equal probability (50%) to choose either to stay at the same transmission channel or to change to the other channel. However, the handoff delay and the total service time can be calculated in the same way as in the previous handoff schemes. In fact, the proposed prioritised principle will give a higher priority to interrupted SUs to resume transmission as explained in the switching-handoff procedure.

In this scheme it is assumed that the sender (SU1) will generate a list of target channel sequences randomly for spectrum handoff and send it to the receiver (SU2) during the connection setup time. As a result, there will be no need for any handshaking time (T_{ha}), which is the time taken to reach consensus on the target channel between SU1 and SU2.

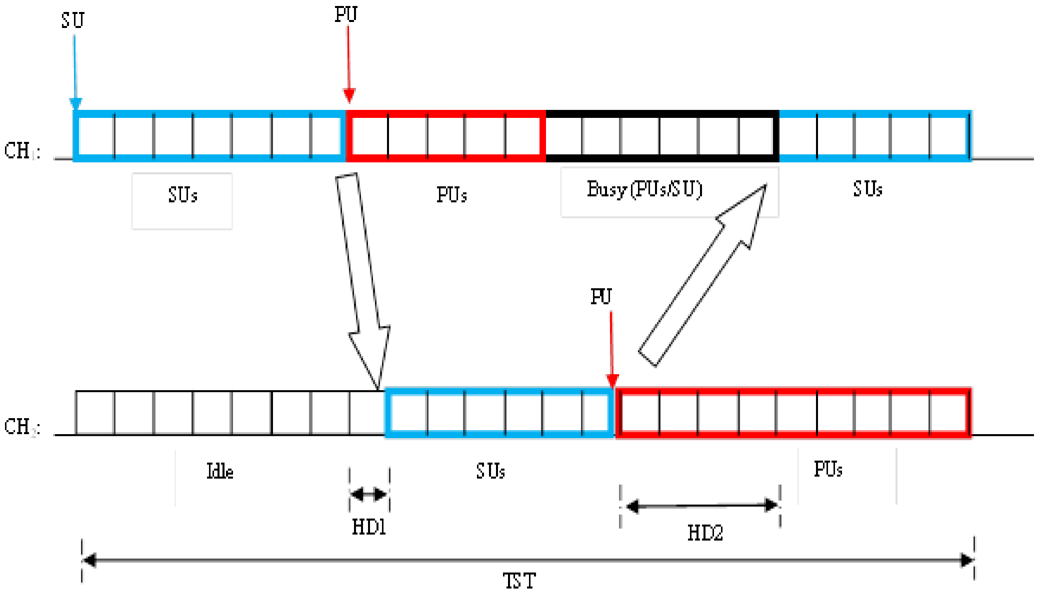

Figure 3.6: Switching-handoff process.

3.3.2.4 Reactive-handoff Spectrum Decision Scheme

In [57, 59], the spectrum sensing delay refers to the time it takes, such that the interrupted secondary user finds an idle channel for transmission after an interruption event occurs. It is clear that the sensing delay plays a major role in this type of handoff scheme since it increases the handoff delay and the total service time.

In this type of scheme, the interrupted users will perform wideband sensing for some time, say T_{se}, to search for candidate idle channels. If they find more than one idle channel, then the handoff procedure will randomly select one out of those channels to resume the transmission. In the case where there are no idle channels, interrupted SUs will have to wait at the head-of-line of the SUs' queue of their operating channel until the channel becomes idle [57, 59]. Indeed, sensing delay is directly proportional to the number of candidate wireless channels to be sensed, i.e., if it takes c time units to sense one wireless channel, we need nc time unites for sensing n channels. However, when a secondary user senses a small number of candidate channels, this can lead to a minimised total service time. On the other hand, it is more difficult to find a free channel when sensing a small number of channels and consequently the handoff delay as well as the total service time could be increased [87, 88]. Perhaps it is also worth considering the handshaking time (T_{ha}) to reach consensus on the target channel between the communicating SUs. Essentially, these delays should be added on to the previous delays mentioned in order to evaluate the total service time.

In general, the sum of sensing time (T_{se}), handshaking time (T_{ha}) and switching time (T_{sw}) is known as the total processing time. Since the interrupted secondary users may stay or change their operating channel depending on other channels' conditions, two types of processing time can be defined. The first type is $T_{pr\text{-}stay}$, which is associated with the 'stay' case, and the second is $T_{pr\text{-}change}$, which is associated with the 'change' case:

$$T_{pr-stay} = T_{se} + T_{ha} \qquad 3.1$$

$$T_{pr-change} = T_{se} + T_{ha} + T_{sw} \qquad 3.2$$

Equation 3.2 implies switching time as the interrupted secondary users change their operating channels. If the switching time is assumed to be zero, then the total processing time (T_{pr}) is:

$$T_{pr} = T_{pr-stay} = T_{pr-change} \qquad 3.3$$

In this reactive model of the spectrum handoff decision, the prioritised principle is not applicable because the interrupted user will only change its operating channel to an idle target channel. For the other types of spectrum handoff schemes, the handshaking time does not exist because the target channel for spectrum handoff is already determined before the communication starts between the intended secondary users. In other words, both communicating SUs know in advance the target channel for resuming the transmission when an interruption event occurs, thus, they switch together to the target channel without any handshaking time.

3.4 Spectrum Handoff Analytical Modelling

We apply the preemptive resume priority (PRP) M/G/1 queuing network model shown in Figure 3.4 to derive the closed form expression for the total service time for the newly implemented (switching-handoff and random-handoff) schemes.

In these models, the effective transmission time is an important parameter. It is defined as the time from starting the transmission or resuming communication until the time an interruption event occurs. In addition, let $\lambda_s^{(k)}$ and $\lambda_p^{(k)}$ be the initial arrival rates of the secondary users' and the primary users' connections to their default wireless channel *k*, respectively; both arrival rates are modelled by the Poisson processes. Also, let $b_s^{(k)}(x)$ and $b_p^{(k)}(x)$ be their service time distributions, with means $E\left[X_s^{(k)}\right]$ and $E\left[X_p^{(k)}\right]$, respectively. We take into consideration the effect of the traffic load of the interrupted secondary users coming from other wireless channels on each channel, i.e. a secondary user with *i* interruptions ($i \geq 1$) will arrive to target channel *k* with rate $\lambda_i^{(k)}$ and effective transmission time $b_i^{(k)}(x)$ with a mean $E\left[X_i^{(k)}\right]$, to resume its unfinished transmission. Note that, SU's parameters with zero interruptions ($i = 0$) are denoted with $\lambda_s^{(k)}$, $E\left[X_s^{(k)}\right]$, etc.

The detection of newly arrived primary users is assumed to be perfect, which means no false alarms. In addition to this, there is an infinitesimal delay for the SU to terminate transmission so that the existence of the secondary users is virtually transparent to that of the primary users.

The primary and secondary users' utilisation factors are defined as:

$$\rho_p^{(k)} = \lambda_p^{(k)} E\left[X_p^{(k)}\right] \tag{3.4}$$

and

$$\rho_i^{(k)} = \lambda_i^{(k)} E\left[X_i^{(k)}\right] \tag{3.5}$$

respectively, and the aggregate system utilisation is given by:

$$\rho^{(k)} = \rho_p^{(k)} + \sum_{i=0}^{\infty} \rho_i^{(k)} \qquad 3.6$$

For simplicity, we suppose that all the channels are identical and have the same traffic parameters. So, dropping the notation (*k*) in all system parameters yields:

$$\rho = \rho_p + \sum_{i=0}^{\infty} \rho_i \qquad 3.7$$

The necessary and sufficient conditions for system stability are:

$\rho < 1$, $\sum_{i=0}^{\infty} \rho_i < 1$ and, $(\rho_p + \sum_{i=0}^{\infty} \rho_i) < 1$

Let us assume that $E[X_s]$ is the average service time, $E[D]$ is the average handoff delay, and $E[N]$ is the average number of interruptions for a secondary user's connection during a period of $E[X_s]$. The total estimated service time of the secondary users is defined [52, 64] as:

$$E[T] = E[X_s] + E[N]\,E[D] \qquad 3.8$$

The second term in 3.8 refers to the average cumulative handoff delay ($E[D_{cum}]$), i.e.,

$$E[D_{cum}] = E[N]\,E[D] \qquad 3.9$$

Thus,

$$E[T] = E[X_s] + E[D_{cum}] \qquad 3.10$$

where, $E[N]$ can be expressed as:

$$E[N] = \lambda_p\, E[X_s] \qquad 3.11$$

3.4.1 New Switching-handoff Model

Let Q_p be the mean number of primary users in the high priority queue, and Q_i be the mean number of secondary users in the interrupted SUs' priority queue with i interruptions ($i \geq 1$). Then the mean waiting time (W_s) that the interrupted secondary users have to wait before receiving the service can be expressed as follows:

$$W_s = R_s + \sum_{i=1}^{\infty} Q_i E\left[X_i\right] + Q_p E\left[X_p\right] + \lambda_p W_S E\left[X_p\right] \qquad 3.12$$

The first term in Equation 3.12 (R_s) represents the residual service time of the user (primary or secondary) in service upon a secondary user's arrival, whereas the second term represents the service time obtained due to existing interrupted secondary users with i interruptions in the interrupted users' queue. The third term describes the service time resulting from the primary users in the high-priority queue. Finally, the last term represents the time that a secondary user has to wait due to the primary users' arrival before commencing its service. The next steps will derive the first two parts of Equation 3.12 one by one.

The first term R_s can be derived as follows:

$$R_s = \frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{1}{2}\sum_{i=0}^{\infty}\lambda_i E\left[\left(X_i\right)^2\right] \qquad 3.13$$

Where, the first term of Equation 3.13 represents the residual service time for the primary user and the second term represents the residual service time for the secondary users with i interruptions.

From [52, 64], $E\left[(X_i)^2\right]$ and λ_i (both based on exponential distribution) can be expressed as:

$$E\left[(X_i)^2\right]=\frac{2}{(\lambda_p+\mu_s)^2} \qquad 3.14$$

$$\lambda_i=\lambda_s\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)^i \qquad 3.15$$

Substituting Equations 3.14 and 3.15 into Equation 3.13 yields:

$$R_s=\frac{1}{2}\lambda_p\,E\left[(X_p)^2\right]+\frac{\lambda_s}{(\lambda_p+\mu_s)^2}\sum_{i=0}^{\infty}\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)^i \qquad 3.16$$

We assume the service time of the secondary users and the primary users follow the exponential distribution i.e., $b_s^{(k)}(x)=\mu_s\,e^{-\mu_s x}$ and $b_p^{(k)}(x)=\mu_p\,e^{-\mu_0 x}$ with the mean $\mu_s=1/E[X_s]$ and $\mu_p=1/E[X_p]$, respectively. As a consequence the remaining service time of the interrupted secondary users' connection also follows the identical exponential distribution.

For simplicity we assume that:

$$\frac{\lambda_p}{\lambda_p+\mu_s}=C \qquad 3.17$$

Then, substituting the expression 3.17 into Equation 3.15 yields

$$\lambda_i=\lambda_s\,C^i \qquad 3.18$$

Thus, R_s can be rewritten as:

$$R_s=\frac{1}{2}\lambda_p\,E\left[(X_p)^2\right]+\frac{\lambda_s}{(\lambda_p+\mu_s)^2}\sum_{i=0}^{\infty}C^i \qquad 3.19$$

Using the well-known series:

$$\sum_{i=0}^{\infty} C^{i} = C^{0} + C^{1} + C^{2} + C^{3} + \ldots\ldots = \frac{1}{1-C} \qquad 3.20$$

Again, we can rewrite R_s as follows:

$$R_s = \frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\lambda_s}{\left(\lambda_p + \mu_s\right)^2} \cdot \frac{1}{1-C} \qquad 3.21$$

After some simplifications we get:

$$R_s = \frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\lambda_s}{\mu_s\left(\lambda_p + \mu_s\right)} \qquad 3.22$$

and, using $\rho_s = \lambda_s/\mu_s$, we have:

$$R_s = \frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\rho_s}{\left(\lambda_p + \mu_s\right)} \qquad 3.23$$

According to Little’s law, Q_i in the second term of 3.12 can be expressed as:

$$\sum_{i=1}^{\infty} Q_i E\left[X_i\right] = \sum_{i=1}^{\infty} W_s \lambda_i E\left[X_i\right] \qquad 3.24$$

where,

$$Q_i = W_s \lambda_i \qquad \text{where } i \geq 1 \qquad 3.25$$

and W_s is the mean waiting time of the secondary users.

By substituting Equation 3.15 into 3.24 we get:

$$\sum_{i=1}^{\infty} Q_i E\left[X_i\right] = W_s \lambda_s \sum_{i=1}^{\infty} \left(\frac{\lambda_p}{\lambda_p + \mu_s}\right)^{i} E\left[X_i\right] \qquad 3.26$$

According to [52, 64], $E[X_i]$ is determined as:

$$E\left[X_i\right]=\frac{1}{\left(\lambda_p+\mu_s\right)} \tag{3.27}$$

Thus,

$$\sum_{i=1}^{\infty} Q_i\, E\left[X_i\right]=\frac{W_s\,\lambda_s}{\left(\lambda_p+\mu_s\right)}\sum_{i=1}^{\infty}\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)^i \tag{3.28}$$

Since C = $\lambda_p/(\lambda_p + \mu_s)$

$$\sum_{i=1}^{\infty} Q_i\, E\left[X_i\right]=\frac{W_s\,\lambda_s}{\left(\lambda_p+\mu_s\right)}\sum_{i=1}^{\infty} C^i \tag{3.29}$$

Using the series:

$$\sum_{i=1}^{\infty} C^i = C^1+C^2+C^3+\ldots\ldots=\frac{C}{1-C} \tag{3.30}$$

Therefore,

$$\sum_{i=1}^{\infty} Q_i\, E\left[X_i\right]=\frac{W_s\,\lambda_s}{\left(\lambda_p+\mu_s\right)}.\frac{C}{1-C} \tag{3.31}$$

After some manipulations we can derive the following formula:

$$\sum_{i=1}^{\infty} Q_i\, E\left[X_i\right]=\frac{W_s\,\rho_s\lambda_p}{\left(\lambda_p+\mu_s\right)} \tag{3.32}$$

By substituting the value of Q_p obtained from [52, 64] in 3.12, the third term of 3.12 can be rewritten as:

$$Q_p E\left[X_p\right]=\frac{\lambda_p^2\, E\left[\left(X_p\right)^2\right]}{2\left(1-\rho_p\right)} E\left[X_p\right] \tag{3.33}$$

By substituting 3.23, 3.32 and 3.33 into 3.12 we can get:

$$W_s = \frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\rho_s}{\left(\lambda_p + \mu_s\right)} + \frac{\lambda_p^2 E\left[\left(X_p\right)^2\right]}{2\left(1-\rho_p\right)} E\left[X_p\right]$$

$$+\frac{W_s \rho_s \lambda_p}{\left(\lambda_p + \mu_s\right)} + \lambda_p W_s E\left[X_p\right] \qquad 3.34$$

Where, $\lambda_p E[X_p] = \rho_p$, after some simplifications, 3.34 can be rewritten as:

$$W_s = \frac{\frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\rho_s}{\left(\lambda_p + \mu_s\right)} + \frac{\lambda_p^2 E\left[\left(X_p\right)^2\right]}{2\left(1-\rho_p\right)} E\left[X_p\right]}{\left(1-\rho_s\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)-\rho_p\right)} \qquad 3.35$$

The handoff delay in this case is the sum of the waiting delay and the channel switching delay:

$$E\left[D\right] = W_s + T_{sw} \qquad 3.36$$

Or,

$$E\left[D\right] = \frac{\frac{1}{2}\lambda_p E\left[\left(X_p\right)^2\right] + \frac{\rho_s}{\left(\lambda_p + \mu_s\right)} + \frac{\lambda_p^2 E\left[\left(X_p\right)^2\right]}{2\left(1-\rho_p\right)} E\left[X_p\right]}{\left(1-\rho_s\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)-\rho_p\right)} + T_{sw} \qquad 3.37$$

Also, putting Equation 3.37 into Equation 3.9, cumulative handoff delay can be estimated as:

$$E[D_{cum}]=E[N]\Bigg[\frac{\frac{1}{2}\lambda_p E\big[(X_p)^2\big]+\frac{\rho_s}{(\lambda_p+\mu_s)}+\frac{\lambda_p^2 E\big[(X_p)^2\big]}{2(1-\rho_p)}E[X_p]}{\left(1-\rho_s\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)-\rho_p\right)}+T_{sw}\Bigg] \qquad 3.38$$

In general, the total service time is expressed as:

$$E[T]=E[X_s]+E[N](W_s+T_{sw}) \qquad 3.39$$

Finally, the total service time can be found by substituting Equation 3.35 into Equation 3.39 as:

$$E[T]=E[X_s]+E[N]\Bigg[\frac{\frac{1}{2}\lambda_p E\big[(X_p)^2\big]+\frac{\rho_s}{(\lambda_p+\mu_s)}+\frac{\lambda_p^2 E\big[(X_p)^2\big]}{2(1-\rho_p)}E[}{\left(1-\rho_s\left(\frac{\lambda_p}{\lambda_p+\mu_s}\right)-\rho_p\right)}+T_{sw}\Bigg] \qquad 3.40$$

The proposed spectrum handoff scheme belongs to the proactive decision approach, due to the fact that target channel for resuming the transmission in case of appearing of the PU is known in advance, since we assume only two channels in this model.

3.4.2 Random Handoff Model

Considering the switching-handoff scheme and the non-switching-handoff scheme, [52] determines the total service time for random handoff as follows:

$$E[T] = E[X_s] + \frac{E[N]}{2} Y_p + \frac{E[N]}{2} (W_s + T_{sw}) \qquad 3.41$$

Where, Y_p is the primary users' busy period in each wireless communication channel.

In this type of spectrum handoff, a target channel for resuming interrupted transmission will be selected uniformly among available channels. In fact, this formula can be applied to compute the total service time of our new random model by substituting the value of W_s derived in Equation 3.35 into Equation 3.41. The only difference between the two models resides in the queue disciplines, as we proposed the priority principle, and definitely this will not affect the way of calculating of $E[T]$.

In this scheme, we suppose that the sender sends a random selected list of the target channels for spectrum handoffs to the receiver in the connection establishment time (proactive-decision).

3.5 Simulation and Numerical Results

In this section we present the simulation results that have been achieved using the discrete event simulator (MATLAB) tool to analyse the cumulative handoff delay. Table 3.1 summarises various implemented handoff models with corresponding features. A summary of simulation parameters are shown in Table 3.2.

Model-Name	**Symbol**	**Decision's Behavior**
Non-switching-handoff	NSWH	Proactive
Old Switching-handoff	SWH-OLD	
New Switching-handoff	SWH-NEW	
Old Random-handoff	RAH-OLD	
New Random-handoff	RAH-NEW	
Reactive-handoff	REH	Reactive

Table 3.1: Implemented handoff models.

Parameter	**Symbol**	**Value(s)**
PU arrival rate	λ_p	0.05-0.30
SU arrival rate	λ_s	0.15
PU service rate	μ_p	0.60
SU service rate	μ_s	0.40

Table 3.2: Simulation parameters.

The presented simulation results cover a high range of channel utilisation (up to ~90%) according to the simulation parameters shown in Table 3.2.

3.5.1 Simulation Setup

In order to evaluate the performance of the proposed handoff schemes, we carried out an extensive number of simulation experiments for different PUs arrival rates and PUs and SUs service rates. We considered a cognitive radio system with two wireless channels and each of these wireless channels is assumed to be collision-free. We neglected the effect of T_{sw} and T_{ha}. In this thesis, a 95% confidence interval (which mean that 95% of simulation results will be situated in its range) is used to evaluate the accuracy of the achieved results. The simulation is executed 6 times and in each simulation round 2 million arrival events are generated. The average of those simulation results is used to draw the graphs. As a result the accuracy in drawing the graphs is very high and the variations are not visible because they are very small in most graphs. In addition to this, we assumed that secondary users had the ability to perfectly sense the available spectrum bands, meaning that the detection of PUs is perfect.

3.5.2 Performance Calculations

In general, for simulation experiments the average total service time *E[T]* can be calculated for each wireless channel using the following formula

$$E\left[T\right]=E\left[X_s\right]+E\left[N\right]*\frac{(\sum Handoff\ Delays)}{Number\ of\ \ Interruptions}+T_{sw}+T_{pr} \qquad 3.42$$

The average handoff delay *E[D]* is just:

$$E\left[D\right]=\frac{(\sum Handoff\ Delays)}{Number\ of\ \ Interruptions}+T_{sw}+T_{pr} \qquad 3.43$$

And the average number of interruptions *E[N]* can be defined as:

$$E\left[N\right]=\frac{Number\ of\ \ Interruptions}{Number\ of\ \ SUs\ Arrivals} \qquad 3.44$$

It is worth noting that, here, the term T_{pr} in Equations 3.42 and 3.43 is a general term and should be defined carefully and separately for each of the spectrum handoff schemes. For example, in proactive handoff schemes, the sensing delay should be zero. Alternatively, the switching delay in a non-switching handoff scheme does not exist at all. In order to achieve credible simulation results, the average statistics of the two channels have been taken in order to draw the figures.

3.5.3 Numerical Results

The numerical results presented in Figure 3.7 to Figure 3.14 show comparisons between analytical and simulation results. In general, from the graphs, it is clear that the analytical and simulation results are approximately the same in the case of NSWH, SWH-OLD, RAH-NEW, and REH schemes. On the other hand, the remaining schemes (SWH-NEW and RAH-OLD) give very close results, especially at low PUs arrival rates of about (0.05-0.20) and (0.05-0.15), respectively. Above these ranges, the difference between the two curves increases as the PUs' arrival rate increases.

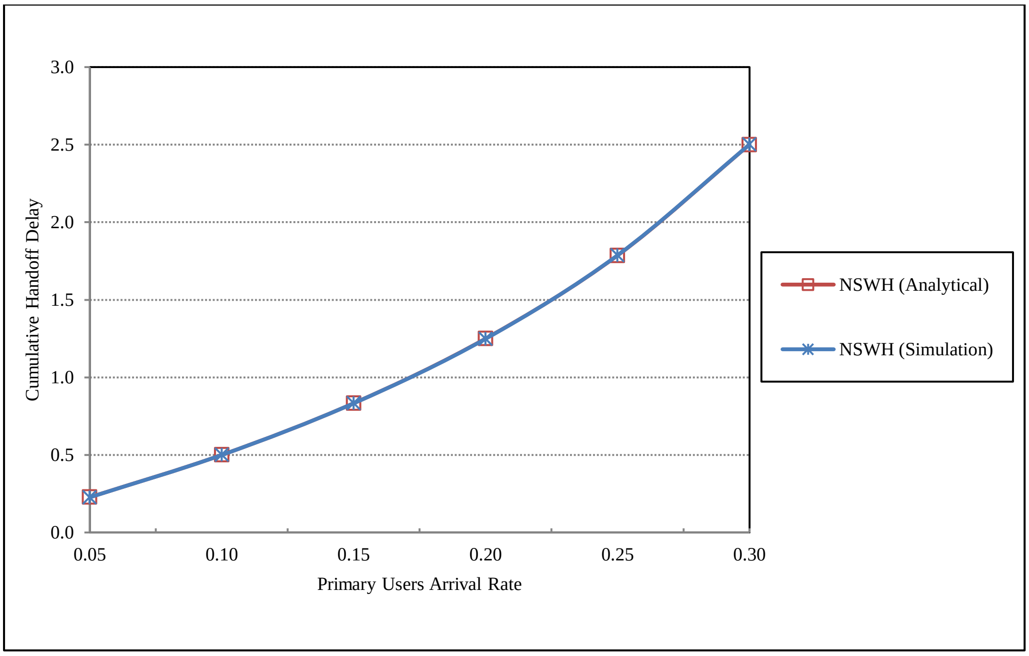

Figure 3.7: Non-switching handoff scheme.

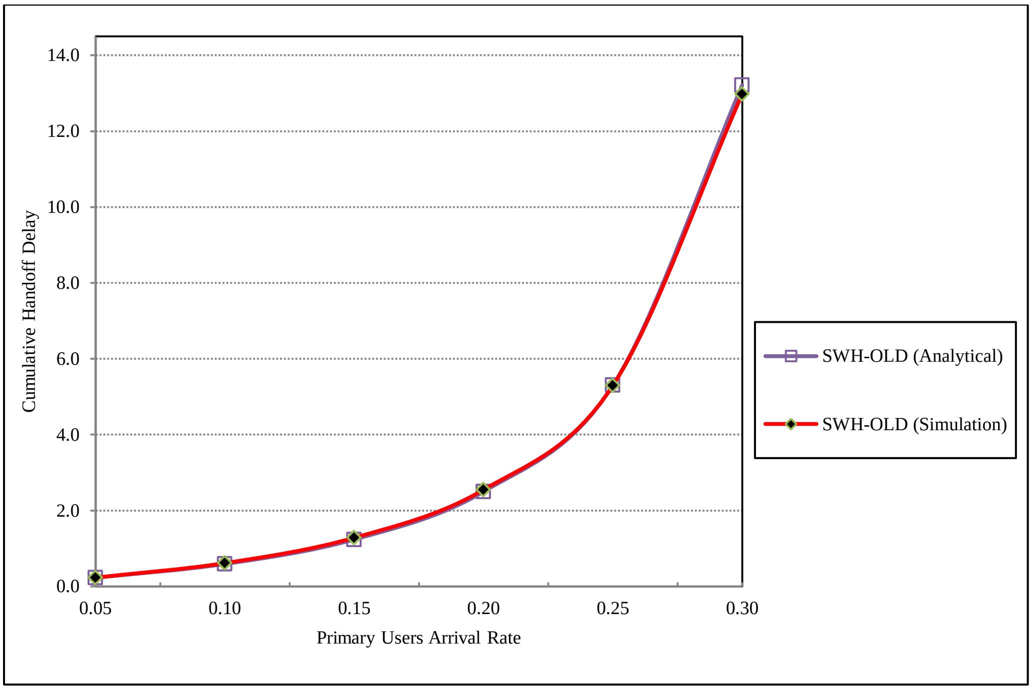

Figure 3.8: Old switching-handoff scheme.

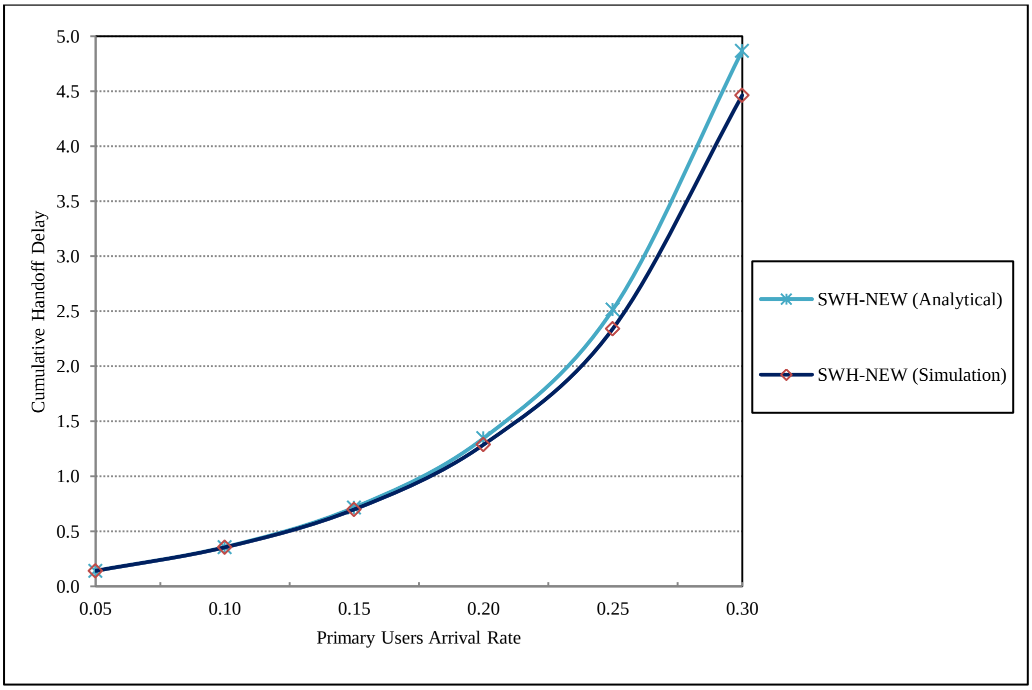

Figure 3.9: New switching-handoff scheme.

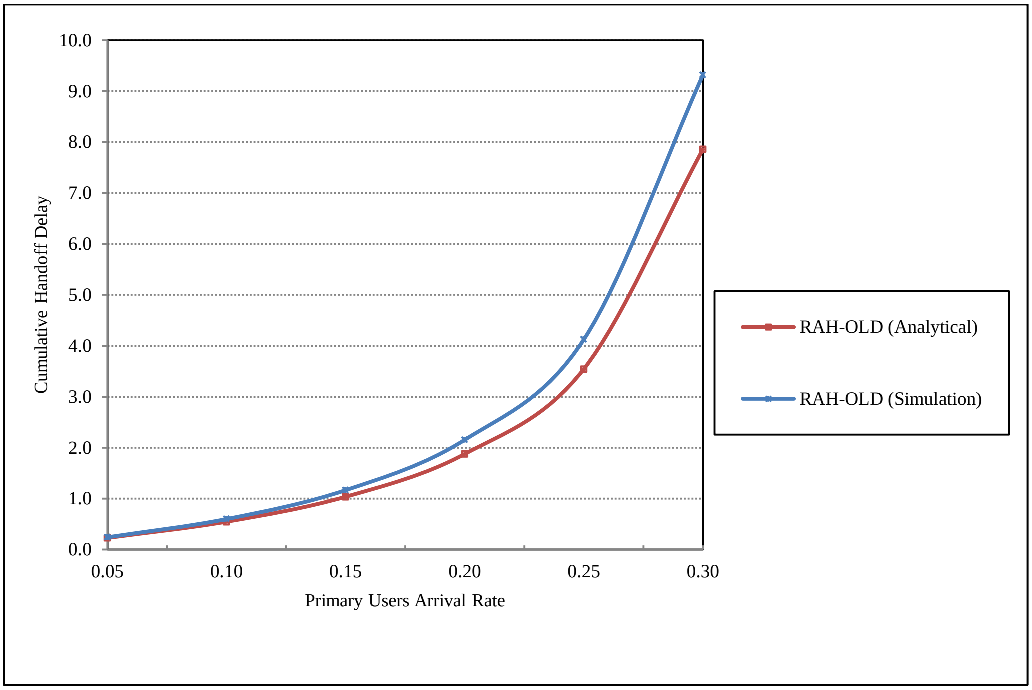

Figure 3.10: Old random-handoff scheme.

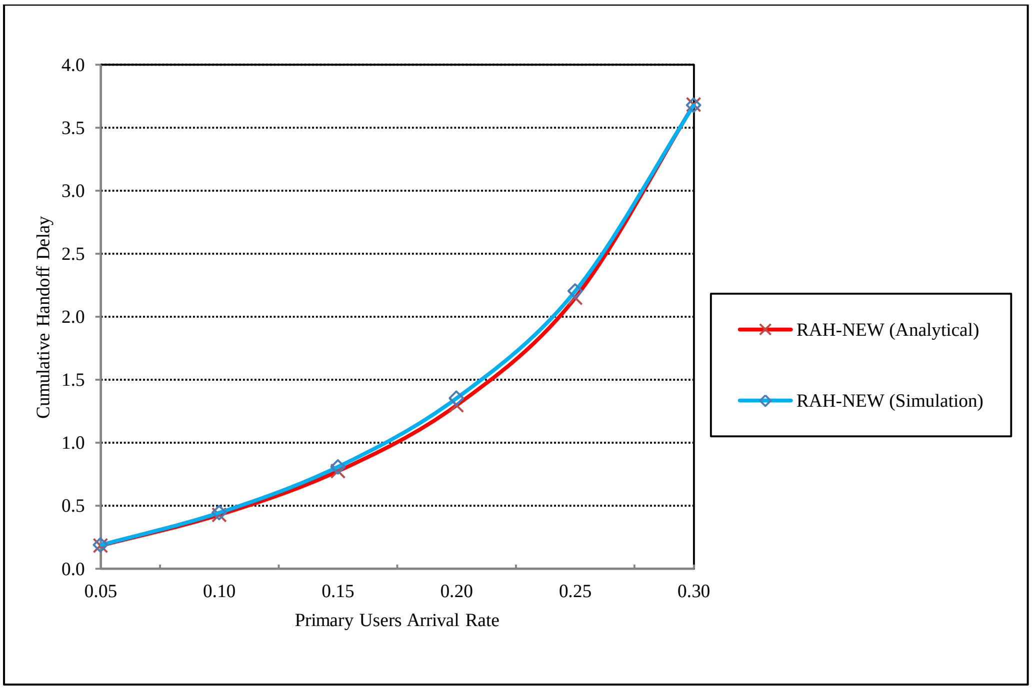

Figure 3.11: New random-handoff scheme.

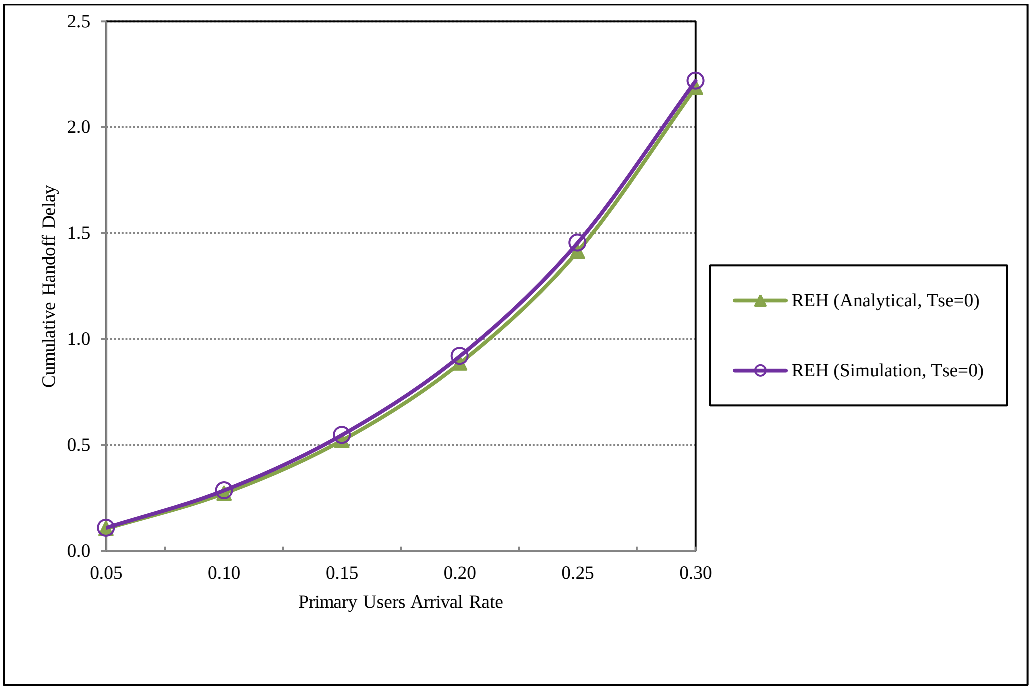

Figure 3.12: Reactive-handoff scheme (T_{se}=0).

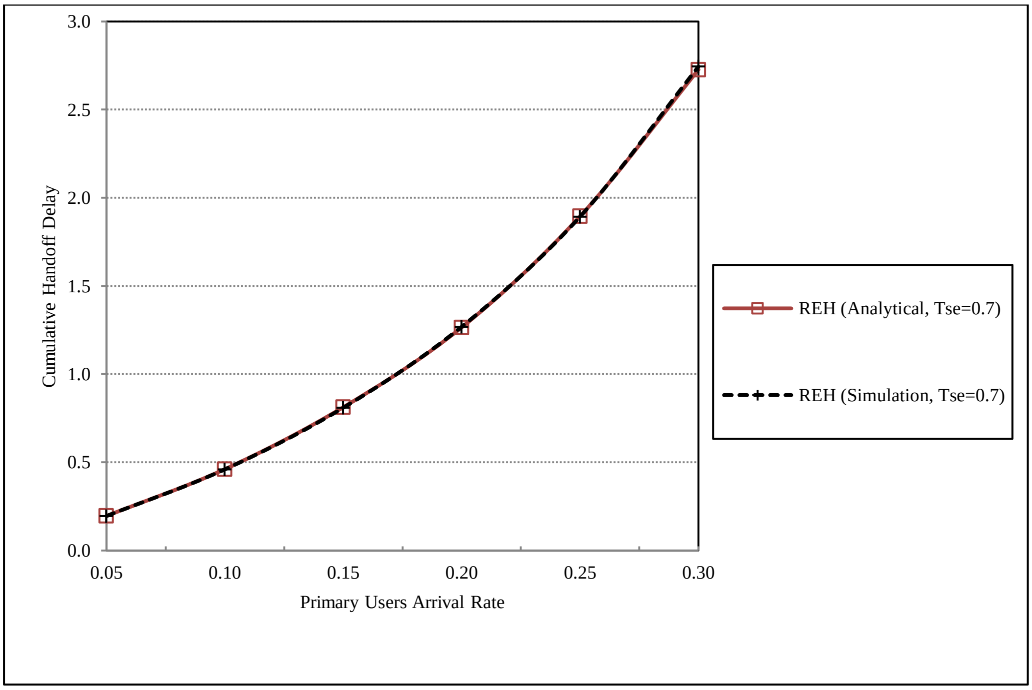

Figure 3.13: Reactive-handoff scheme (T_{se}=0.7).

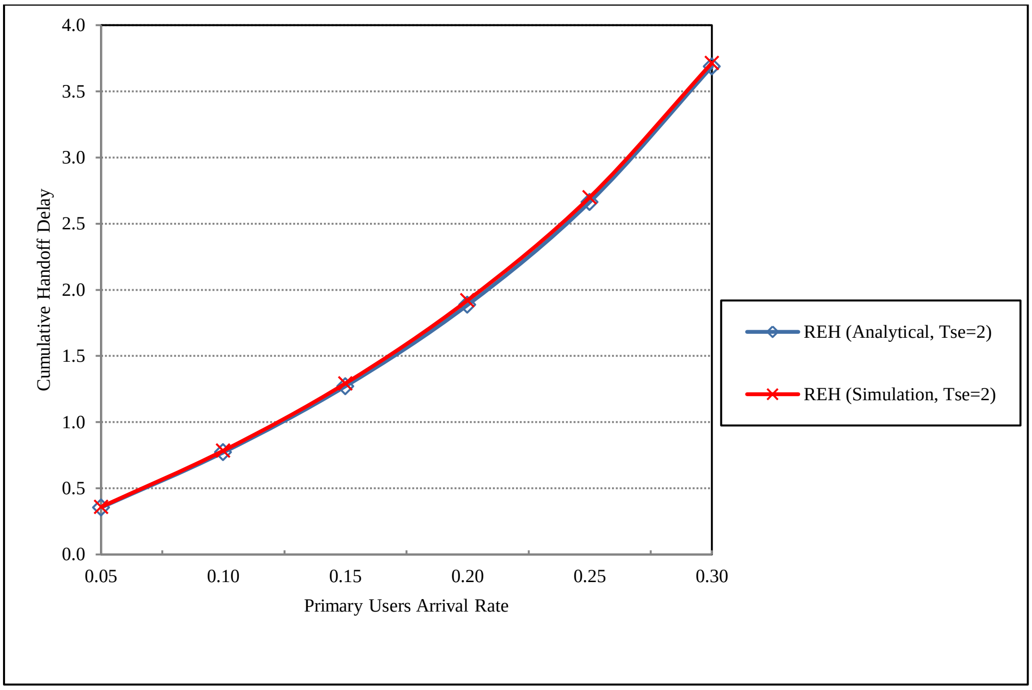

Figure 3.14: Reactive-handoff scheme (T_{se}=2).

Figure 3.15 compares the performance of the reactive-handoff decision schemes (REH) for a range of sensing delay values (T_{se}). The graph shows that, as the sensing time increases, the cumulative handoff delay increases.

Figure 3.16 and Figure 3.17 show that the switching-handoff (SWH-NEW) and random-handoff (RAH-NEW) models implemented with the prioritised criteria outperform their old corresponding schemes (SWH-OLD and RAH-OLD) for every PUs’ arrival rate. For example, when PUs’ arrival rate is 0.3, the SWH-NEW scheme can significantly improve the cumulative handoff delay by 65% (Figure 3.17) and the RAH-NEW scheme considerably by 60% (Figure 3.17). This is the case as the interrupted users in the new models precede any uninterrupted users in the receiving service. This will decrease the cumulative handoff delay for the interrupted secondary users.

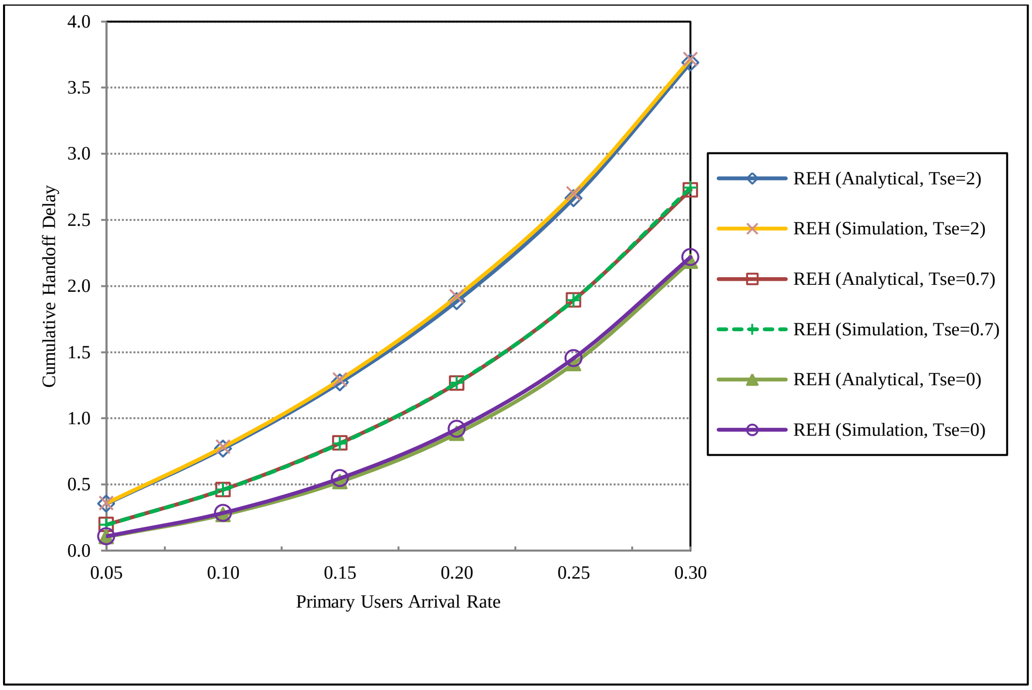

Figure 3.15: Reactive-handoff scheme with different values of T_{se}.

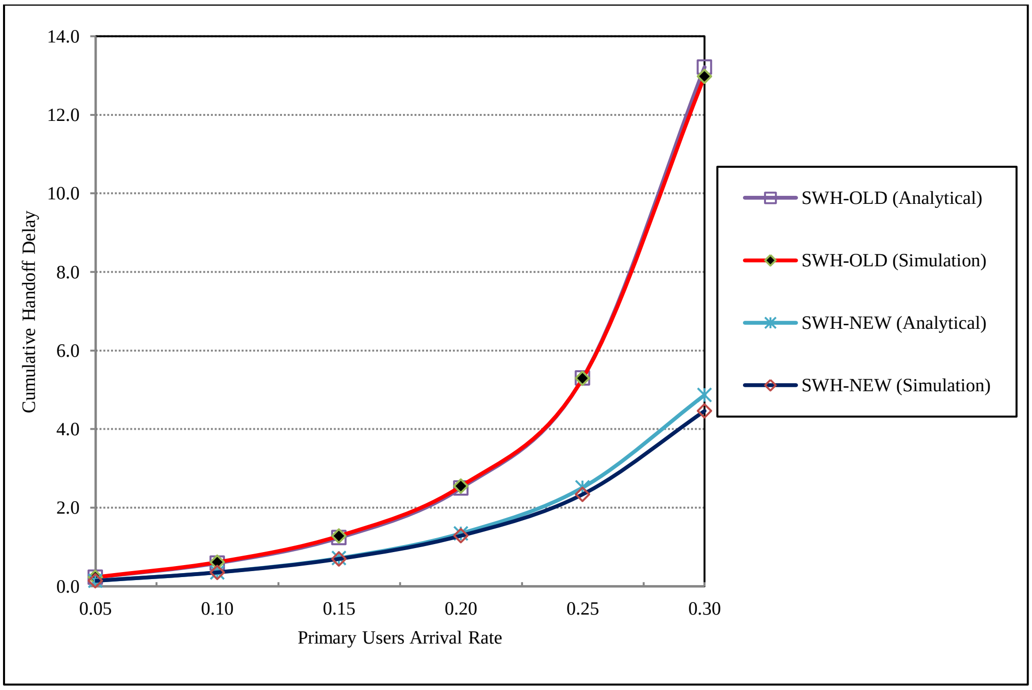

Figure 3.16: Comparison between old and new switching-handoff schemes.

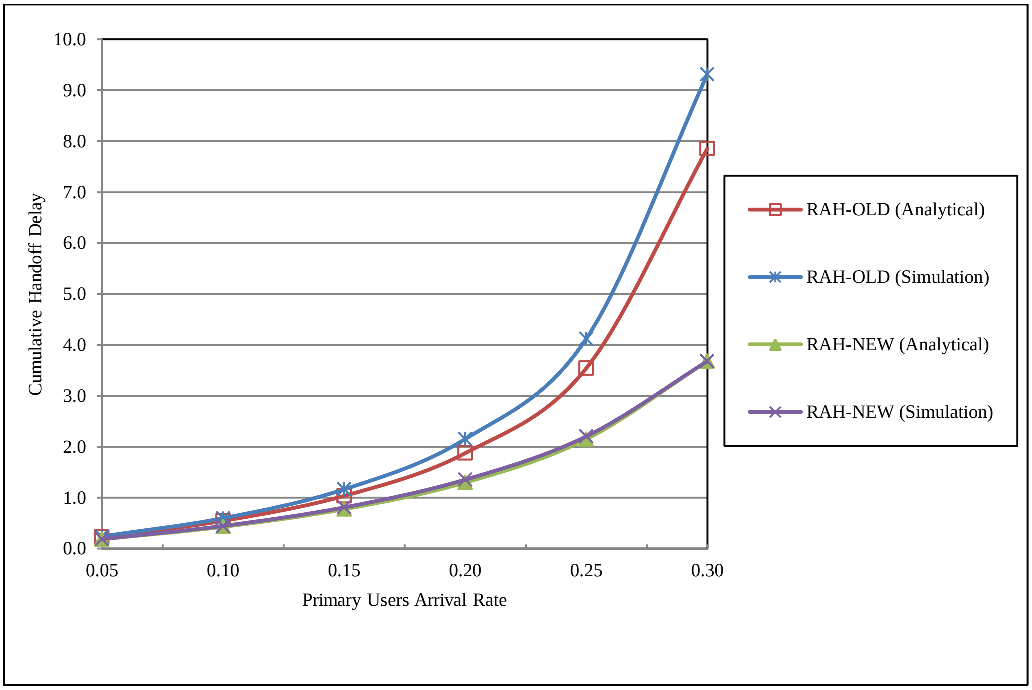

Figure 3.17: Comparison between old and new random-handoff schemes.

Figure 3.18 compares the performance of the reactive-decision schemes (REH) for a range of sensing delay values (T_{se}) with SWH-NEW and RAH-NEW. As it is shown, when T_{se}=0 (not realistic), the reactive scheme achieves the shortest cumulative handoff delay for all the PUs' arrival rates. However, in general, as the sensing delay increases, the reactive model performs poorly compared with other proactive models. For example, when T_{se}=2.0 for the majority of the PUs' arrival rates (0.05-0.27), the REH scheme shows the worst performance in terms of cumulative handoff delay.

Finally, Figure 3.19 compares NSWH, SWH-NEW, RAH-NEW, and REH ($T_{se=}$0.7). The results show that for the PUs arrival rate of 0.05-0.20, SWH-NEW performs the best. However, for the remaining range, NSWH provides the best performance. It is perhaps unsurprising that for lower PUs' rates, the target channel is more likely to be in an idle state, the waiting delay will be decreased which, in turn, reduces the cumulative handoff delay. This could arguably be the reason why SWH-NEW performs better than the other schemes. However, for the higher PUs' arrival rates, the opposite is true and NSWH shows a better performance.

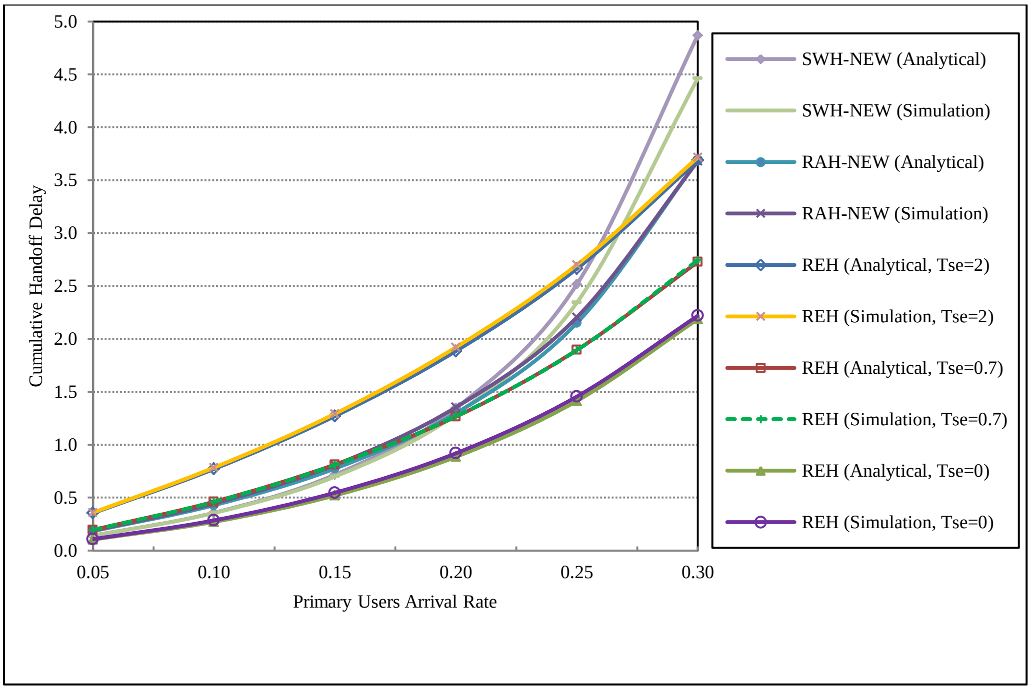

Figure 3.18: Comparison of reactive scheme for different values of sensing delay with new proactive schemes.

Figure 3.20 shows the effects of the secondary users' service rate μ_s on the cumulative handoff delay ($E[D_{cum}]$), when considering the SWH-NEW scheme. From the graph, it is clear that $E[D_{cum}]$ increases as μ_S decreases. $E[D_{cum}]$ This can be interpreted as: when the rate of μ_S decreases, the average service time $E[X_s]$ increases, thus, $\mu_s = 1/E[X_s]$, therefore, the secondary user in service will be interrupted with a high probability, which leads to an increase in the cumulative delay.

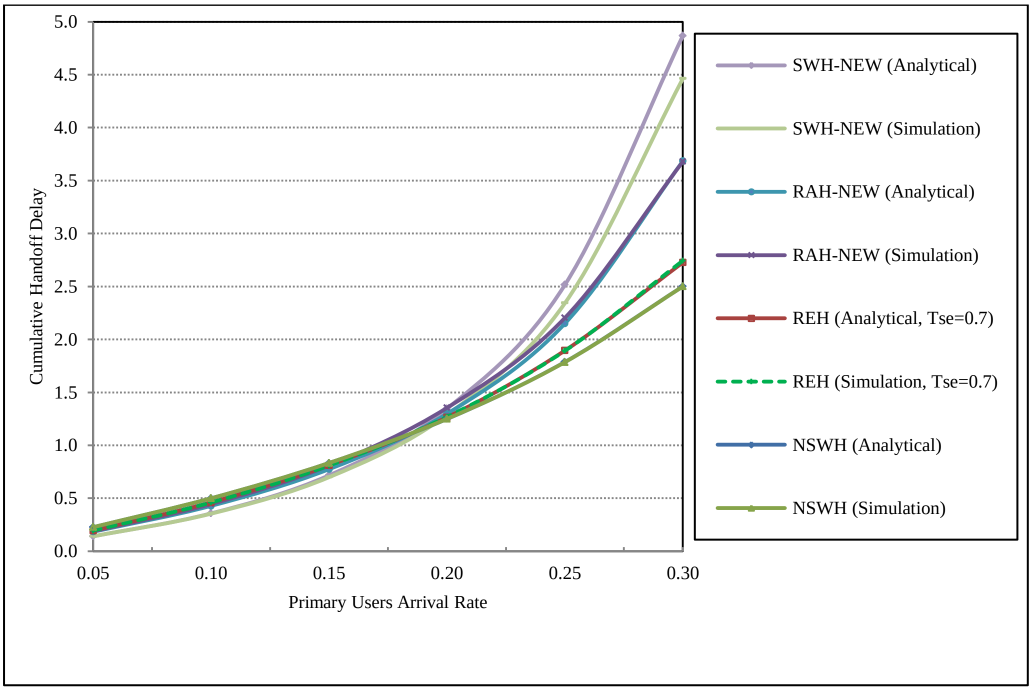

Figure 3.19: Performance of various handoff schemes.

Figure 3.21 shows the effects of the primary users' service rate μ_p on the cumulative handoff delay, in the case of the SWH-NEW scheme. Again, from the graph, it is clear that $E[D_{cum}]$ increases as μ_p decreases. This can be understood as, when the rate μ_p decreases the average service time $E[X_p]$ increases, since, $\mu_p = 1/E[X_p]$, therefore, the interrupted secondary users in their associated queues will wait for long periods before the channel becomes idle, which leads to an increase in the cumulative delay.

The analytical results derived in this chapter can be used to design an admission control policy for the arriving secondary users' rates according to certain average cumulative handoff delay. In general, each application has some delay requirements which should not be exceeded to maintain required level of QoS. Suppose for an application the maximum acceptable level of average handoff delay is 0.002. Assume that $\lambda_s = (0 \div 0.01)$, $\mu_s = 0.01$ and $\mu_p = 0.02$. Figure 3.22 shows an admission region to control the handoff delay. When $\lambda_p = 0.003$, secondary users' arrival rates become higher than 0.005, which restrict the handoff to occur in order to satisfy the delay constraint. In another scenario, when $\lambda_p < 0.001$, the CR network can nearly accept all arriving secondary users. For the given parameters, arriving secondary users are only accepted when ($0 < \lambda_p < 0.006$).An effective policy for controlling handoff delay can be easily designed based on various control techniques, such as call admission control mechanisms [89, 90] and the p-persistent carrier sense multiple access protocol (CSMA) [91].

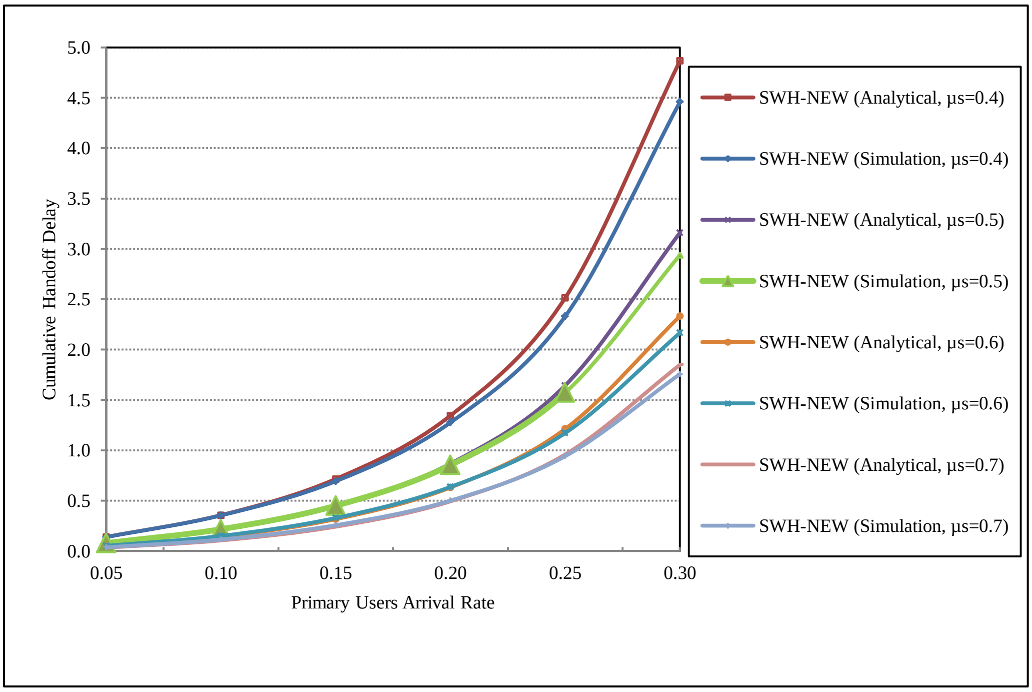

Figure 3.20: Effect of SUs service rate on the cumulative handoff delay.

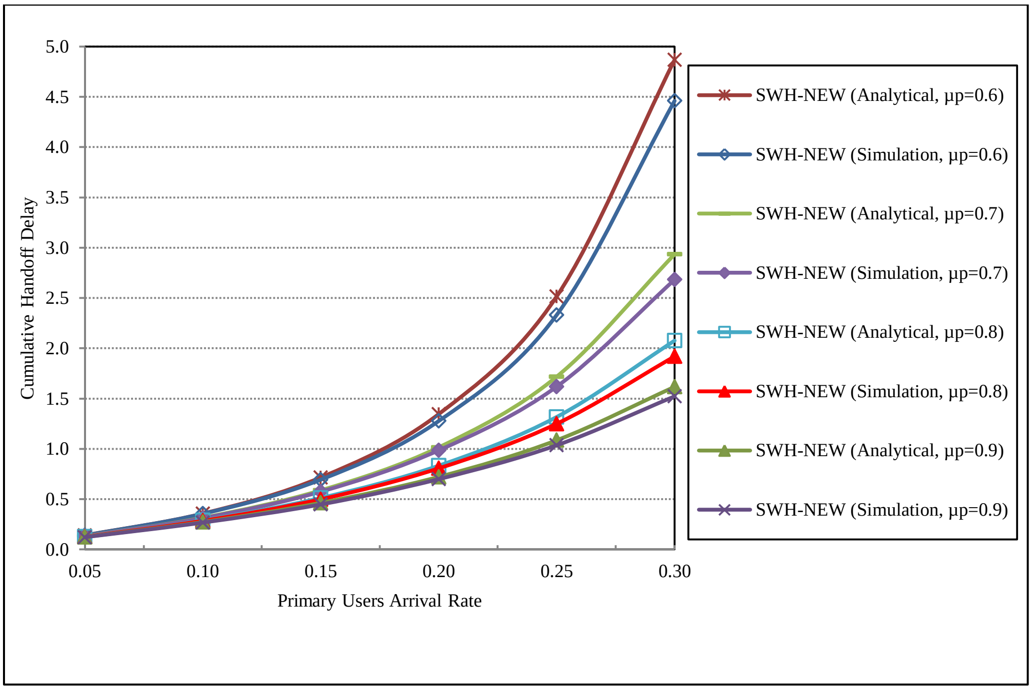

Figure 3.21: Effect of PUs service rate on the cumulative handoff delay.

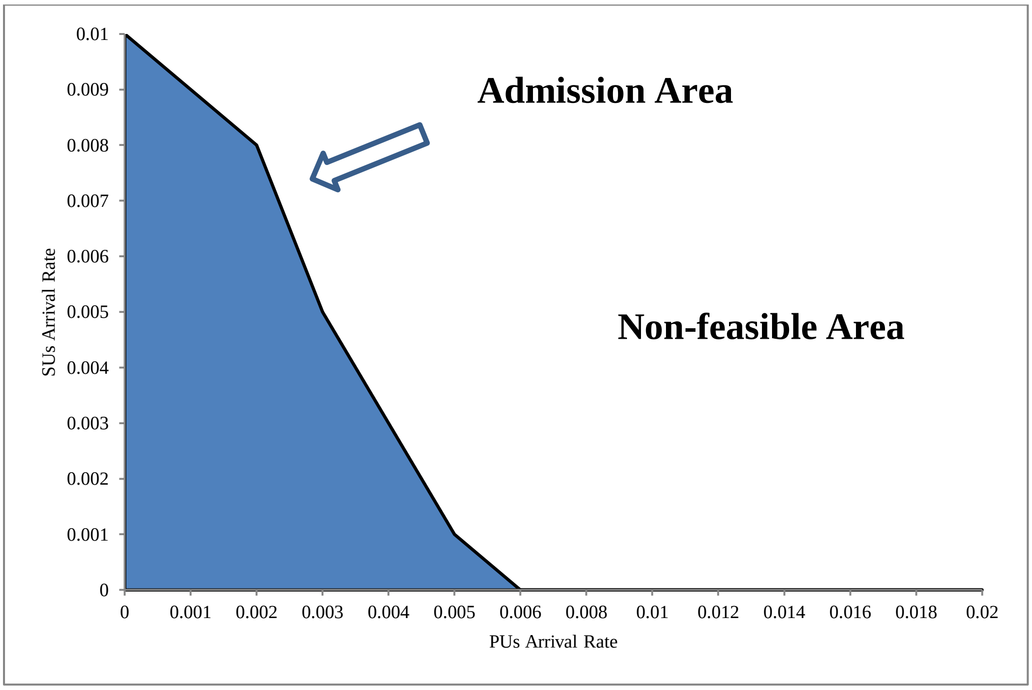

Figure 3.22: Admission region.

3.6 Summary

In this chapter, we presented a prioritised proactive decision handoff scheme in cognitive radio networks. Existing work does not consider any preferences for interrupted secondary users to resume their unfinished transmission on the target channel under the case of a handoff process. The proposed prioritised schemes provide the interrupted secondary users with a higher priority to utilise unused licensed channels. Results confirm that our proposed prioritised schemes reduce the cumulative handoff delay and hence the total service time of the interrupted secondary users.

Chapter 4.

NON SENSING-BASED SPECTRUM HANDOFF MODELLING

4.1 Introduction

In cognitive radio networks (CRNs), spectrum handoff procedure will be initiated and performed whenever the spectrum owner returns to its licensed band. Accordingly, operating secondary users (SUs) working in that band have to leave the band immediately and transfer to another idle band. The quality of the on-going communication can be degraded by this. In fact, simple spectrum handoff strategies can accomplish reasonable performance for various communication desires, whereas more innovative adaptive strategies are essential to achieve the highest benefit. In general, spectrum handoff can be implemented using two different strategies: reactive spectrum handoff and proactive spectrum handoff, as explained in the previous chapters. In the reactive-decision spectrum handoff approach, interrupted secondary users perform channel search for a backup channel in an on-demand manner once detecting the primary users' (PUs) return to the current operating channel. Mostly, instantaneous wideband spectrum sensing will be performed to help the secondary users search for idle channels in order to resume their unfinished transmission. The advantage of such a reactive approach is that sensing results are accurate. Nevertheless, the cost of this accuracy is the longer handoff delay and greater power consumption [92]. Sensing delay refers to the time period that the interrupted secondary user spends until finding an idle channel for transmission after an interruption event happens [35, 57, 59]. What is more, sensing delay widely depends on the number of channels to be sensed. As a consequence, sensing delay plays a key role in this type of spectrum handoff scheme as it increases the handoff delay and the total service time.

4.2 Proposed Spectrum Handoff

In cognitive radio networks, the spectrum agility functionality aims to support the secondary users selecting the best channel(s) to send and receive their data in case of the spectrum owners’ return to their licensed bands. This process is called spectrum handoff. Spectrum handoff is considered as a very big issue in cognitive radio networks.

As mentioned before, generally, spectrum handoff approaches can be categorised into two key approaches from the point of view of the decision timing for choosing the goal channel for future spectrum handoffs: proactive-decision spectrum handoff and reactive-decision spectrum handoff [47, 48]. The goal channel is the channel on which the transmission of the interrupted secondary user will be continued in after discovering the return of the spectrum owner (PU) to the current operating channel. Seamless spectrum handoff is achieved, which is smooth and fast without performance degradation during handoff events, which can be a very difficult issue. In general, the performance decreasing problem is associated with the handoff delay. The more handoff delay, the more degradation on the performance. The former approach has been discussed broadly in Chapter 3, and new schemes have been implemented and compared with other existing schemes. There have been several earlier researches on the latter approach, reactive-decision spectrum handoff.

In [35, 57, 59], the corresponding reactive-decision handoff scheme is basically the sensing-based scheme and is expressed in detail in Chapter 3, too. However, sensing delay plays the most important role in this type of handoff scheme. In general, in such a scheme, handoff users will perform wideband spectrum sensing for some time to search for candidate idle channels; if there is more than one empty channel, then the handoff procedure will choose at random one out of those channels to continue the transmission. In case there is no idle channel, interrupted SUs will have to wait at the SUs' queue of the working channel until the channel becomes idle again [35, 57, 59]. The problem here is that when the number of channels which handoff users have to sense is high, in this case, the sensing time will be high as well because of the fact that the sensing time is directly proportional to the number of channels that have to be sensed. However, sensing a few of the nominee channels makes finding empty channels a hard task, even though it reduces the handoff delay.

However, if it is assumed that the primary system provides the secondary system with some information regarding the current spectrum usage (such as expected duration of usage) by a centralised database [45], then the spectrum sensing time can be eliminated from counting the handoff delay which improves the performance of the secondary users. Moreover, there will be no need for the handshaking process as the target channel for spectrum handoff is known for the sender and the receiver.

This work is an extension of our previous work [58]. However, the presented work in [58] is not supported by any of the results and no comparisons with other existing spectrum handoff schemes have been executed. Moreover, when we compare the new reactive-decision suggested spectrum handoff scheme, with the existing sensing-based reactive-decision scheme, and with the other proposed schemes in Chapter 3, the new scheme shows an improvement in terms of cumulative handoff delay at most situations.

4.3 System Model

This chapter proposes a Preemptive Resume Priority M/M/1 queuing model to implement a cognitive radio network with two wireless channels, where the accessing point is implemented with a database service, as illustrated in Figure 4.1 and Figure 4.2. In this model we examine the conditions in which the present transmission channel or the other channel should be used to resume the interrupted communications dependent on the state of the channels; namely, the secondary users' instantaneous queue length. However, to estimate the state of each channel, a central controller will collect information regarding all of the SUs' instantaneous queue length and stored them in a database service.

Preliminary properties for the proposed PRP M/M/1 queuing network model:

- SUs or PUs arrive at their default channel, i.e., *k*, according to Poisson distribution, with mean rates $\lambda_s^{(k)}$ and $\lambda_p^{(k)}$, respectively.
- The service times are modeled by the exponential distributions, with mean rates of $E\left[X_s^{(k)}\right]$ and, $E\left[X_p^{(k)}\right]$, respectively.

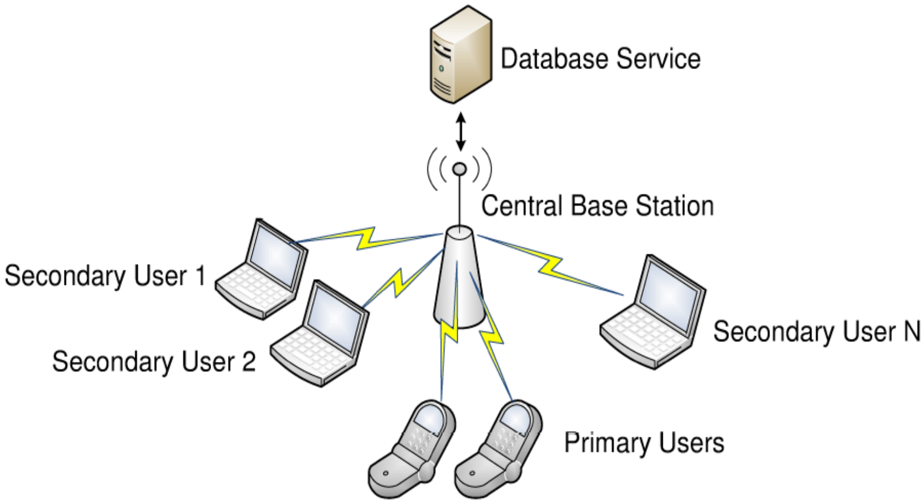

Figure 4.1: Centralised entity with database service [93].

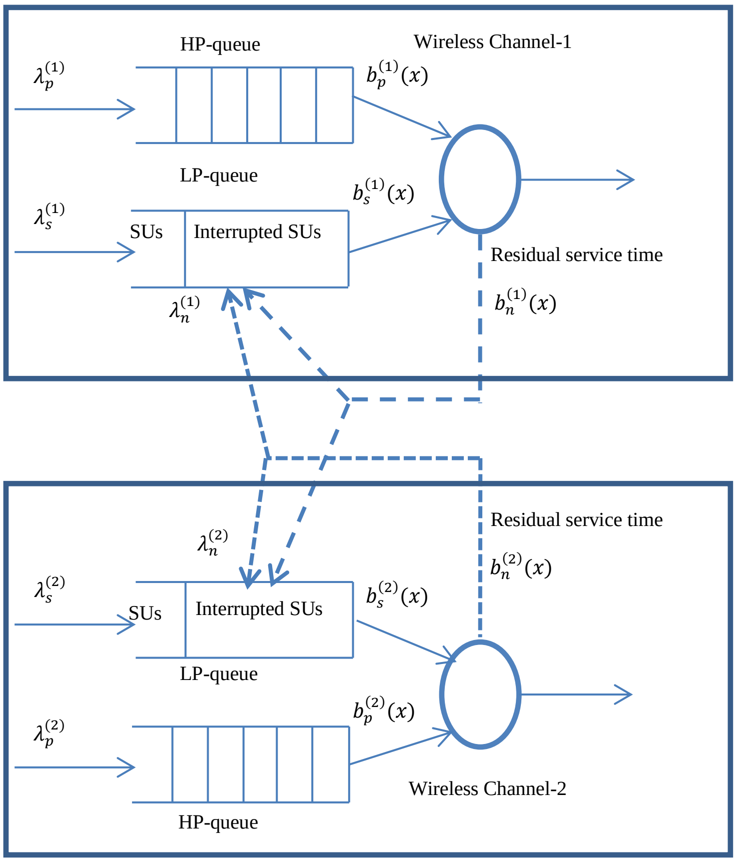

Figure 4.2: Proposed CRN network (where n denotes n^{th} interruption).

- Both classes of users will compete to utilise each channel on the basis of the PRP M/M/1 queuing disciplines.
- Within the same class of users, SUs and PUs will compete to utilise the frequency channel on the basis of the FCFS order in each channel. Therefore, each wireless channel is assumed to be collision-free.
- This model uses preemptive priority queue disciplines which characterises the inherent traffic structure in CRN's.
- PUs can preempt the transmission of secondary users.
- Interrupted users arrive at their target channels with the mean rate $\lambda_i^{(k)}$ and finish their transmission with a mean effective transmission time $E\left[X_i^{(k)}\right]$.
- The algorithm selects the goal channel depending on the shortest instant queue length of the secondary users.

Note that, it is important to incorporate the instantaneous primary users' queue length as well in determining the target channel for spectrum handoff. However, this is not practical in cognitive radio network since it is always assumed that no interaction is possible between the secondary users' network and the primary users' network, and the communications of the secondary users is totally transparent to the primary users [94].

4.4 Spectrum Handoff Decision

It is impractical to execute the spectrum handoff procedure reactively according to the delay metric as decision factor, as the required goal of the application is to satisfy specified delay restrictions [95]. For example, one of the challenges is how to know the instant queuing delay of the users in the queues at the moment of arriving exactly, in order to choose the channel with the shortest queuing delay. However, queuing delay can be estimated based on the number of users in the queue, as expressed in Little's' law as following:

$$W_s = \frac{L_s}{\lambda_s} \qquad 4.1$$

Where,

λ_s: Interrupted secondary users' arrival rate.

W_s: Secondary users' queuing delay.

L_s: Secondary users' Mean queue length.

As it can be seen from Equation 4.1, the queue length (L_s) is directly proportional with the queuing delay (W_s). Therefore, in this model, the instantaneous queue length of the secondary users is used as the key factor in determining the spectrum handoff decision. In other words, handoff secondary users will choose the channel with the shortest queue length in order to resume the interrupted communication. However, if the channels' queues have the same queue length, then the interrupted users will stay at their operating channel to resume the transmission as soon as it becomes available.

4.5 Experiments Setup

A Preemptive Resume Priority M/M/1 queuing model is proposed to implement a cognitive radio network with two identical wireless channels. In order to examine the performance of the proposed handoff model, extensive simulation experiments for various PUs' arrival rates have been conducted using the MATLAB simulation package. Simulation parameters are stated in Table 4.1.

Simulation parameters	**Symbol**	**Value(s)**
PU arrival rate	λ_p	0.05……..0.30
SU arrival rate	λs	0.15
PU service rate	μ_p	0.60
SU service rate	μ_s	0.40
PU arrival rate Increment	Δ	0.05

Table 4.1: Simulation parameters and symbols.

4.6 Performance Calculations

In simulation experiments, to calculate the cumulative handoff delay $E[D_{cum}]$ for each wireless channel, the following formula has been used:

$$E\left[D_{cum}\right]=E\left[N\right]\left(\frac{(\sum Handoff\ Delays)}{Number\ of\ \ Interruptions}+T_{sw}\right) \quad 4.2$$

Where, T_{sw} represents the switching delay which is assumed to be very small and can be neglected. Whereas, $E[N]$ represents the average number of interruptions which can be defined as:

$$E\left[N\right]=\frac{Number\ of\ \ Interruptions}{Number\ of\ \ SUs\ Arrivals} \quad 4.3$$

However, In order to achieve reliable simulation results, the average statistics of the two channels have been taken in order to draw the figures.

4.7 Results

Figure 4.3 to Figure 4.6 compare the performance of the proposed reactive spectrum handoff scheme with other handoff schemes. As expected, the proposed scheme can improve the performance of the suggested cognitive radio network in terms of the cumulative spectrum handoff delay. This is because the handoff secondary users choose the channel with the shortest instant queue length as a backup channel to resume their disturbed transmissions. Furthermore, the proposed scheme provides interrupted users with a higher priority over the new users to utilise the target channel.

The graphs draw the relationship between the cumulative handoff delay and the primary users' traffic loads for various spectrum handoff schemes. Figure 4.3 compares the existing reactive spectrum handoff scheme (which is denoted here as REH-OLD, also denoted as just REH in Chapter 3) under various given sensing times with the new implemented scheme in this chapter (REH-NEW). When we assume the idle case in which the sensing time is zero (*Tse*=0), which is unrealistic, REH-OLD outperforms REH-NEW for all primary user traffic loads. However, as the sensing time increases, the resultant cumulative handoff delay of REH-OLD increases as well. When the sensing time is set to 0.7, the two handoff schemes achieve approximately the same cumulative handoff delay for loads between (0.05-0.25). On the other hand, REH-OLD has the lowest cumulative handoff delay for the other rest loads. For relatively high values of sensing time; i.e., 2 and 3, REH-OLD performs poorly compared to REH-NEW for all traffic loads and experiences average cumulative handoff delay up to 21% and 34% higher than what the REH-NEW achieves. From the graph, it is clear that as the sensing time increases, the cumulative handoff delay of REH-OLD increases as well. This is a natural result as REH-OLD is a sensing-based algorithm and is affected directly by increasing the sensing time, whereas REH-NEW does not rely on sensing techniques.

The proposed spectrum handoff scheme REH-NEW spends ~~a~~ period of time to access the database centre. According to [96], the database access time is 4 msec, on the other hand, it takes 160 msec [97] to sense only one wireless channel. Furthermore, it is mentioned in Subsection 3.3.2.4, the sensing delay is a cumulative parameter which means if the number of channels to be sensed is 10, then the required sensing time to sense them is 1600 msec which is ten times the sensing time needed to sense one channel, and so on. As a result, the database access time can be neglected comparing with the channel sensing time.

Figure 4.4 provides a comparison between REH-NEW and some other new spectrum handoff schemes that are implemented in Chapter 3, such as switching handoff (SWH-NEW) and random-handoff (RAH-NEW) schemes. From the graph, it is clear that SWH-NEW has the lowest cumulative handoff delay for the traffic loads between (0.05-0.18). Above 0.18, REH-NEW experiences low cumulative handoff delay compared with the other schemes, and thus can improve the average

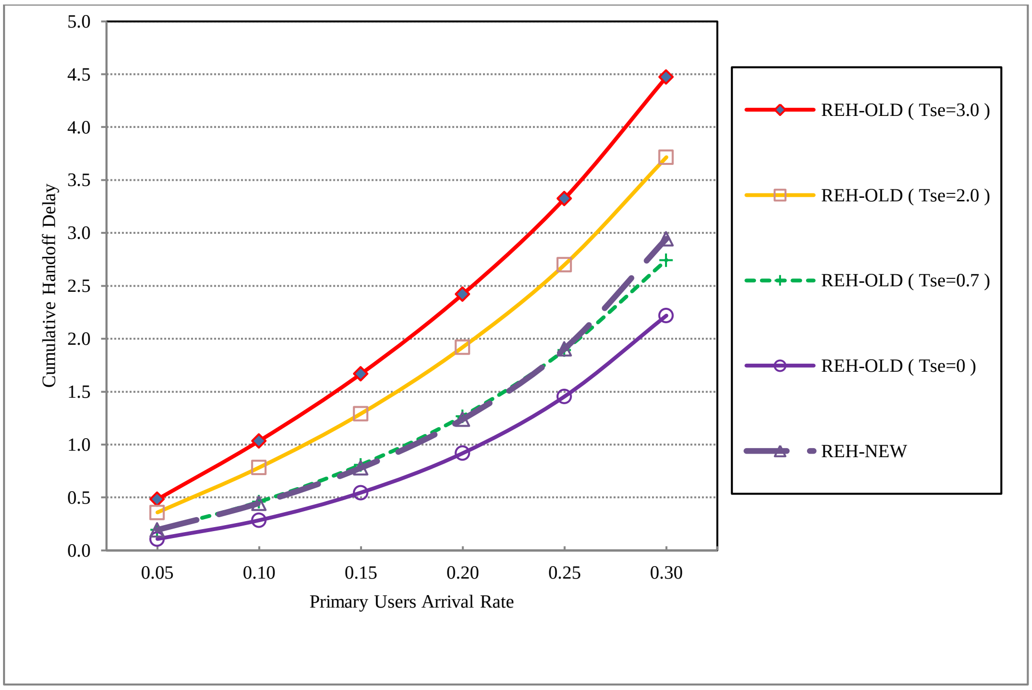

Figure 4.3: Comparison of new and old reactive handoff schemes under different sensing times.

cumulative handoff delay up to about 20% and 34% compared with RAH-NEW and SWH-NEW, respectively. This can be explained as:

In the case of SWH-NEW, handoff users change their operating channel whenever an interruption event occurs - as explained in Chapter 3. However, at the low traffic loads of primary users (0.05-0.18), it is more likely to find the target channel in the idle state. As a consequence, the cumulative handoff delay will reduce in comparison with the other schemes. On the other hand, at the higher traffic loads, REH-NEW works better as SWH-NEW experiences high cumulative handoff delay because it is more likely for an interrupted secondary user to find the target channel in a busy state. As a result, interrupted secondary users have to wait at the corresponding queues which increase their cumulative handoff delay. Whereas, in the case of RAH-NEW, the scheme selects the target channel for spectrum handoff randomly and does not consider any real effective decision factors, hence, REH-NEW performs better because of the fact that it considers the innovative rule to choose the target channel.

Generally, at low traffic loads, there is not much of a difference between the performances of the three schemes. This is because the channels are not utilised most of the time and a high percentage of handoff users will not wait for very long in order to get service.

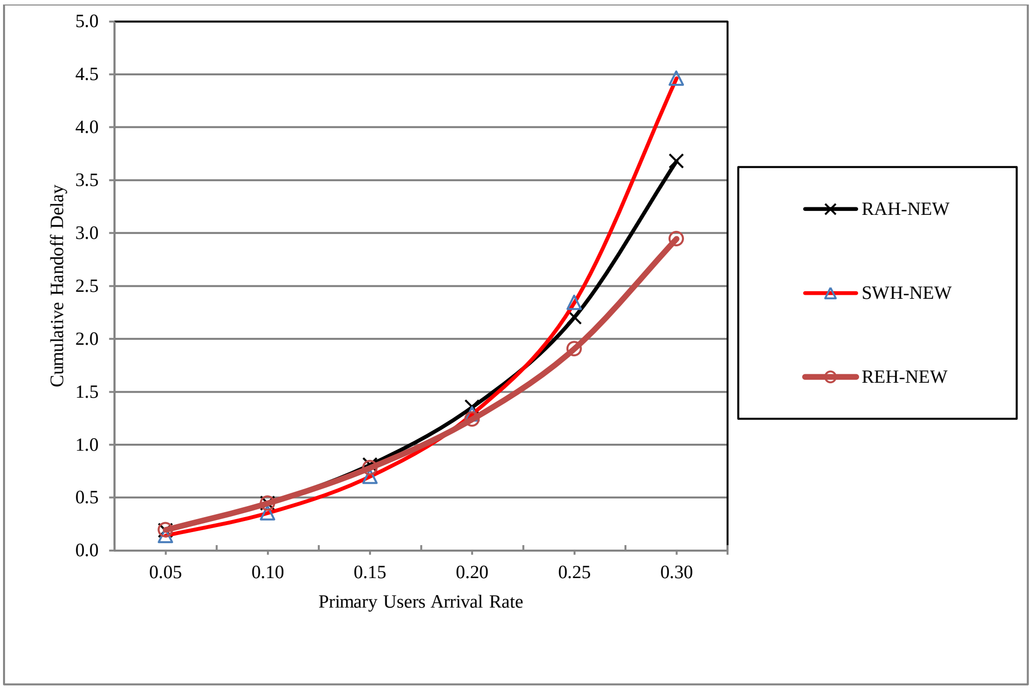

Figure 4.4: Comparison of various new handoff schemes.

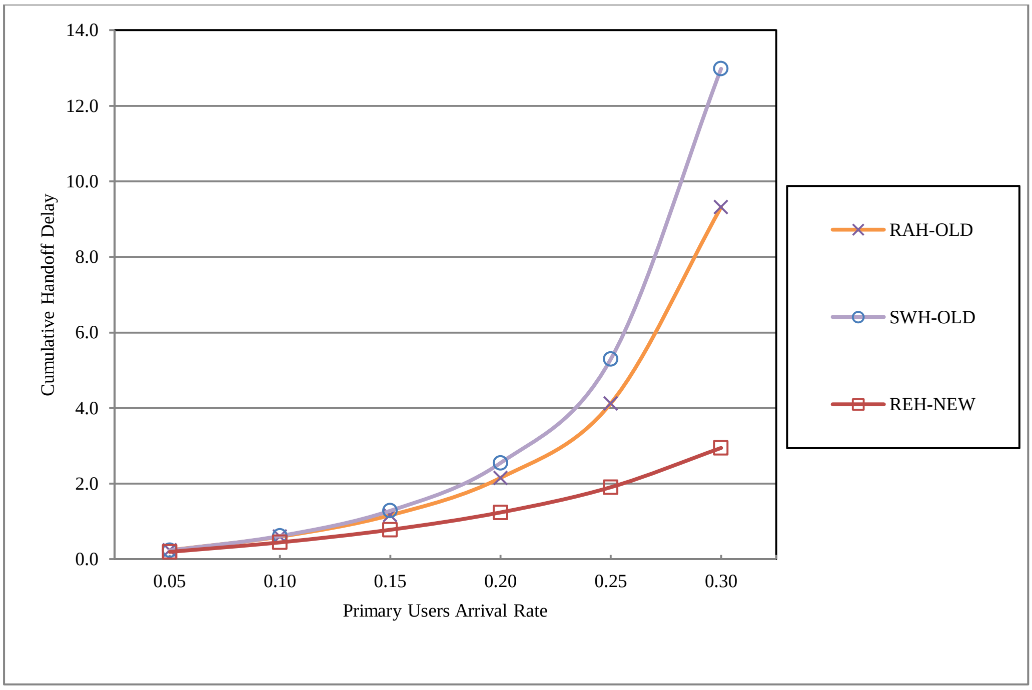

Figure 4.5: Comparison of new reactive-handoff scheme and other old schemes.

Figure 4.5 displays a comparison between the existing old handoff schemes (RAH-OLD and SWH-OLD) with the new proposed reactive scheme REW-NEW. The new proposed scheme proves to have the lowest cumulative handoff delay amongst the presented schemes in the graph. REW-NEW achieves cumulative handoff delay less than approximately 68% and 77% compared with RAH-OLD and SWH-OLD, respectively.

Figure 4.6 compares the performance of the new reactive handoff scheme (REH-NEW) with the existing non-switching-handoff scheme (NSWH). The graph shows that at low and moderate traffic loads approximately between (0.05-0.22), REH-NEW performs better than NSWH, while above this range, NSWH shows an improvement in the cumulative handoff delay. In NSWH scheme, interrupted secondary users always stay at their operating channel in order to finish their transmission once the channel becomes idle (see Subsection 3.3.2.1). At the low and moderate traffic loads, the other channel is more likely to be in the idle state, and since NSWH is not allowed to change the operating channel, the interrupted users always have to wait whenever an interruption event occurs, which will increase their cumulative handoff delay. On the other hand, REH-NEW allows the interrupted users to change their operating channel and select the channel with the shortest queue length which will decrease their cumulative handoff delay. However, at high traffic loads, interrupted users in the NSWH scheme achieved the lowest cumulative handoff delay as it is most likely for the other channel to be in the busy state. Here it is important to note that in REW-NEW, the channel with the shortest secondary users' queue length does not always achieve the lowest cumulative handoff delay, as the primary users' queues are not taken into consideration in counting the shortest queue length. For example, it might be the selected channel (of course, with the shortest secondary users' queue) for spectrum handoff that has a number of primary users waiting for service bigger than the other channels have. Consequently, handoff users will encounter a long waiting time even though they select the channel with the shortest secondary users' queue length.

In general, as expected, the proposed scheme can improve the performance of the suggested cognitive radio network in terms of the cumulative spectrum handoff delay. This is because handoff secondary users choose the channel with the shortest instant queue length as a backup channel to resume their disturbed transmissions.

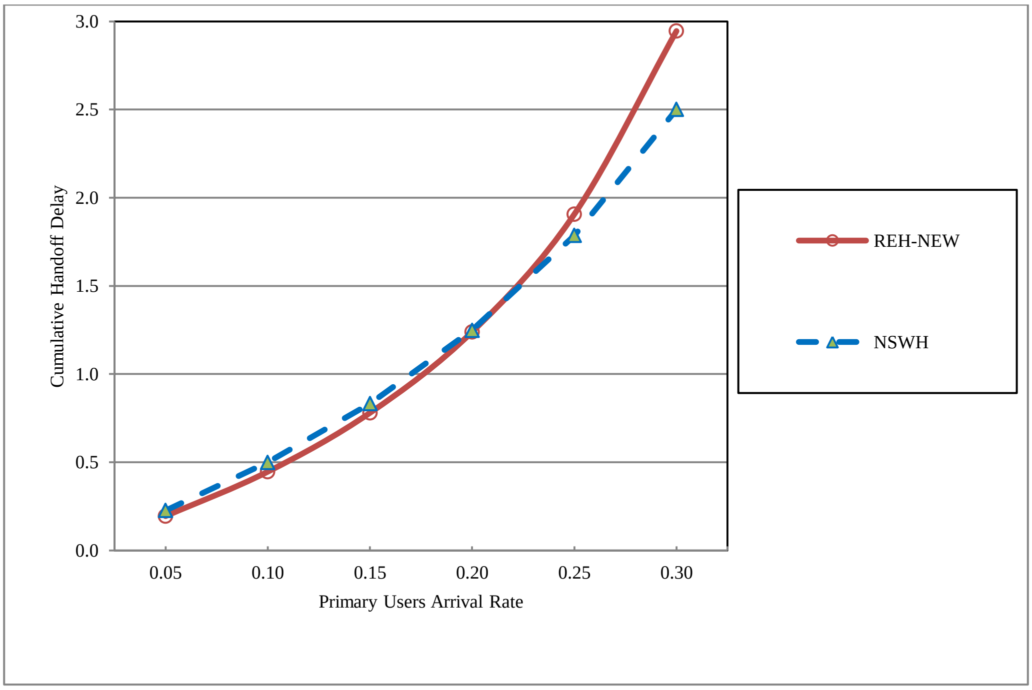

Figure 4.6: Comparison of new reactive-handoff scheme and non-switching handoff scheme.

4.8 Summary

In this chapter, a Preemptive Resume Priority M/M/1 queuing model is proposed to investigate the effect of the spectrum handoff strategies in the cognitive radio networks. Different spectrum handoff strategies are studied and compared quantitatively and qualitatively. The suggested scheme is designed to switch reactively between the current channel and the targeted channel at the events of interruptions depending on the shortest instant queue length. The comparison study shows that our novel scheme can improve the performance of the implemented cognitive radio network in terms of cumulative handoff delay.

Chapter 5.

SPECTRUM HANDOFF MODEL UNDER SHARED HANDOFF-QUEUE

5.1 Introduction

By giving priority to handoff users, the QoS of the secondary users in the network will certainly improve. Sometimes, traffic conditions are different from channel to channel according to the random arrival process to each channel. For example, it is possible to find at a given time a channel with long low priority and/or handoff queues compared with the other channels' associated queues. At the end, some uninterrupted users in a channel might have received service before some other interrupted users in another channel. From the queuing theory point of view, systems with a single common queue are generally believed to be fairer and more efficient than systems with separate queues in terms of delay and performance measures [12]. If we borrow this idea and incorporate it with the implemented model in the previous chapters, then it could compensate the random arrival process effect and therefore reduce the transmission latency of the handoff users and improve the spectrum utilisation of the network as well.

In previous chapters, the suggested cognitive radio network was implemented with separate handoff queues for each wireless channel. In this chapter, a slight modification is applied to the proposed models in the previous chapters. Namely, the number of handoff queues is convergence to only one common queue instead of separate queues, as illustrated in Figure 5.1. The model consists of two priority queues for each wireless channel: low-priority queues (mainly for secondary users' transmissions), and high-priority queues (for primary users' transmissions). Furthermore, a handoff queue (shared-queue) is a common queue and can serve the interrupted users coming from both the system channels.

Our study shows that, by applying the general queue technique with FCFS queue disciplines to all the spectrum handoff schemes implemented in the previous chapters, they are converged to a single spectrum handoff scheme, as illustrated in Table 5.1

The proposed spectrum handoff scheme is modelled using a Preemptive Resume Priority M/M/1 queuing network model. We compare the proposed model with the existing and new implemented models in the previous chapters. Achieved results approved that the shared-queue spectrum handoff model outperforms other models in terms of the total service time under the variation of the primary users' traffic loads.

5.2 System Model

Some general basic properties of the shared-queuing spectrum handoff model are given in the following points:

- Secondary users and primary users will arrive randomly with mean rates of $\lambda_s^{(k)}$ and $\lambda_p^{(k)}$, respectively, according to the Poisson process.
- On the other hand, they will departure after receiving exponential service times with mean rates of $E\left[X_s^{(k)}\right]$ and, $E\left[X_p^{(k)}\right]$, respectively.

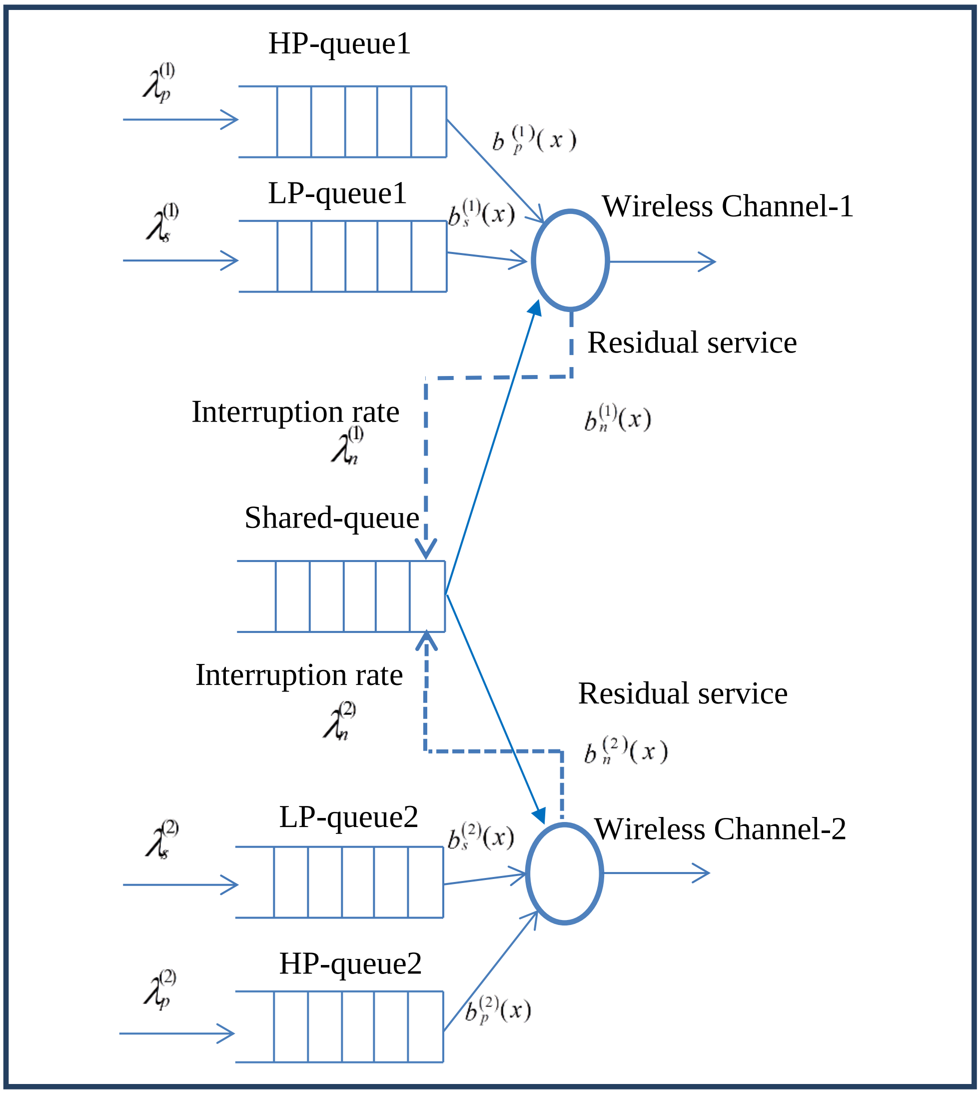

Figure 5.1: Shared-queue (SQ) spectrum handoff model (where *n* denotes n^{th} interruption).

- They will compete for exploiting available spectrum channels opportunistically based on the PRP M/M/1 queuing disciplines in which primary users will preempt the transmission of secondary users at the instance of their arrival if they find their channels being used by the secondary users.

- Within the same channel, SUs and PUs will contend to exploit the frequency channel based on preemptive queuing scheduling disciplines. Consequently, each wireless channel is assumed to be collision free.
- Interrupted secondary users will arrive at the shared-queue (SQ) with a mean rate $\lambda_i^{(k)}$ and will finish their transmission with a mean effective transmission time $E\left[X_i^{(k)}\right]$, where $i \geq 1$.
- If any of the available channels is in idle state, namely the other wireless channel as it is assumed here that there are only two channels, then the interrupted user will resume its unfinished transmission in the idle channel, otherwise it will join the tail of the shared-queue and wait for service.
- Interrupted secondary users at the head of SQ will be handled by the first available wireless channel according to the FCFS queue discipline - of course, after all the primary users in the high-priority queue of the selected channel - and before any secondary users in the low-priority queues of both channels. This means, the first secondary user at the SQ will be served by the first channel that completes servicing all of its primary users. This can done using the dispatching scheduler of the base station. We assume here the service time is much larger than the response time of the dispatching scheduler. So we ignore the effect of the response time.
- The new model is categorised under the reactive spectrum handoff decision approach by its nature. In other words, at the moment of interruption, there is no ready target channel for spectrum handoff to complete the on-going transmission, as mentioned above.

5.3 Experiments Setup

The suggested cognitive radio network consists of two identical wireless channels and is modelled by a Preemptive Resume Priority M/M/1 queuing model. All simulation results presented in this part have been performed with the discrete event simulator (MATLAB) tool to analyse the total service time of the interrupted secondary users.

Table 5.1summarises the various implemented spectrum handoff schemes in this and pervious chapters with related features, while,

Table 5.2 summarises the simulation parameters that are used to achieve the presented results.

5.4 Performance Calculations

In general, the system average total service time *E[T]* for simulation experiments can be estimated using the following formula:

$$E[T] = E[X_s] + E[N] * \frac{(\sum Handoff\ Delays)}{Number\ of\ Interruptions} + T_{sw} + T_{ha} \qquad 5.1$$

Where, T_{ha} is the handshaking time and T_{sw} represents the switching delay which are both assumed to be neglected. Whereas, $E[X_s]$ and *E[N]* refer to the mean service time and the average number of interruptions of the secondary users, respectively. Hence, *E[N]* can be written as:

$$E[N] = \frac{Number\ of\ Interruptions}{Number\ of\ SUs\ Arrivals} \qquad 5.2$$

Ordinary model	**Equivalent shared-queue model**
NSWH (Proactive)	REH-SQ (Reactive)
SWH-OLD (Proactive)	
SWH-NEW (Proactive)	
RAH-OLD (Proactive)	
RAH-NEW (Proactive)	
REH-OLD (Reactive)	
REH-NEW (Reactive)	

Table 5.1 Implemented handoff models.

Simulation parameters	**Symbol**	**Value(s)**
PU arrival rate	λ_p	0.05-0.30
SU arrival rate	λ_s	0.15
PU service rate	μ_p	0.60
SU service rate	μ_s	0.40
PU arrival rate Increment	Δ	0.05

Table 5.2 Simulation parameters.

5.5 Results

As it can be seen from Figure 5.2 - Figure 5.6, a comparison study has been conducted in order to investigate the performance of the new spectrum handoff scheme which was implemented using shared-queue (SQ) for handoff users with various spectrum handoff schemes that are presented in the previous chapters with separate handoff users' queues. The comparison has been made in terms of total service time against primary users' arrival rate. In general, the accomplished results show that the SQ scheme outperforms all the other spectrum handoff schemes for all the primary users' traffic loads.

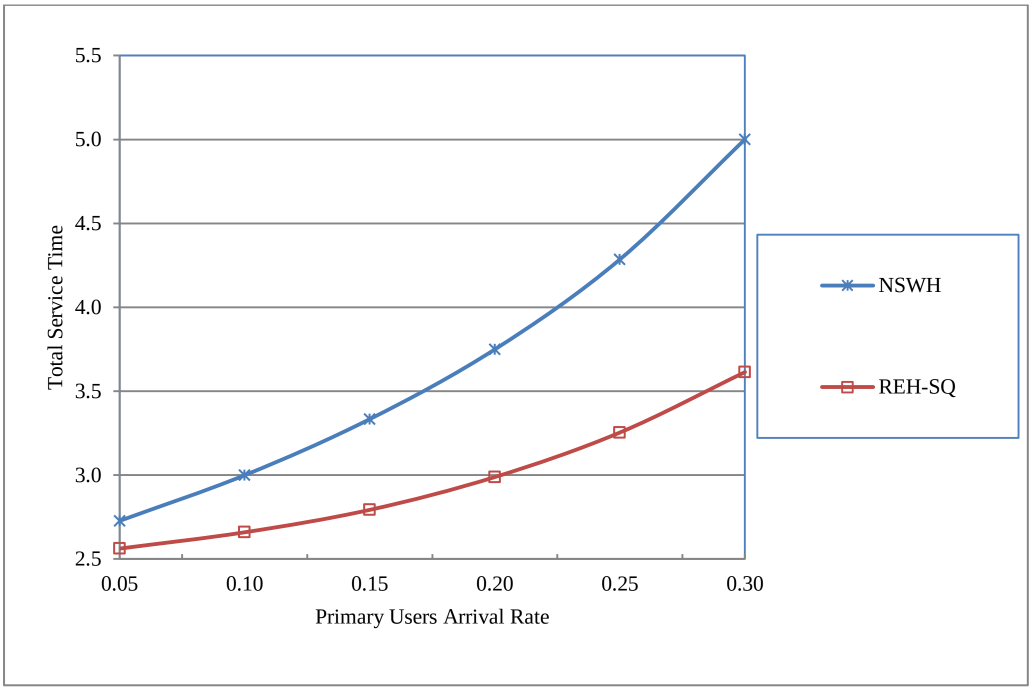

Figure 5.2: Comparison between non-switching and SQ handoff schemes.

Figure 5.2 compares the implemented SQ scheme with the non-switching handoff (NSWH) scheme. From the graph, it is clear that SQ scheme can considerably improve the total service time. This improvement increases as the primary users' traffic loads increase. For example, when the primary users' traffic loads is 0.2, the SQ scheme improves the total service time by 20%, and when the traffic load is 0.3, the improvement reaches approximately 30%.

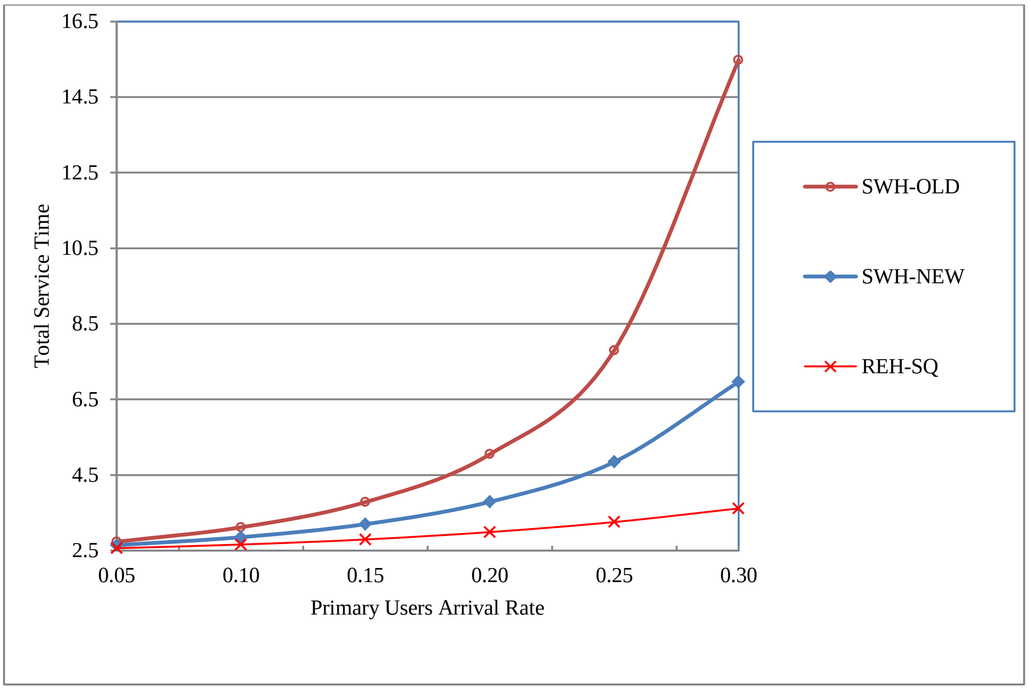

Figure 5.3: Comparison between switching and SQ handoff schemes.

Figure 5.3 and Figure 5.4 illustrate the performance of switching-handoff and random-handoff schemes presented in Chapter 3, and REH-SQ for every PU's arrival rate. Clearly, REH-SQ reduces considerably the total service time of the secondary users compared with the other handoff schemes. For example, when the PU arrival rate is 0.3, the REH-SQ scheme can significantly improve the total service time approximately by 50% compared with the SWH-NEW scheme (see Figure 5.3), and by more than 40% compared with the RAH-NEW scheme (see Figure 5.4). However, this improvement in the total service time can be increased much more when we compare REH-SQ with the old corresponding schemes as shown in the graphs. For instance, the improvement in the total service time is significant and is approximately 75% and 70% compared with the SWH-OLD and RAH-OLD schemes, see Figure 5.3 and Figure 5.4, respectively.

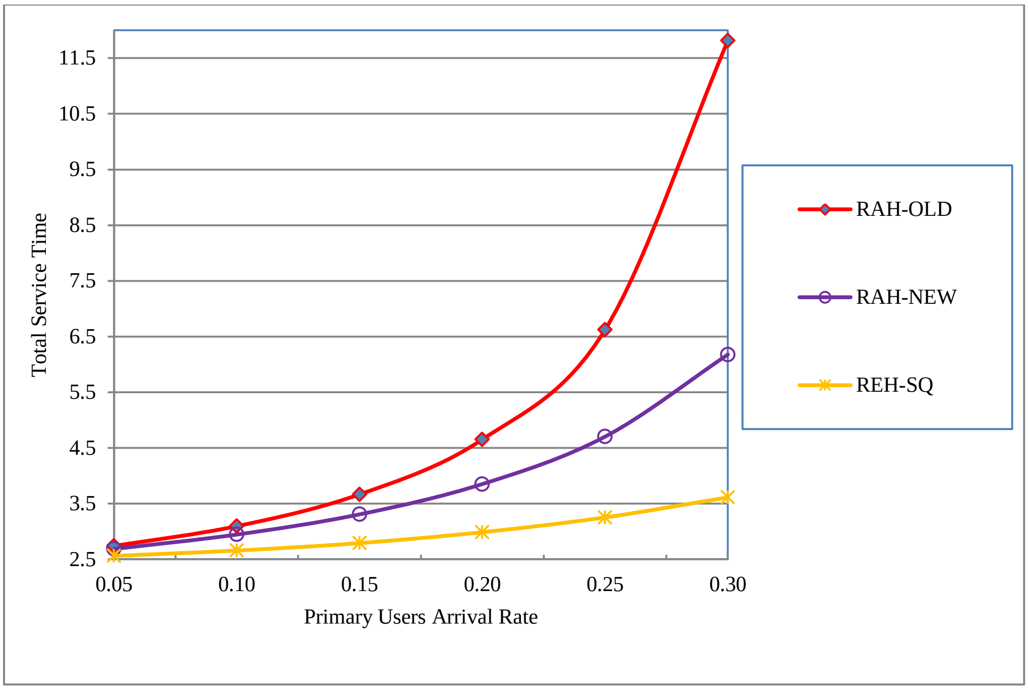

Figure 5.4: Comparison between random and SQ handoff schemes.

Figure 5.5 compares the performance of REH-SQ with the existing and developed reactive-handoff spectrum decision schemes REH-OLD and REH-NEW, respectively. REH-OLD is compared with a range of sensing delay values (T_{se}). The graph shows that even when the idle case is considered, i.e., T_{se}=0, REH-SQ still achieves the lowest total service time compared with the other schemes. For example, when REH-SQ is compared with the idle case of REH-OLD (T_{se}=0), then the improvement in the total service time is nearly 23%. However, as the sensing time increases, this improvement increases as well; when the sensing time is increased to 0.7, and 2, then the improvement increases to about 31% and 42%, respectively. Sensing-based schemes such as REH-OLD waste some time on searching for free channels. As a consequense, the handoff delay and total service time increases, which leads to degradtion in the quality of service of on going transmissions. On the other hand, REH-NEW performs less than REH-SQ by nearly 34%.

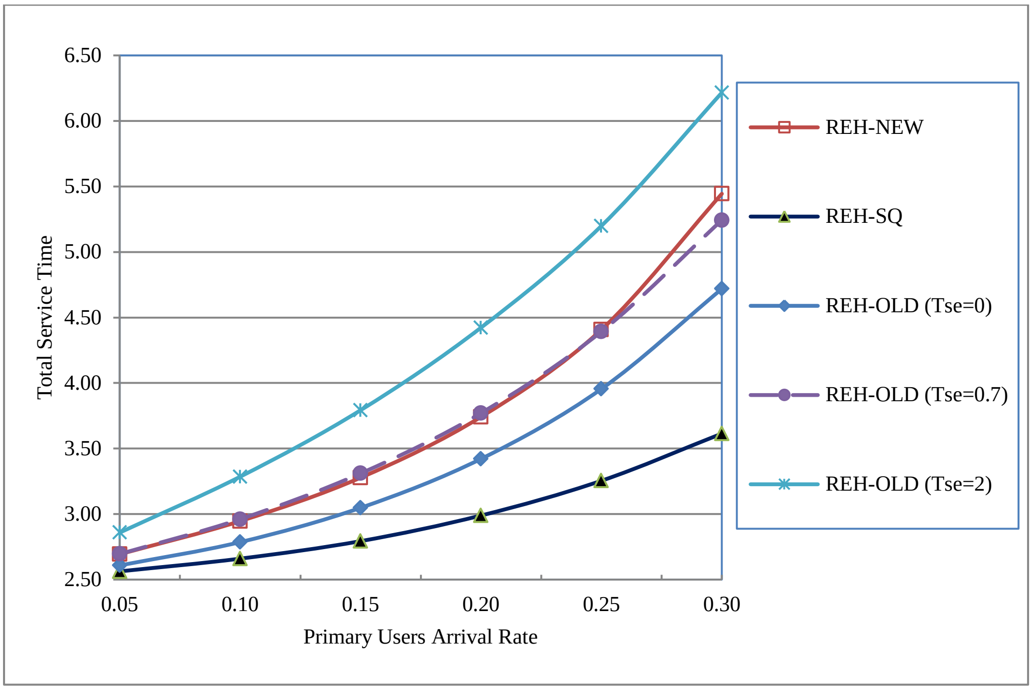

Figure 5.5: Comparison between reactive handoff schemes.

Figure 5.6 compares the NSWH, SWH-NEW, RAH-NEW, and REH-NEW handoff schemes with REH-SQ. The graph shows that REH-SQ outperforms all the other schemes for all PUs' traffic loads. In particular, when we compare REH-SQ with NSWH, RAH-NEW and SWH-NEW, REH-SQ shows an improvmet of about 28%, 42% and 48%, respectively.

This is true, since in the separate queue case, it is possible to find some interrupted users waiting for service in the queue while the other channels are in the idle state, or it has only secondary users in the low-priority queue waiting for service. Of course this will waste some transmission opportunities for the interrupted users which will lead to an increase in the handoff delay and the total service time as well. On the other hand, when we consider the shared-queue case, interrupted users in the shared queue will be handled by the first channel, which becomes idle. Also, there will be no secondary user served before any interrupted users exist in the system. Obviously, handoff delay and total service time will be lowered in the case of using shared queues compared with the case of using separate queues.

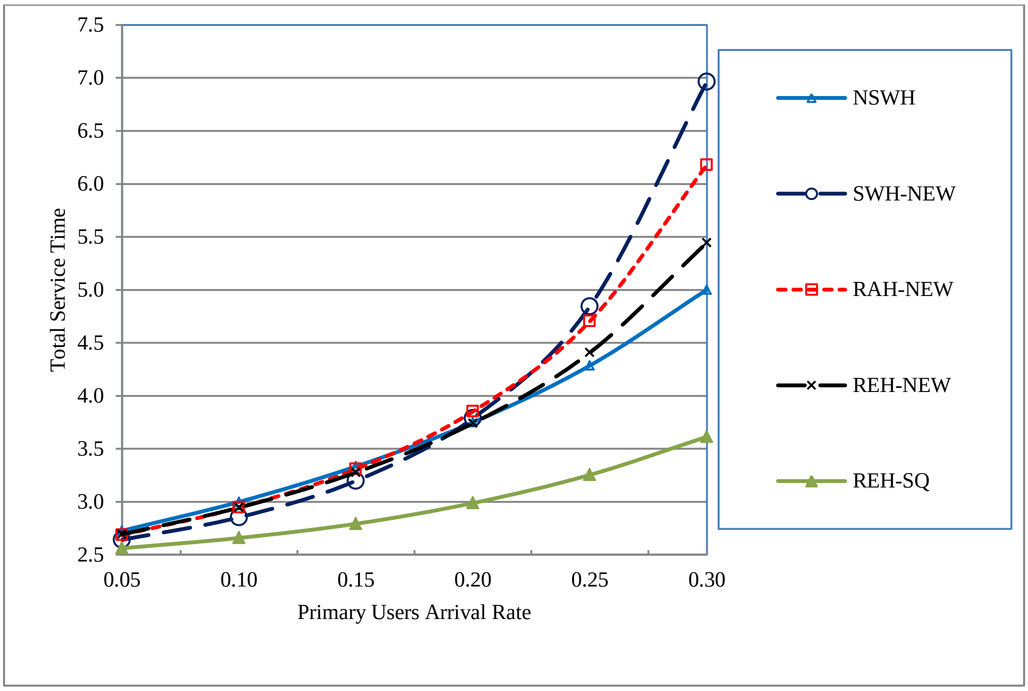

Figure 5.6: Comparison between shared-queue scheme with other implemented schemes.

5.6 Summary

In this chapter, a new reactive-decision spectrum handoff scheme has been implemented. Existing work only considers separate queues for handoff users. The new presented scheme uses a common queue to serve handoff secondary users, in order to increase the efficiency of the utilisation of the available spectrum channels. Achieved results prove that the presented reactive scheme implemented with a common handoff queue reduces the total service time for the secondary users significantly.

Chapter 6.

MODELLING OF COGNITIVE RADIO NETWORKS UNDER HETEROGENEOUS SPECTRUM ENVIRONMENT

6.1 Introduction

6.1.1 Dynamic Spectrum Access (DSA) and Spectrum Handoff

Dynamic spectrum access in a heterogeneous spectrum environment consisting of multi-licensed and multi-unlicensed wireless channels with various access priorities is not explored well yet. Given that the licensed spectrum band is underutilised [98] and the unlicensed spectrum band is overcrowded, there is need to explore this complex situation. The next subsection introduces the users' priority in cognitive radio networks

6.1.2 Users Priority

Wireless systems assign priority to their customers to utilise the spectrum in different ways. As mentioned in Chapter 3, some conventional wireless systems provide higher priority for handoff calls over originating calls, as it is not acceptable to terminate the handoff calls than blocking the new originating calls.

In cognitive radio networks which operate under licensed spectrum bands and centralised architecture (e.g., our works [53], [58], and [56]), a high priority is provided for interrupted secondary users over new arriving secondary users to utilise the available licensed channels. Theoretically, this is achieved by applying the preemptive resume priority queuing network model.

However, in a shared spectrum environment and decentralised network architecture, the challenge is that accessing the unlicensed spectrum bands is a random process. As a result, applying traditional queuing scheduling disciplines such as pre-emptive, FCFS, HOL priority queues, etc. is not applicable. In a random access environment, channel accessing is based on a distributed medium access control (MAC) protocol, such as carrier sense multiple access (CSMA). As a result, channel contention time in CSMA should be taken into consideration when calculating the delay performance measures of the contended secondary users [48]. To provide different priorities for handoff and new secondary users, we adopt the prioritised contention access (PCA) [99] in which different random retrial (also, called backoff) times are assigned to each type of secondary users. The higher the secondary users' priority, the shorter the random back off time. The retrial time of the secondary users is the period of time between two consecutive retrials performed by a secondary user to get the service [76]. According to PCA, handoff users in our model should be assigned a shorter retrial time than the new secondary users. This means, in average, handoff users will wait in the orbit less than the new secondary users before getting a chance to utilise the unlicensed channels. As a result, the quality of service of handoff users can be improved. Fortunately, in a random CSMA environment, communicating SUs do not need to be synchronised to access available channels [32, 76, 100].

This chapter proposed and implemented a prioritised queuing network model to inspect the effects of the spectrum handoff process on the performance of the secondary users in a heterogeneous spectrum environment. Because the analysis of spectrum handoff delay, mathematically, is a very complicated issue, especially under such dynamic circumstances, we decomposed the model into two main parts and validated analytically the resultant parts whenever it is possible in order to have a minimum level of confidence.

6.2 Spectrum Handoff Modelling in a Heterogeneous Spectrum Environment

In Chapter 2, existing work in spectrum handoff in CRNs with the concept of OSA in both licensed and heterogeneous spectrum bands is discussed in detail.

However, most of the existing spectrum handoff models with OSA did not investigate the performance of secondary users in a mixed spectrum environment. This means that these models do not take into consideration that secondary users can operate on both spectrum band environments. In other words, they ignore the possibility of unlicensed bands to become available after some time, and hence can be used for transmission.

As it is known, multiple spectrum handoffs can reduce the QoS of the secondary users, and in order to reduce the handoff delay of the secondary users especially for real time applications, interrupted secondary users should be provided with a higher priority to utilise the available spectrum bands over the new secondary users. This can mitigate the adverse effect of the handoff process. However, existing work does not differentiate between handoff and new originating secondary users, as no priority has been given to handoff users in getting services, which leads to an increase in the handoff delay and total service time of the interrupted users.

Furthermore, all previous studies did not broadly discuss the issue of spectrum handoff operations in terms of spectrum handoff delay.

6.3 Proposed Spectrum Handoff Model

6.3.1 System Model

In this subsection, a prioritised cognitive radio spectrum handoff model is presented. The proposed model analyses the performance of the secondary users operating in a heterogeneous and very dynamic spectrum situation composed of licensed channels, unlicensed channels, and three types of users: primary users, new (original) secondary users, and handoff secondary users. In general, each secondary user is assumed to be implemented with a cognitive radio to allow them to sense the spectrum environment and detect the empty spectrum channels. Secondary users can utilise both types of spectrum channels. This can improve the spectrum usage for cognitive radio networks.

The main objective of the proposed model is to increase the spectrum usage of the secondary users. Moreover, to decrease the number of spectrum handoff for the secondary user. Finally, to decrease the associated spectrum handoff delay.

The proposed system model suggests that PUs and SUs will compete for utilising the spectrum channels using the centralised access point (AP) and random access techniques, as illustrated in Figure 6.1. The model consists of a number of independent wireless channels. There are N_1 licensed channels and N_2 unlicensed channels, as shown in Figure 6.2. The licensed channels share two priority queues: low-priority queues and high-priority queues for secondary and primary users' transmissions, respectively. The queues are implemented with an infinite size for simplicity. In addition, a Pre-emptive Resume Priority (PRP) M/M/C queuing model and a M/M/C retrial queuing model are suggested to manage the spectrum handoff process in the system.

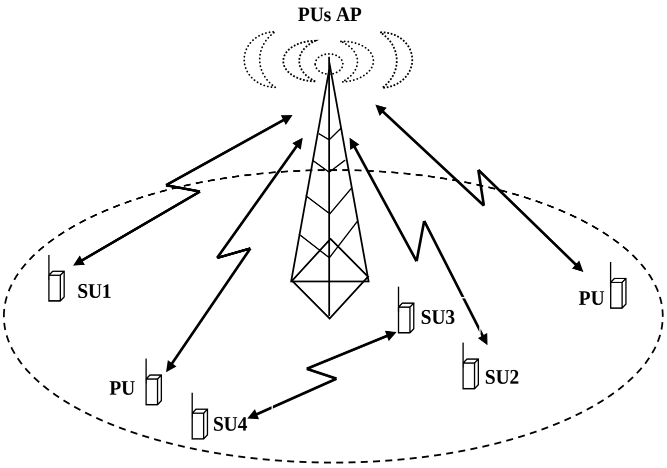

Figure 6.1: Proposed heterogenous CR network.

In general, when a secondary or primary user arrives to the licensed channels (N_1) and finds no idle channel for utilisation, the arriving user will be queued in the corresponding queue type low-priority queue or high-priority queue, respectively. On the other hand, when a secondary user (either handoff or new user) arrives to the unlicensed channels (N_2) and finds no idle channel for usage, the arriving secondary user enters the retrial group (also called the orbit). However, there is no rule of queuing order within the orbit. Instead, the retrial group becomes a source of repeated calls which may be regarded as a sort of infinite queue. But, the order of getting the service depends on the random order in which secondary users return back to check the channels' status and the random chance that a channel is found empty at the time the secondary user returns to the system [101].

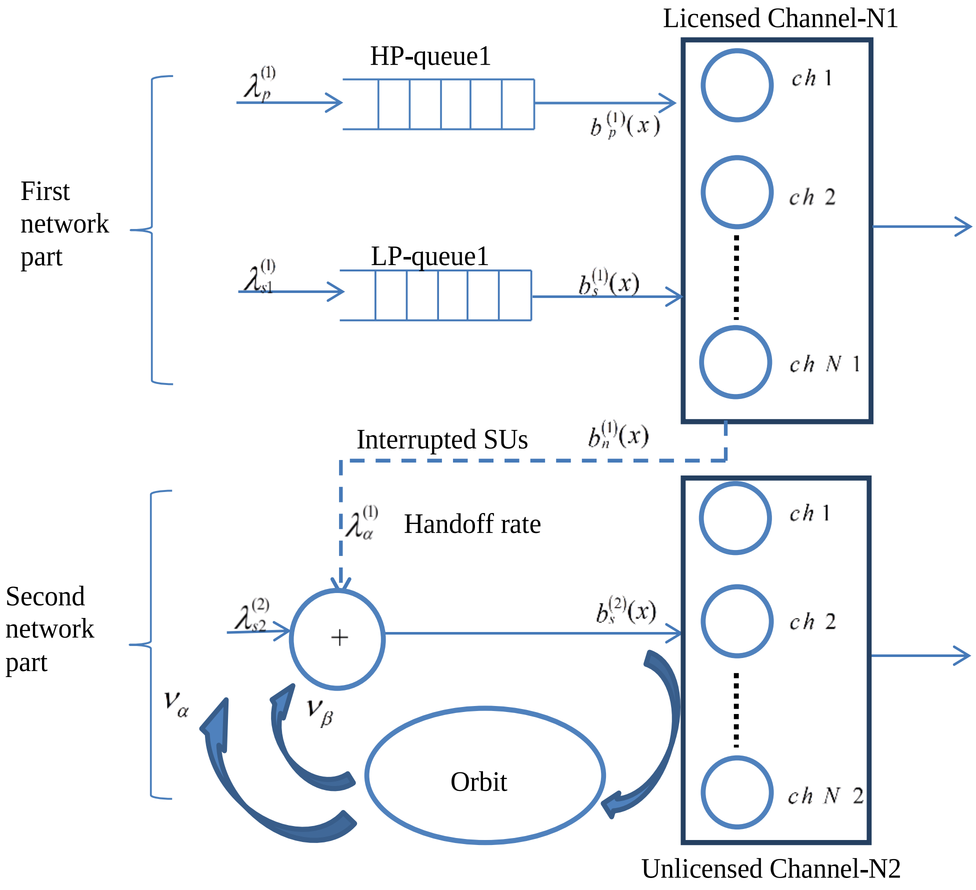

Figure 6.2: Proposed cognitive radio queuing network model operating in a mixed wireless channels (where $n = 1$).

The general assumptions for the proposed model are given below:

- There are two types of spectrum channels that can be utilised in an opportunistic manner: licensed channels and unlicensed channels.

- The available number of licensed channels and unlicensed channels for transmission is N_1 and is N_2, respectively.
- Primary users arrive according to Poisson distribution, with rate λ_p, and depart according to the exponential distribution with rate μ_p, and can utilise only the licensed channels. Upon arriving at the instance of a primary user, if there is at least one idle licensed channel, then the arriving primary user will occupy this idle channel and send its transmission. If all licensed channels are busy and at least one licensed channel is occupied by a secondary user, then the arriving primary user will select and pre-empt one of the licensed channels occupied by the secondary users. If all the licensed channels are busy by primary users, then the arriving primary user joins the high-priority queue of the licensed channels and will be served according to the FCFS queue disciplines.
- Secondary users arrive according to Poisson process, with initial rates λ_{s1} and λ_{s2} and utilise available licensed (N_1) and unlicensed (N_2) radio spectrum channels, respectively, with exponential service rates μ_{s1} and μ_{s2}, respectively.
- If newly arrived secondary users at the licensed channels (N_1) find empty channels, then they will occupy one of those empty channels and transmit their data and leave the system after finishing the data transmission, otherwise if all the licensed channels are busy, either by primary users or secondary users, then the arriving secondary users join the low-priority queue

of the licensed channels and will be served according to the FCFS queue disciplines but after all the primary users exist in the high-priority queue.

- Newly arrived secondary users at the unlicensed channels (N_2) always perform spectrum sensing for some time, say (T_{se}), before starting their transmission. If they find empty channels, then they will occupy one of those empty channels, transmit their data and leave the system after finishing their transmission, otherwise they will enter the retrial group and retry for service after waiting for an exponentially distributed period of time with retrial rate v_{β} .
- Interrupted secondary users will perform the spectrum handoff procedure and arrive at the unlicensed channels with rate λ_{α}. Since the initial secondary users' arrival rate (λ_{s1}) is a Poisson process and will be interrupted by randomly arriving primary users, the resultant interruption rate λ_{α} is also a Poisson process.
- Moreover, since the service time of the secondary users follows the exponential distribution and the handoff users will resume their interrupted transmissions, the remaining service time ($1/\mu_{\alpha}$) of the handoff SUs also follows the exponential distribution with the same rate; i.e. $\mu_{\alpha} = \mu_{s1}$.
- If the handoff secondary user senses some empty channels, the secondary user utilises one of those empty channels and transmits its data and leaves the system after finishing its transmission, otherwise it will enter the retrial group and retry for service after waiting for an exponentially distributed amount of

time with retrial rate v_α. Handoff secondary users require a period of time to switch to the new channel (T_{sw}). Also, in order to notify the intended receiver regarding the selected target channel, the communicating users spend some time to exchange handshaking messages (T_{ha}).

Since interrupted secondary users will handoff to unlicensed channels, this can limit the number of experienced interruptions by each secondary user to one ($n = 1$). Secondary users will keep repeating trials until they get the opportunity and complete sending their transmission data.

6.4 Spectrum Handoff Analysis in a Heterogeneous Spectrum Environment

In this scheme, primary users can utilise only licensed spectrum channels (N_1) while secondary users can use both licensed and unlicensed spectrum channels. In the case of interruptions, interrupted secondary users have to move to the unlicensed spectrum channels (N_2) and will not be interrupted again since all the users in the unlicensed channels will be served with equal rights. In other words, all secondary users have the same right to utilise the available unlicensed channels.

Upon the instant they arrive at the N_2 channels, both the handoff and new SU users will be served immediately if they find empty channels, otherwise they will enter the retrial queue and will retry for service after waiting for a random period of time. In general, to apply some kind of priority for a particular class of users over another class in a random access environment, the length of the random period of waiting time in the orbit is used. For example, users with higher priority must wait shorter periods of time than users with a lower priority before they can retry for service. As a result, on average, high priority users will be served before any other low priority users. In this model, handoff users will be given high priority over new SUs to utilise available unlicensed channels (N_2) by assigning them with shorter retrial times which lead to shorter average waiting times in the orbit.

Generally, the analysis of retrial queues is a difficult problem except for limited simple cases [101]. The degree of difficulty further increases if there is more than one class of users [14], as proposed in this chapter. Moreover, when a heterogeneous radio spectrum environment, which consists of licensed and unlicensed channels is considered, it implies the usage of two different channel access technologies; contention-free and contention-based. Therefore, in order to facilitate validation of our proposed cognitive radio network queuing model which operates in a heterogeneous spectrum environment and consists of multi-channel and multi-class of users, we decomposed our proposed queuing network model into two main parts as shown in Figure 6.3 and Figure 6.4; the first part illustrated in Figure 6.3 is just the ordinary M/M/C queue with preemptive resume priority disciplines, whereas the second part (Figure 6.4) is the ordinary M/M/C retrial queue. The interruption rate (also called the handoff rate) output from the first part of the model will be the input to the second part of the model, besides the ordinary input to the second part. In other words, interrupted secondary users and new secondary users will arrive at the second network part - as two inputs - with mean rates λ_α and λ_{s2}, respectively.

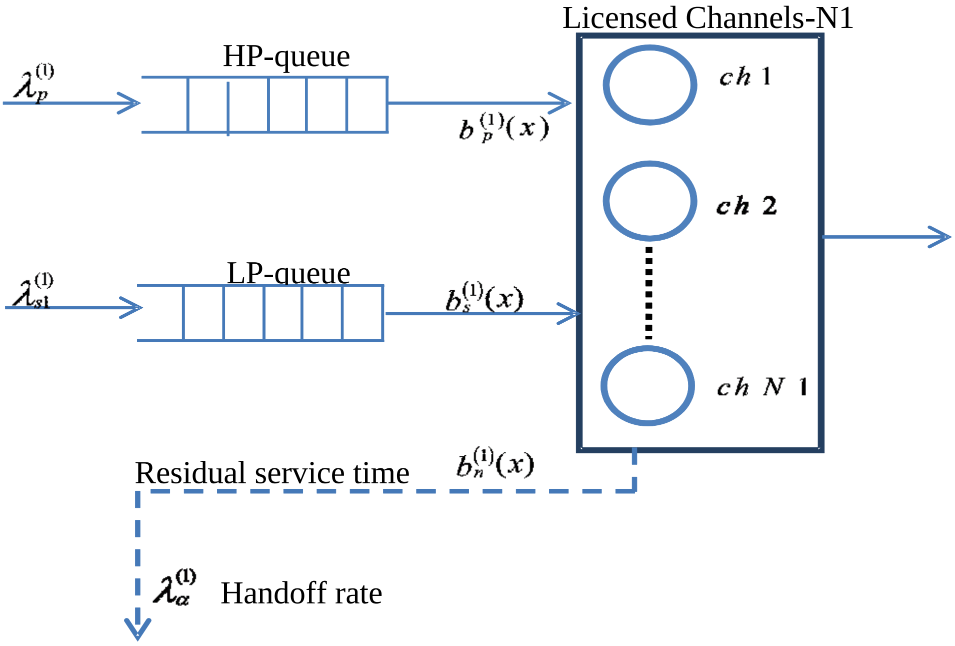

Figure 6.3: First network part: M/M/C queue with preemptive resume priority disciplines (n = 1).

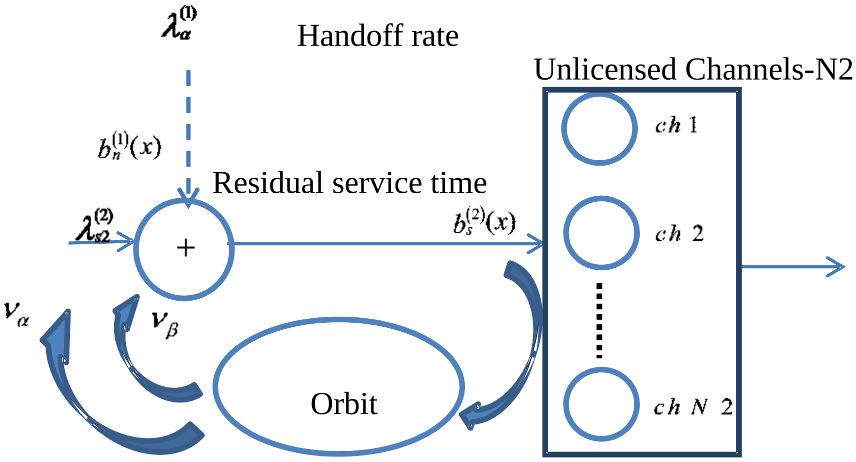

Figure 6.4: Second network part: M/M/C retrial queue with two input arrivals and two retrial rates (n = 1).

Single-licensed and single-unlicensed channel, and multi-licensed and multi-unlicensed channel scenarios will be analysed and discussed in the next sections.

Moreover, the former scenario is considered with identical and non-identical traffic rates for handoff and new arrival SUs'. Whereas, the latter scenario is considered with different numbers of licensed and unlicensed spectrum channels.

6.4.1 Scenario 1: Single Licensed and Unlicensed Channels (N_1=1, N_2=1)

In this scenario, an analytical model for the proposed queuing network model, which comprises of two parts, can be found in the literature, but separately and sometimes with needing to some assumptions to be valid for our case.

For the first part of the network model illustrated in Figure 6.5, the interesting parameter is the interruption rate (λ_α) since interrupted users will arrive at the second part of the network and will affect the performance of the network directly. The expression of λ_α is mentioned in [64] and can be written as:

$$\lambda_\alpha = \alpha \lambda_{s1} \tag{6.1}$$

Where α refers to the interruption probability which is equal to:

$$\alpha = \frac{\lambda_p}{\lambda_p + \mu_{s1}} \tag{6.2}$$

From the above equation it is clear that the interruption probability α increases as the primary user arrival rate λ_p increases or when the secondary user tends to use the transmission channel for longer periods. Interruption probability refers to the probability that the SU in service will be interrupted by arriving PUs.

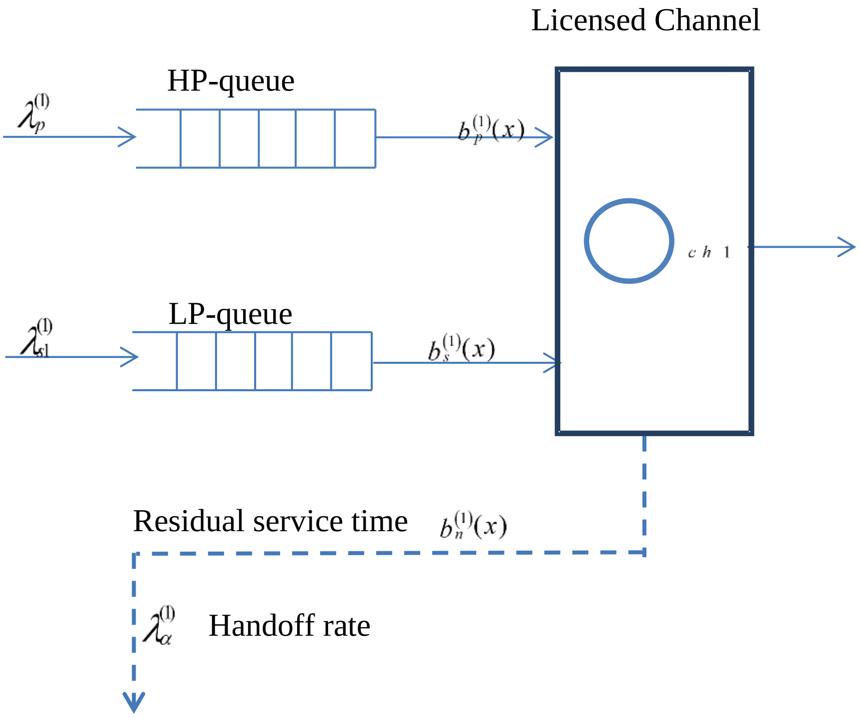

Figure 6.5: First network part: M/M/1 queue with preemptive resume priority disciplines (n = 1).

For the second part of the network model illustrated in Figure 6.6, existing analytical models derive the average waiting delay in the orbit for only a single-class of users. The average waiting time in the orbit W can be defined as the average time spent in the orbit until finding the channel empty and starting the transmission [101]. The average waiting time W is mentioned in [101, 102] and is expressed as follows:

$$W = \frac{1}{\mu}\left(\frac{\rho}{1-\rho}\right)\left(1+\frac{\mu}{\nu}\right) \qquad 6.3$$

Where μ and ν refer to the service and retrial rates of the assumed single-class of users, and ρ refers to the channel utilisation which is equal to the ratio of the arrival rate to the service rate:

$$\rho = \lambda / \mu \qquad 6.4$$

The second part of the proposed network model has two classes of users which will arrive and compete randomly for the service. As a result, some assumptions should be made to the existing analytical formula (6.3) in order to become valid for analysing the proposed model. Actually, the proposed model, assumes two classes of users (handoff and new SUs) instead of one class - as existing model considers-. As a result the parameters λ, μ and ν should be defined again to fit the proposed model as shown below.

Therefore, λ will represent the aggregate arrival of the handoff and new SUs, and can be expressed as:

$$\lambda = \lambda_{\alpha} + \lambda_{s2} \qquad 6.5$$

Also, since, the service time of the secondary users μ_{s2}^{-1} follows the exponential distribution the remaining service time of the handoff secondary users μ_{α}^{-1}, as a consequence, also follows the exponential distribution with the same rate, which is equal to μ_{s2}^{-1}. This also means that all marginal ($\mu_{s1}, \mu_{s2}, \text{and}\, \mu_{\alpha}$), and aggregate ($\mu$) service rates are equal:

$$\mu = \mu_{s1} = \mu_{s2} = \mu_{\alpha} \qquad 6.6$$

Moreover, the equivalent retrial rate ν can be expressed using the retrial rate for the handoff and the new arriving SUs, say ν_{α} and ν_{β}, respectively, as follows:

$$\nu = \left(\frac{\lambda}{\lambda_{\alpha}}\right)\nu_{\alpha} + \left(\frac{\lambda}{\lambda_{s2}}\right)\nu_{\beta} \qquad 6.7$$

To sum up, λ, μ and ν are defined as the aggregate arrival, service and retrial rates, respectively of the handoff and the new SUs.

Now two handoff schemes will be considered as non-prioritised and prioritised spectrum handoff, as explained below.

6.4.1.1 Non-prioritised Spectrum Handoff Scheme

Consider the special case where both the handoff and new secondary users have identical traffic parameters, as given in Equations 6.6, and the next two Equations 6.8 and 6.9.

$$\nu_{\alpha} = \nu_{\beta} \qquad 6.8$$

$$\lambda_{\alpha} = \lambda_{s2} \qquad 6.9$$

As a result, it is expected to achieve same average waiting delay for both the handoff and the new users in the unlicensed channel since all their traffic rates are identical.

If we denoted the marginal average waiting delays as W_{α} and W_{β} for the handoff and the new arriving SUs, respectively, then W_{α} and W can be written as:

$$W_{\alpha} = \frac{1}{\mu}\left(\frac{\rho_{\alpha}}{1-\rho_{\alpha}}\right)\left(1+\frac{\mu}{\nu_{\alpha}}\right) \qquad 6.10$$

$$W_{\beta} = \frac{1}{\mu}\left(\frac{\rho_{\beta}}{1-\rho_{\beta}}\right)\left(1+\frac{\mu}{\nu_{\beta}}\right) \qquad 6.11$$

Where, ρ_{α} and ρ_{β} denotes the marginal channel utilisation factors by the handoff and new SUs, respectively, and can be defined as:

$\rho_{\alpha} = \lambda_{\alpha}/\mu$ and $\rho_{\beta} = \lambda_{s2}/\mu$

While the aggregate utilisation factor (ρ) of the unlicensed channel can be computed as:

$$\rho = \rho_{\alpha} + \rho_{\beta} \qquad 6.12$$

Clearly, the aggregate average waiting delay (W) can be expressed using the marginal delays W_{α} and W_{β} as:

$$W = \left(\frac{\lambda}{\lambda_{\alpha}}\right)W_{\alpha} + \left(\frac{\lambda}{\lambda_{s2}}\right)W_{\beta} \qquad 6.13$$

, where,

$$W = W_\alpha = W_\beta \tag{6.14}$$

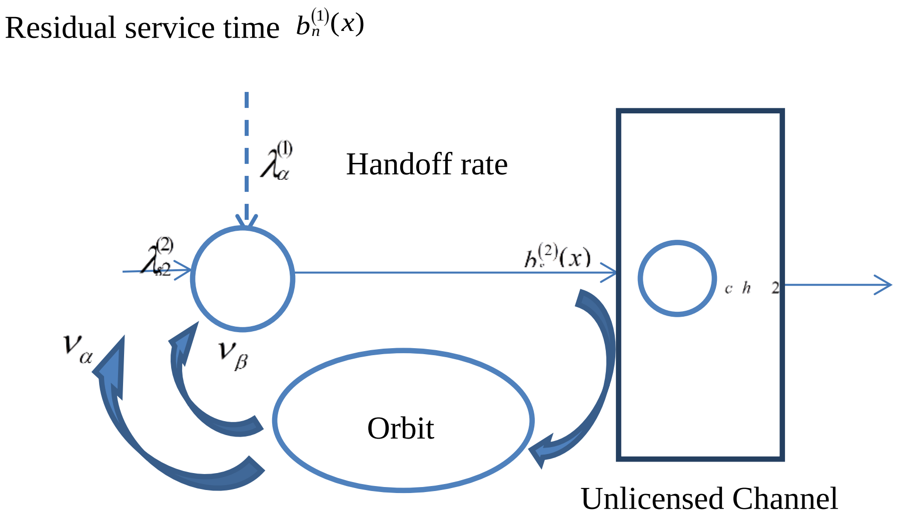

Figure 6.6: Second network part: M/M/1 retrial queue with two input arrivals and two retrial rates (*n* = 1).

To sum up, Equation 6.3 can represent the aggregate and marginal average waiting delays for handoff and new secondary users by using corresponding traffic rates.

To consider the handoff delay ($E[D]$), additional incurred delays, such as the wideband sensing time (T_{se}), the switching time (T_{sw}), and the handshaking time (T_{ha}), should be included in the average waiting delay as well.

$$E[D] = W_\alpha + T_{se} + T_{sw} + T_{ha} \tag{6.15}$$

If we neglect the effect of the additional delays, then *E[D]* can be written as :

$$E\left[D\right]=W_{\alpha} \qquad 6.16$$

Henceforth, W_{α} will either be called the handoff delay or the waiting (queuing) delay for handoff secondary users.

6.4.1.2 Prioritised Spectrum Handoff Scheme

When the traffic rates of handoff and new SUs are not all identical, in particular, the case when $v_{\alpha} > v_{\beta}$ is considered, handoff users are given a higher priority over the original SUs to utilise unlicensed channels. The existing analysis for the average waiting delay can only represent the aggregate value. In this case, we would expect that handoff SUs will experience less average waiting delay than the new SUs in the unlicensed channels, since the retrial rate of the former is always higher than that of the latter.

6.4.2 Scenario 2: Multi-licensed and Multi-unlicensed Channels (N_1 = n1, N_2 = n2)

When we extend the suggested queuing network model to the general case in which the system operates under multi-licensed and multi-unlicensed channels, then the complexity of the proposed model will increase besides the issues mentioned above, and so, as a consequence, only the simulation results are provided in this case.

Obviously, marginal utilisation factors can be computed analytically using $\rho_\alpha = \lambda_\alpha / n_2\mu_{s2}$ and $\rho_\beta = \lambda_{s2} / n_2\mu_{s2}$ besides the aggregate factor by $\rho = \lambda / n_2\mu_{s2}$ for all of the above mentioned scenarios and cases.

6.5 Simulation Experiments and Numerical Results

In order to evaluate the performance of the proposed CR network model, we can perform extensive simulation experiments in a heterogeneous spectrum environment with single-channels and multi-channels. The simulation experiments have been achieved with the discrete event simulator (MATLAB) tool to analyse the aggregate and the marginal waiting delay for the secondary users, for both of the considered scenarios. Moreover, non-prioritised and prioritised spectrum handoff schemes are considered, too. Although a 95% confident interval is used to evaluate the accuracy of the achieved results, the variations of the confidence interval is relatively high in Figure 6.20 and Figure 6.21, this due to the reduction of the number of handoff events by increasing the number of licensed and unlicensed channels from 1 to a higher number. In general, this leads to decreases in the number of samples used to calculated the average in the confidence interval. Also, we assume secondary users have the ability to sense perfectly the available spectrum bands, which means that the detection of PUs is perfect. Simulation results have been comprehensively compared with the results achieved from the analytical model whenever it is possible.

6.5.1 Performance Calculations

In general, for simulation experiments the marginal, W_α and W_β, and aggregate W average waiting times spent in the orbit before getting the service can be calculated using the following formulae:

$$W_\alpha = \frac{\sum Waiting\ Delay\ of\ Handoff\ SUs}{Number\ of\ Handoff\ Users} \tag{6.17}$$

$$W_\beta = \frac{\sum Waiting\ Delay\ of\ New\ Arriving\ SUs}{Number\ of\ New\ SUs\ Arriavls} \tag{6.18}$$

Hence, the aggregate average is just the sum of the marginal averages which can be written as:

$$W = \frac{\sum (Delay\ of\ Handoff\ SUs + Delay\ of\ New\ SUs)}{Number\ of\ Handoff\ Users + Number\ of\ New\ SUs} \tag{6.19}$$

6.5.2 Scenario1: (N_1 = 1, N_2 = 1)

In this scenario, two schemes are assumed: non-prioritised and prioritised spectrum handoff schemes. Table 6.1 summarises various simulation parameters with corresponding rates, which are used for the first scenario, unless otherwise stated.

Parameter	**Symbol**	**Value(s)**					
CH1 SU arrival rate	λ_{s1}	6.50					
Interruption/handoff rate	λ_{α}	1.00	2.00	3.00	4.00	5.00	6.00
CH2 SU arrival rate	λ_{s2}	1.00	2.00	3.00	4.00	5.00	6.00
PU arrival rate	λ_{p}	2.36	5.78	11.14	20.80	43.33	156.00
PU service rate	μ_{p}	315.00					
SUs' service rates	μ_{s1}, μ_{s2}	13.00					
Handoff SUs retrial rate	υ_{α}	5,10,15,20					
New SUs retrial rate	υ_{β}	5,10,15,20					

Table 6.1: Simulation parameters for first scenario.

6.5.2.1 Non-prioritised Spectrum Handoff Scheme

Existing models do not differentiate between original and handoff SUs. Both types of SUs have same priority to access the unlicensed channel, which is expressed here as $\nu_{\alpha} = \nu_{\beta}$. Results presented in this part are in line with such assumptions and equality. Graphs provided in Figure 6.7 to Figure 6.10 display a comparison between simulation and analytical results. Figure 6.7 and Figure 6.8 show the corresponding delays for handoff and new users, respectively. The average waiting delay in the orbit is drawn versus the secondary users' arrival rate in order to achieve the comparison. Theoretical results in all of the figures can represent both the marginal and aggregate delay values as they are equal in this case. As expected, as the SU arrival rate increases, the corresponding delay increases as well. Generally, graphs show that the results obtained from simulation experiments and those produced from analytical models are exactly the same.

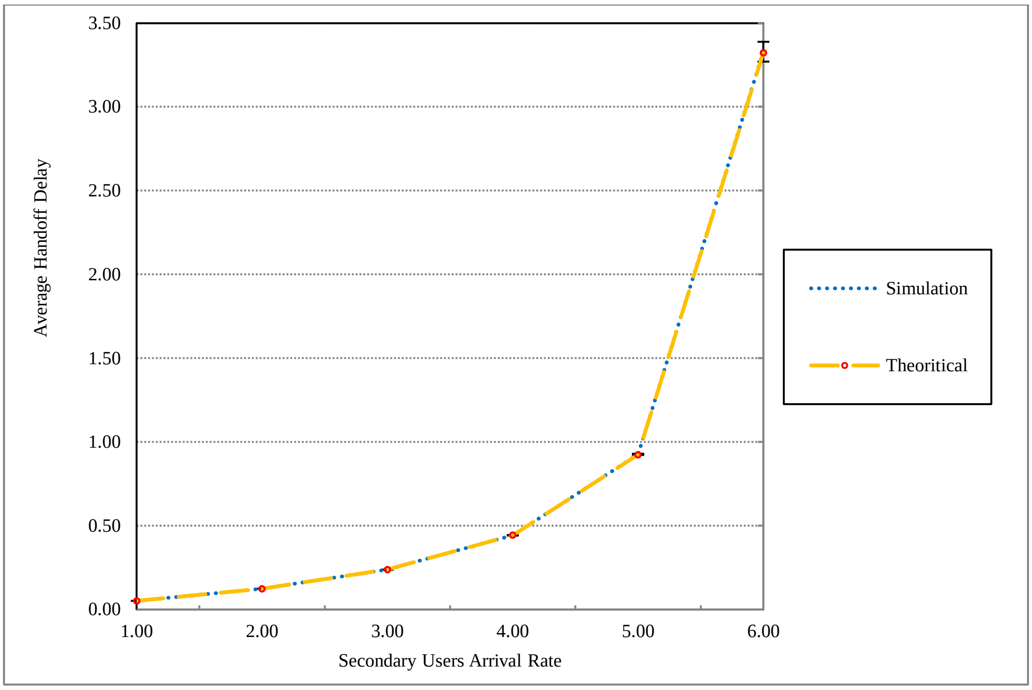

Figure 6.7: Average handoff delay for handoff SUs with ($\nu_\alpha = 5$, $\nu_\beta = 5$).

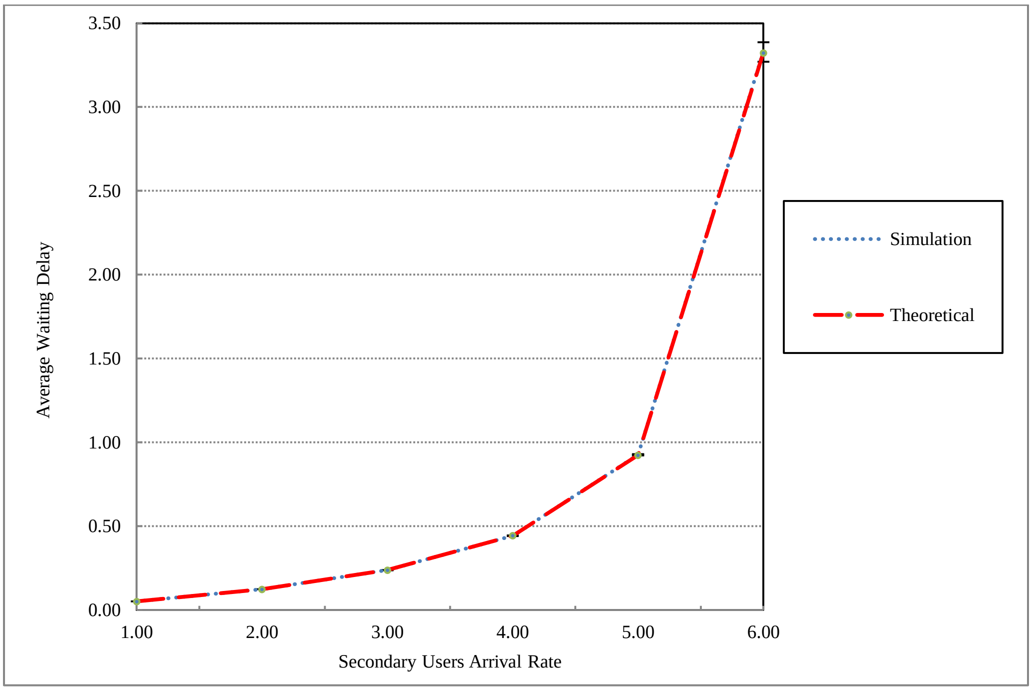

Figure 6.8: Average waiting delay for new SUs with ($\nu_\alpha = 5$, $\nu_\beta = 5$).

Figure 6.9 compares the results obtained from the analytical model and the simulation experiments for the average delay of both the new and handoff SUs. As expected, since all traffic rates are equal for both types of the SUs, the resultant waiting delay in the orbit is equal as well.

Figure 6.10 compares the performance of the proposed network for a range of equal retrial rates of the new and handoff SUs. From the graph, it is clear that the average handoff delay decreases with the increase of the retrial rates. This can be understood as, when the retrial rates ν_α and ν_β increase, the retrial time decreases which leads to a decrease in the average waiting time in the orbit until finding an opportunity to transmit the data.

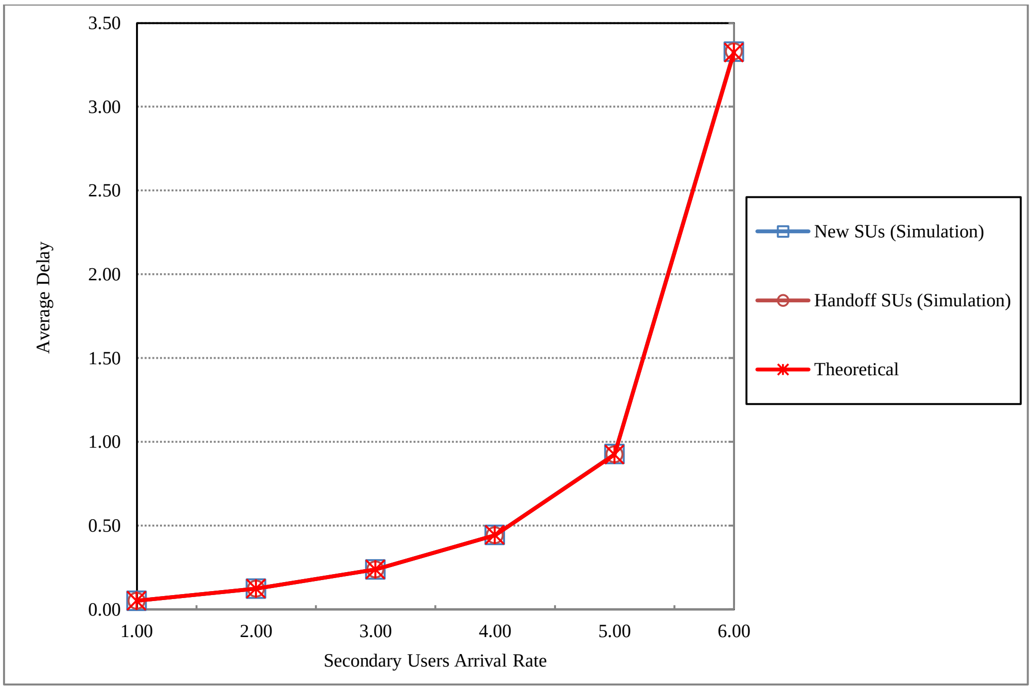

Figure 6.9: Average waiting delay with ($\nu_\alpha = 5$, $\nu_\beta = 5$).

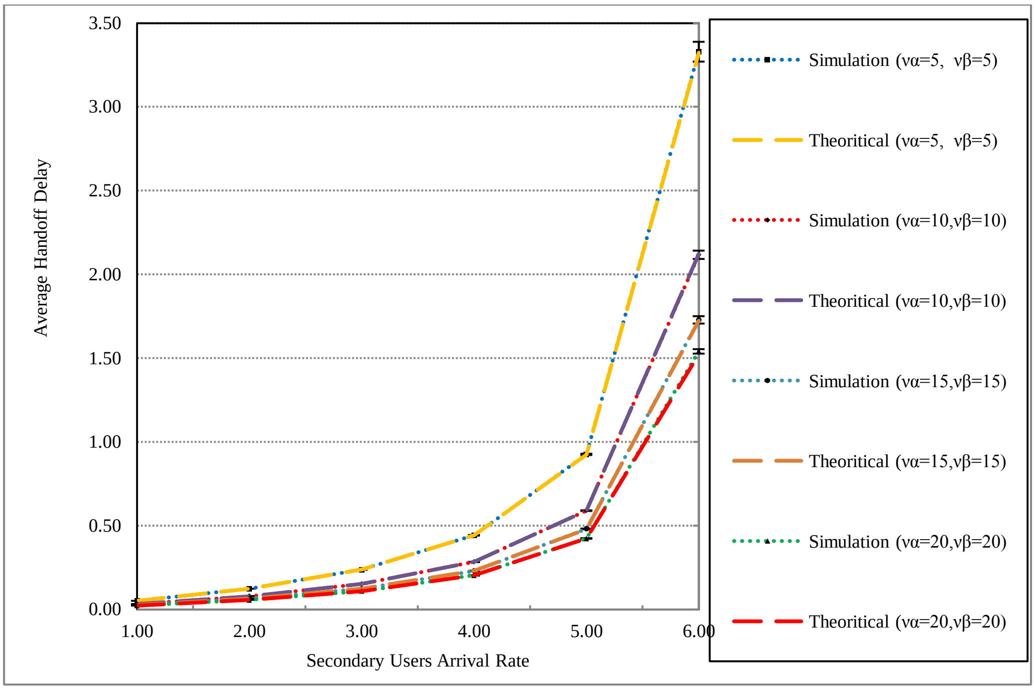

Figure 6.10: Average handoff delay for a range of equal retrial rates.

6.5.2.2 Prioritised Spectrum Handoff Scheme

In this case, simulation parameters illustrated in Table 6.1 are still valid. However, the only change assumed here is that retrial rates are different (i.e., always $\nu_\alpha > \nu_\beta$). In other words, handoff secondary users will generally have a higher priority over the new secondary users to utilise the available spectrum channels, and this can clearly decrease their handoff delay. Moreover, ν_β is assumed to be constant (i.e. $\nu_\beta = 5$) whereas ν_α can be varied, as illustrated in the following graphs.

Figure 6.11 to Figure 6.18 compare simulation and/or analytical results for the original and handoff SUs under a range of retrial rates. For reasons of comparison, the figures also include the non-prioritised scheme results (i.e., when both retrial rates are equal $\nu_\alpha = \nu_\beta$).

The next three figures (Figure 6.11- Figure 6.13) show illuminate results of the proposed model. As shown by the graphs, handoff SUs always experience a less average waiting delay than the new SUs, which can improve their quality of service QoS. This can be interpreted in such a way that, since handoff SUs will wait for shorter average periods of time in the orbit than the new SUs (this is because of $\nu_\alpha > \nu_\beta$), the resultant delay in the orbit for handoff users will be, on average, less than that of the new users.

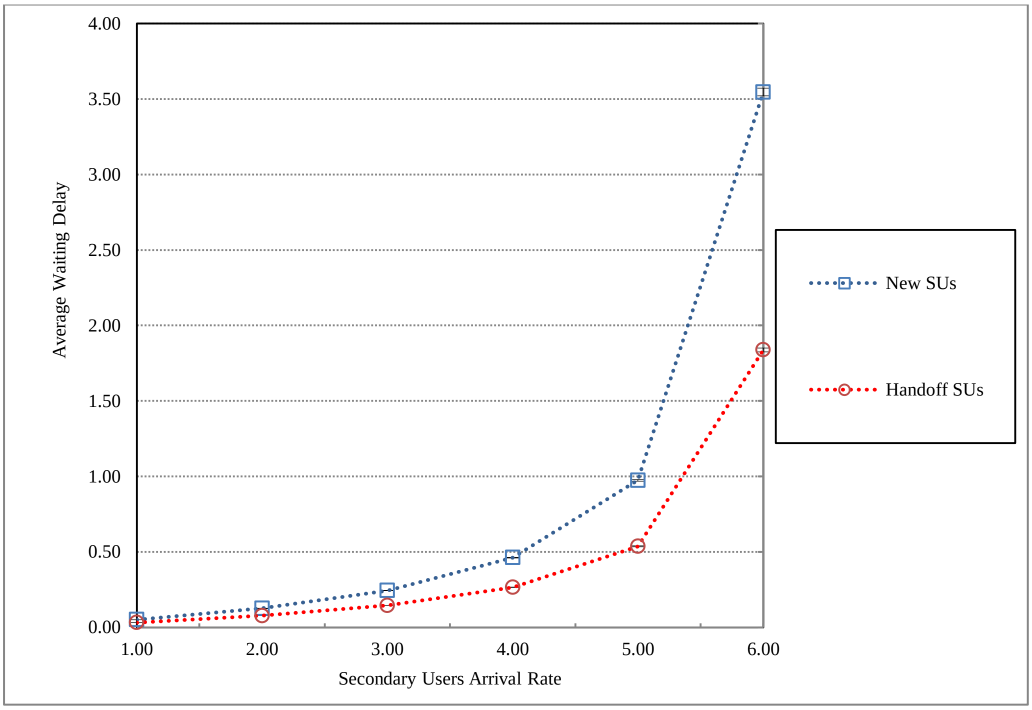

Figure 6.11: Simulation results for average waiting delays ($\nu_\alpha = 10$, $\nu_\beta = 5$).

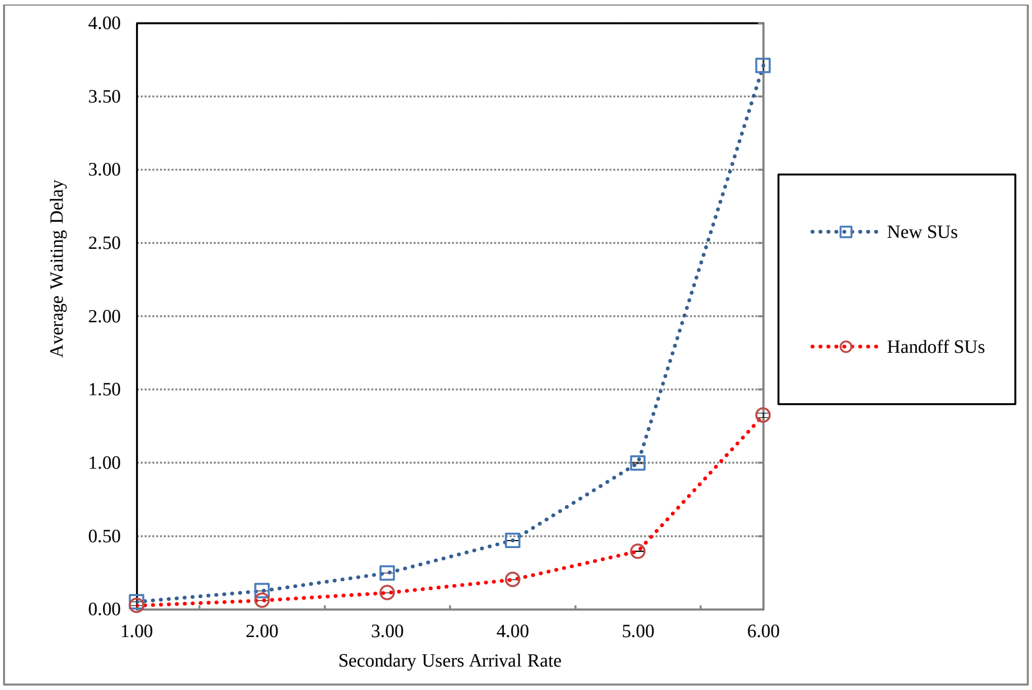

Figure 6.12: Simulation results for average waiting delays ($\nu_\alpha = 15$, $\nu_\beta = 5$).

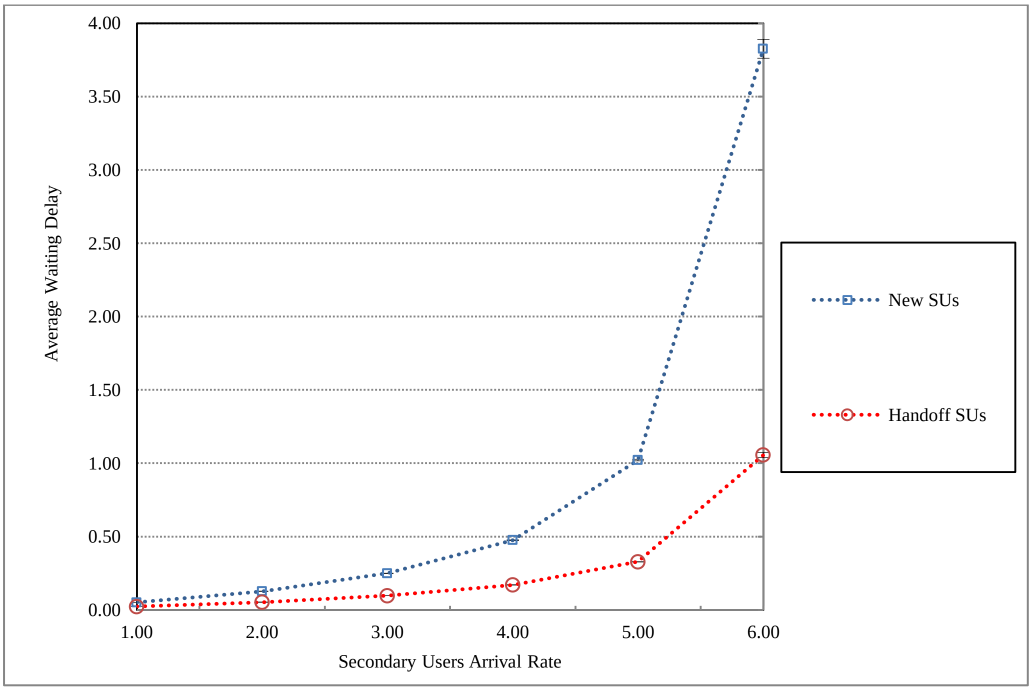

Figure 6.13: Simulation results for marginal delays ($\nu_\alpha = 20$, $\nu_\beta = 5$).

Figure 6.14 shows simulation results to investigate the performance of - only - handoff secondary users under various retrial rates. As shown in the graph, the average handoff delay decreases with the increase of the handoff users' retrial rate (ν_α) for a constant original user retrial rate ($\nu_\beta = 5$). For example, when ν_α increases from 5 to 10, the average handoff delay would decrease by approximately 45% at 6 traffic loads. On the other hand, when ν_α increases from 5 to 20 for the same traffic loads, the improvement in the average handoff delay would exceed 68%. This is obvious and can be understood as, for high values of ν_α, handoff users will experience, on average, short waiting times in the orbit before they retry for service, which leads to a reduction in the average handoff delay. On the contrary, for low ν_α, handoff users will spend longer waiting times on average before they repeat the trial for the service, which causes an increase in the average handoff delay.

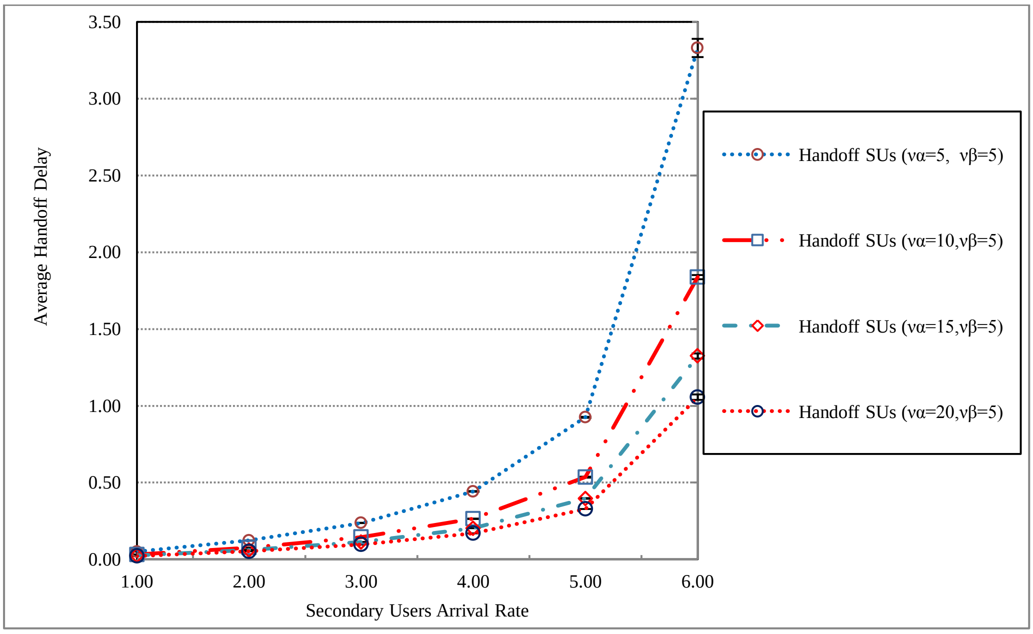

Figure 6.14: Effect of SUs retrial rates on the handoff delay (Simulation results).

Figure 6.15 to Figure 6.17 draw together the simulation results for the marginal and aggregate delays of handoff and new SUs. As expected, the aggregate delay graph is always situated between the two marginal delays. Also, the handoff delay graph is always located under the aggregate graph. On the other hand, the waiting delay for the new SUs graph is always situated above the aggregate graph. This is clear as the aggregate delay represents the average of the marginal delays; consequentially, it should always be located between the two marginal delays. But when $\nu_\alpha = \nu_\beta$, all delays are equal and appear as one graph, as shown in Figure 6.9.

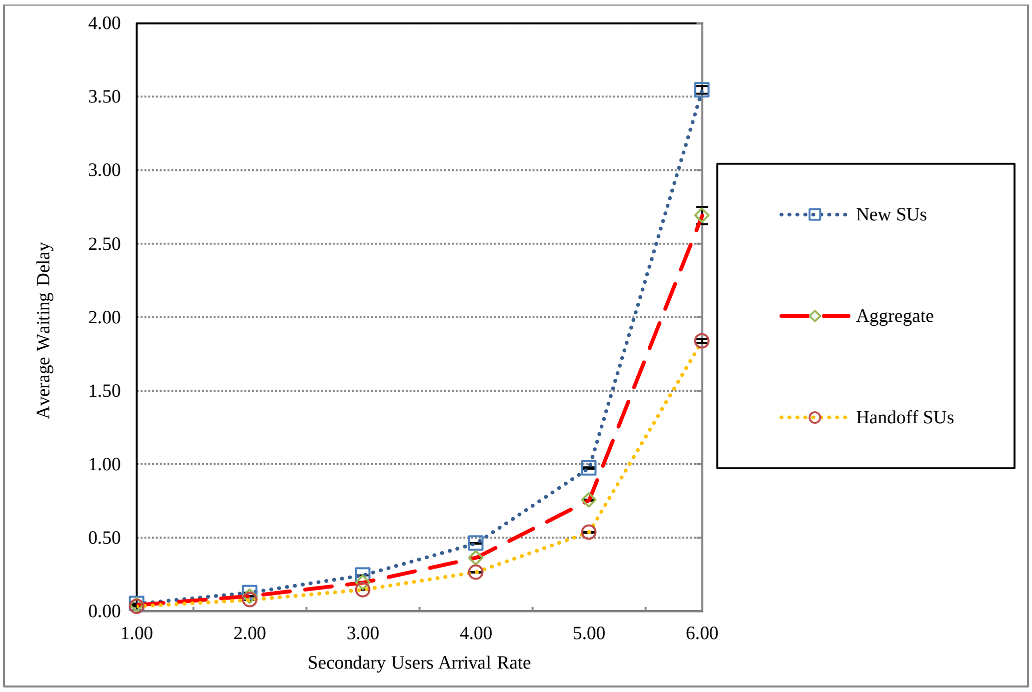

Figure 6.15: Simulation results for marginal and aggregate delay ($\mathcal{V}\boldsymbol{\alpha}$=10, $\mathcal{V}_{\boldsymbol{\beta}}$=5).

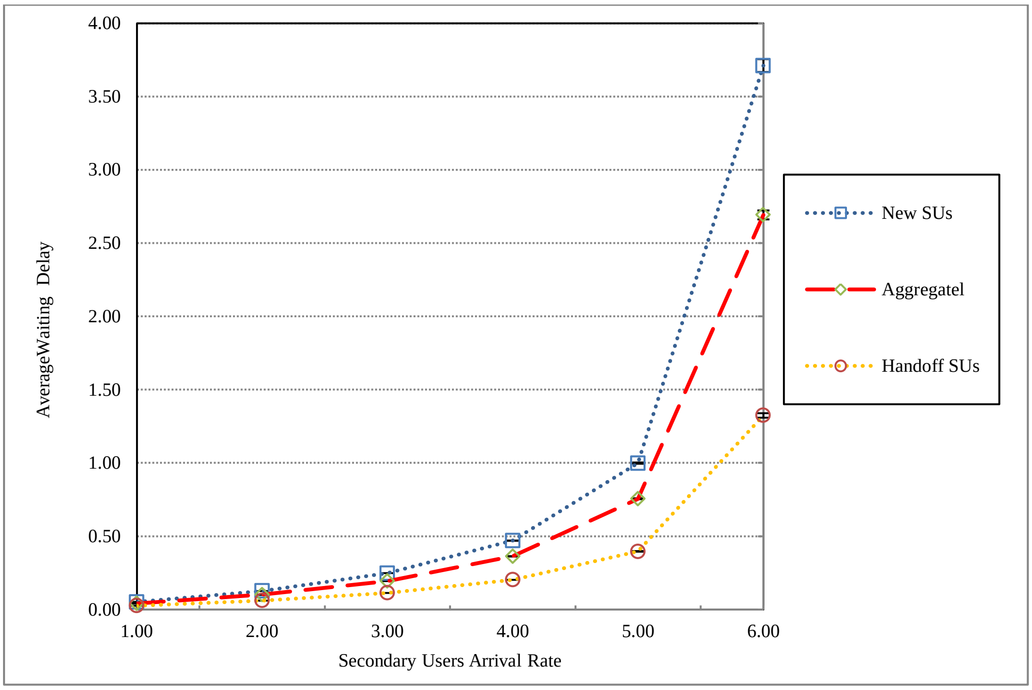

Figure 6.16: Simulation results for marginal and aggregate delay ($\mathcal{V}_{\alpha}$=15, $\mathcal{V}_{\beta}$=5).

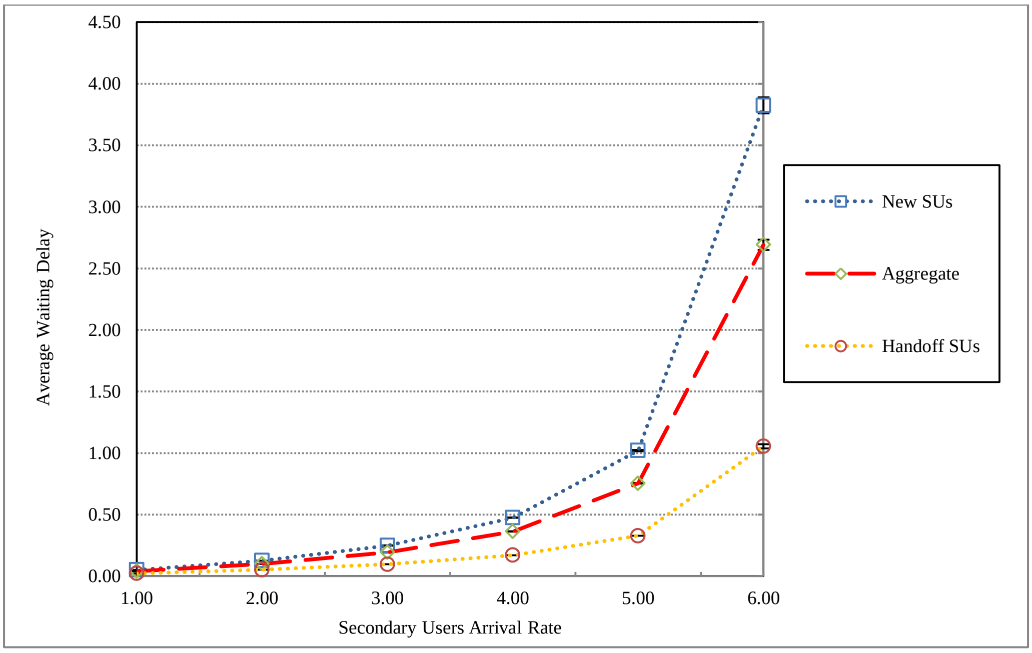

Figure 6.17: Simulation results for marginal and aggregate delay ($v_{\alpha} = 20$, $v_{\beta} = 5$).

The aggregate simulation results are validated using the results produced from the analytical model and give exactly same results as shown in Figure 6.18.

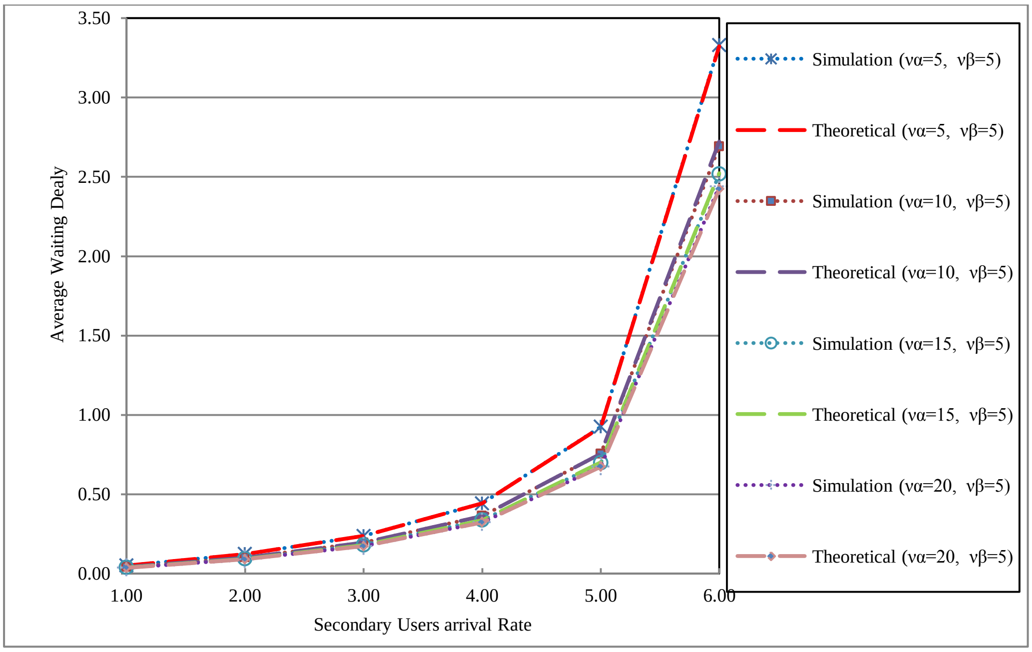

Figure 6.18: Effect of SUs retrial rates on the average aggregate waiting delay.

The achieved results in Figure 6.7 to Figure 6.18 show that, as the SUs' arrival rates increase, the average waiting delays increase as well. This can be interpreted as follows: for a constant service rate, when the arrival rate of the SUs increases, the probability that new, handoff, or retrial users find the channel being busy becomes high. When the user finds the channel busy, a random back off time is generated, and it waits in the orbit, this procedure will be continued until the user gets service. Therefore, this will increase the average waiting delay in the orbit for the SUs.

6.5.3 Scenario 2: (N_1=n1, N_2=n2)

In this section, the performance of the secondary users will be investigated by simulation experiments under different numbers of licenced N_1 and unlicensed N_2 channels. When we consider the special case in which N_2 is set to 1, the second part of the model can also be analysed analytically, as it has been done in the first scenario. In other words, Equation 6.1 is still valid, but with deduction the interruption probability ($\boldsymbol{\alpha}$) from the simulations. Simulation parameters for this section are illustrated in Table 6.2. Table 6.3 shows the relationship between the number of channels that belong to a spectrum band and the interruption rate λ_{α}. The table reveals that the interruption rate only relies on the number of licenced channels N_1 and does not depend on the number of unlicensed channels N_2; this is obvious because the interruption process always occurs in the licensed channels N_1. Also, the table demonstrates that, as the primary user's arrival rate λp increases, the interruption rate $\lambda\boldsymbol{\alpha}$ increases as well. For example, when $N_1 = 2$ and assuming that λ_p increases from 1 to 6, the interruption rate λ_{α} would increase gradually from 0.0802 to 0.7283 as well. This is because of the fact that, as λ_p increases, the probability that the secondary user under the service will be interrupted increases as well, which leads to an increase in the handoff rate.

On the contrary, it can be seen that, as the number of licensed channels N_1 increases, the interruption rate λ_{α} decreases. For example, at the low rates of PU, say $\lambda_p = 2$, and when $N_1 = 2$, the interruption rate is 0.1919 and decreases to 0.0380 when $N_1 = 3$ and to 0.0055 when $N_1 = 4$. This can be understood as follows: for a given PU arrival rate, when the number of licensed channels increases, the equivalent number of secondary users that can be served at the same time increases too. Also, since the arrived PU can interrupt and pre-empt only one SU at the same time, the number of interrupted SUs decreases, which leads to a decrease in the interruption rate.

Parameter	**Symbol**	**Value(s)**
PU arrival rate	λp	1.0 – 6.0
PU service rate	μp	10.00
SU1 arrival rate	λ_{s1}	3.00
SU2 arrival rate	λ_{s2}	1.00 - 6.00
SUs’ service rate	μ_{s1}, μ_{s2}	8.00
Handoff SUs retrial rate	υ_{α}	10.00
New SUs retrial rate	υ_{β}	5.00

Table 6.2: Simulation parameters for second scenario.

λ_p			1	2	3	4	5	6
$N_1 = 2$	$N_2 = n2$	λ_α	0.0802	0.1919	0.3188	0.4556	0.5921	0.7283
$N_1 = 3$			0.0125	0.0380	0.0780	0.1316	0.1982	0.2754
$N_1 = 4$			0.0014	0.0055	0.0136	0.0271	0.0474	0.0742

Table 6.3: Simulation results for spectrum handoff rate (where n2=1, 2, 3).

Figure 6.19 to Figure 6.21 describes the effect of the changing number of channels N_1 and N_2 on the average handoff delay. In each figure, N_2 is set to be constant while N_1 is variable. From the graphs, it can be recognised that, for a given number of unlicensed channels N_2, the average handoff delay decreases with an increase in the number of licensed channels N_1. For example, in Figure 6.19, when the primary users' arrival rate is 6 and N_1 increases from 2 to 3, the average handoff delay would decrease by almost 30%. This can be explained as, when N_1 increases, the number of secondary users that can be served at the same time increases as well; this means that more SUs will finish their transmission in the N_1 channels, which decreases the handoff rate λ_α. Therefore, the decrease in λ_α will lead to a decrease in the number of users which enter the orbit, meaning that there is a decline in the average handoff delay, according to Little's formula.

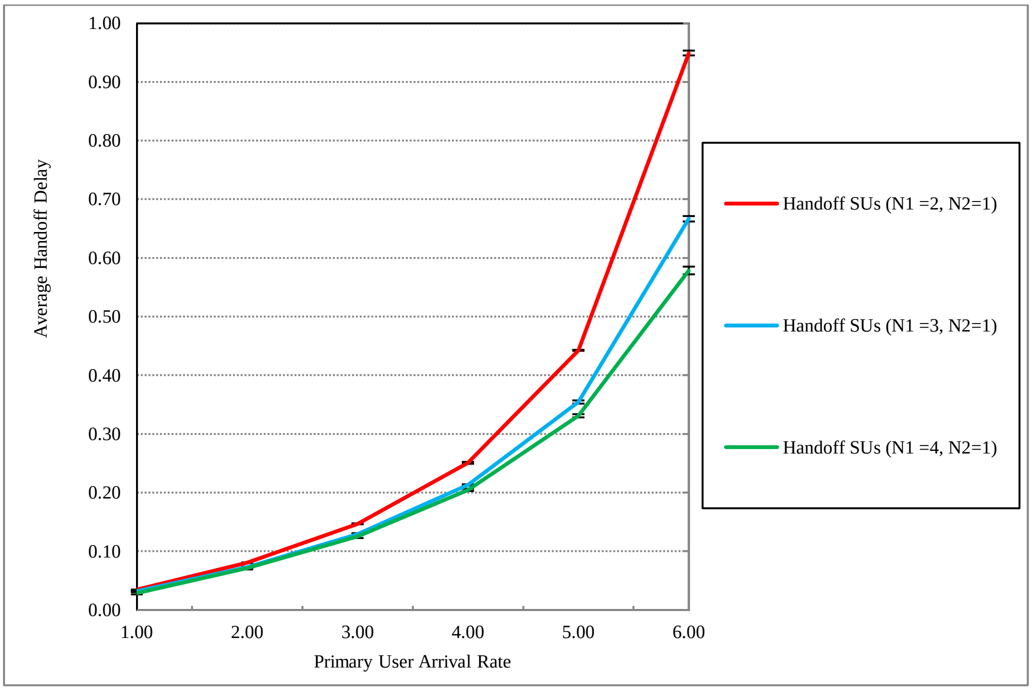

Figure 6.19: Average handoff delay.

In Figure 6.21 the number of unlicensed channels is set to 3 in order to show that our achieved results can be realistic since the standard IEEE802.11b also uses three non-overlapping channels.

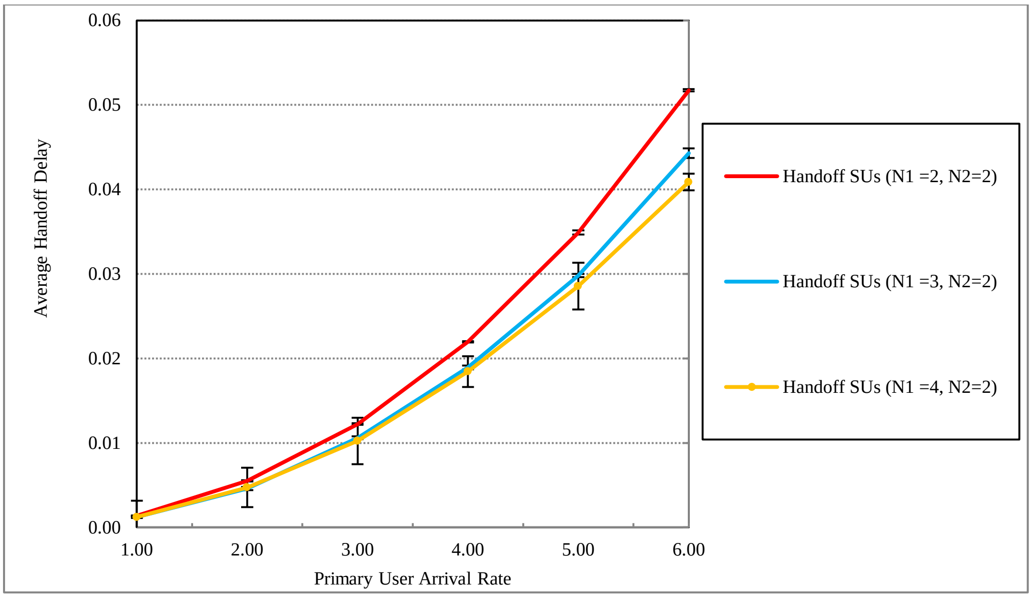

Figure 6.20: Average handoff delay.

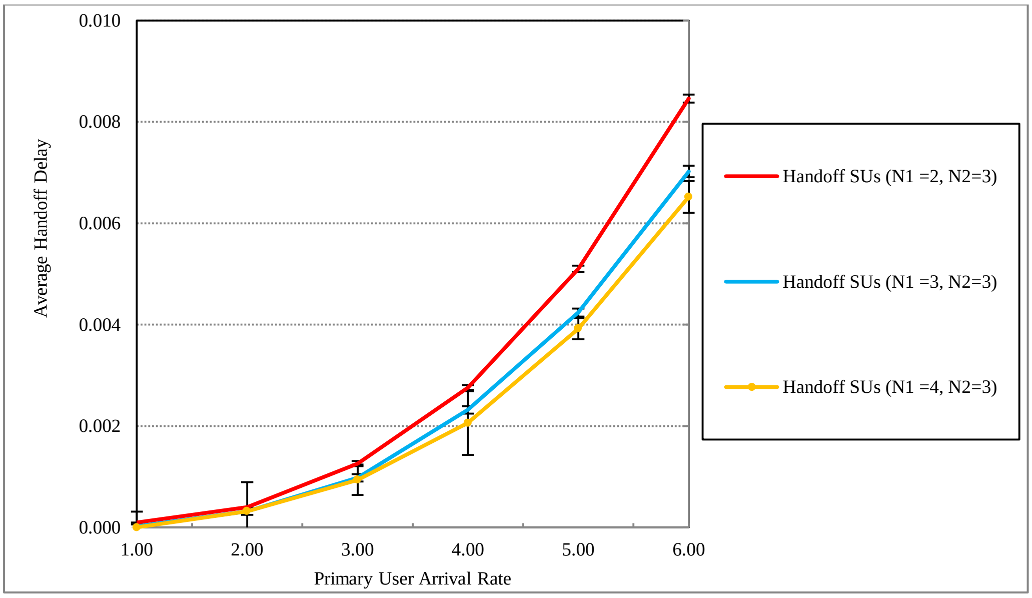

Figure 6.21: Average handoff delay.

6.6 Summary

In this chapter, a novel spectrum handoff scheme has been developed, which operates in a heterogeneous spectrum environment. The scheme employs contention-free and contention-based (random) access methods, which are modelled using the PRP M/M/C queue and the M/M/C retrial queue, respectively. The developed scheme switches all handoff users to the unlicensed spectrum channels in order to resume interrupted transmissions. The main advantage of the unlicensed spectrum band is that all users have the same priority and no more interruptions are allowed, which limits and decreases the handoff delay. Additionally, the interrupted secondary users have been given a higher priority to utilise the available unlicensed channel in order to reduce the average handoff delay and improve their QoS. Achieved results from extensive simulation experiments have been validated whenever it is possible, with the results being obtained from the analytical models. Moreover, a comparison study has been made to compare the performance of the scheme under different scenarios and various companions of simulation parameters, such as retrial rates and the number of licensed and unlicensed channels. Revealed results have been shown that the proposed scheme improves the average handoff delay and increases the performance of the proposed heterogeneous cognitive network.

Chapter 7.

CONCLUSIONS AND FUTURE WORK

Due to the growth and spread of wireless devices in the unlicensed spectrum bands, such as ISM, the bands become more and more crowded and therefore affect the performance of the wireless networks negatively. Thus, the cognitive spectrum access principles are required to utilise the existing spectrum bands more efficiently. The OSA technique is a step towards solving the issue of the spectrum being underutilised in today's wireless networks. Based on OSA, wireless networks can reuse the remainder of the spectrum which is not being used by the spectrum owners therefore, the efficiency of using the spectrum for wireless networks will be improved. However, several challenges need to be resolved in order to gain from the advantages of such innovate spectrum management techniques. Spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility (handoff) are examples of such challenges.

Despite its importance, spectrum handoff isn't explored as much as it should be. Handoff secondary users occupying the licensed spectrum should vacate the spectrum when required by the primary users, while initiating the spectrum handoff procedure to find an empty spectrum where they can complete the interrupted transmissions. On the other hand, spectrum handoff would make the interrupted secondary users experience additional delay. Thus, the principal contribution of this thesis is to present answers to address the aforementioned challenge.

In the next subsection, we conclude the main chapters of the thesis by giving a summary of each chapter and couple them with the achieved results, and finally providing the conclusions reached.

7.1 Conclusions

The most important contribution chapters of the thesis are summarised below:

In Chapter 3, two novel proactive-decision spectrum handoff schemes had been developed. These schemes are the switching-handoff scheme (SWH-NEW) and the random scheme (RAH-NEW) which operate on a centralised contention free environment in the licensed spectrum bands. A Preemptive Resume Priority (PRP) M/G/1 queuing theory was used to model the proposed cognitive radio network. The developed schemes provide handoff secondary users with a higher priority to utilise unused spectrum channels. The proposed schemes have been compared with the existing generated schemes, such as non-switching-handoff (NSWH), switching-handoff (SWH-OLD), random handoff (RAH-OLD), and reactive handoff (REH-OLD). Simulation results have been validated using the developed analytical models. The comparisons demonstrated that the simulation results closely match those achieved using analytical models. Moreover, achieved results have revealed that the improvement in the cumulative handoff delay in SWH-NEW and RAH-NEW is approximately 65% and 60% respectively, compared with the existing generated-corresponding schemes.

In Chapter 4, a spectrum handoff scheme was presented, which is classified under the reactive-decision approach and is denoted as REH-NEW. The proposed cognitive radio network has been modeled using the Preemptive Resume Priority (PRP) M/M/1 queuing theory, which operates on a centralised contention-free environment in the licensed spectrum bands. Handoff users choose the channel with the least number of waiting secondary users to resume the interrupted transmissions, and are provided with higher priority to utilise the idle spectrum channels. Extensive simulation experiments have been conducted to compare REH-NEW with the corresponding existing scheme (REH-OLD) under a range of sensing times. Also, it has been compared with other schemes implemented and generated in Chapter 3. Achieved results, show that the improvement in the cumulative handoff delay in some cases is approximately 34% compared with the existing corresponding scheme (REH-OLD). On the other hand, REH-NEW improves the cumulative handoff delay up to about 20% and 34%, compared with RAH-NEW and SWH-NEW, respectively. Moreover, this improvement can be much larger and can reach up to approximately 68% and 77% in the case of being compared to the existing schemes, RAH-OLD and SWH-OLD, respectively.

In Chapter 5, another reactive-decision spectrum handoff scheme was suggested and was denoted as REH-SQ. A Preemptive Resume Priority (PRP) M/M/1 queuing theory was used to model the proposed cognitive radio network, which works in a centralised, contention free environment in the licensed spectrum bands. The suggested scheme used a common queue to serve the arrived handoff users from all of the channels which encountered interruption events. Also, handoff users preferred to occupy idle channels over the other secondary users. Extensive simulation experiments have been conducted to investigate the performance of the scheme. REH-SQ has been compared with other reactive schemes, such as REH-NEW and REH-OLD, under a range of sensing times. Moreover, it has been compared with other schemes implemented and generated in the previous chapters. The achieved results have demonstrated that REH-SQ can significantly improve the total service time. For example, the improvement in the case of SWH-OLD and RAH-OLD can reach up to approximately 70% and 75%, respectively.

In Chapter 6, a novel spectrum handoff scheme was developed, which operated in a heterogeneous spectrum environment. Since access techniques to the licensed spectrum bands and unlicensed spectrum are different and rely on contention-free and contention-based (random) access approaches, respectively, a mixed queuing network model has been adopted in order to model the proposed cognitive radio network. A Preemptive Resume Priority (PRP) M/M/C queuing theory has been used to model the part of the network which operates on the licensed spectrum band and the M/M/C retrial queuing theory has been used to model the part of the network which operates on the unlicensed spectrum bands. The developed scheme switches all handoff users to the unlicensed spectrum channels in order to resume interrupted transmissions. The main advantage of the unlicensed spectrum bands is that all users have the same priority and no more interruptions are allowed in this band, limiting and decreasing the handoff delay. Moreover, the scheme provides the handoff users with a higher priority to utilise the idle spectrum channels by assigning them with higher retrial rates compared with those of other new users. In other words, this has been achieved by assigning handoff SUs with a shorter average backoff time than that of the new SUs. In general, two scenarios have been assumed: a ‘single-licensed and single-unlicensed channels' (N_1=1, N_2=1) scenario and a 'multi-licensed and multi-unlicensed channels' (N_1=n1, N_2=n2) scenario. The achieved results from extensive simulation experiments have been validated whenever it is possible with the results obtained from the analytical models. Moreover, a comprehensive comparison study has been made to compare the performance of the scheme under different scenarios and various combinations of simulation parameters, such as the retrial rates and the number of licensed and unlicensed channels. Also, the scheme has been compared with the situation where both the original and handoff SUs have the same access priority, and when handoff users have a higher priority to access the unlicensed band. The revealed results have shown that the proposed scheme improves the average handoff delay and increases the performance of the proposed heterogeneous cognitive radio network.

7.2 Future Work

In future, the proposed schemes in this thesis can be extended to practise more general traffic models of the arrival and service processes, rather than just the Poisson and the Geometric distributions, respectively, in order to capture the features of today's traffic better. Also, in general, finite queuing models attract less attention than the infinite models and so, as a result, more research can be done in this area. Based-on this hypothesis, other performance measures besides delay metrics can be investigated such as, the blocking probability and the throughput.

Furthermore, other service completion policy rather than the resumption policies which is used in this these can be adopted. In this thesis, the interrupted secondary users resume their unfinished transmission on the available channels, according to the proposed spectrum handoff scheme. However, in other situations, the interrupted secondary users may decide to drop the interrupted transmission and will not be transmitted any more. Other possibility, may they decide to retransmit the whole connection rather than resuming the interrupted transmission. In this case, the preemptive repeat priority queuing network should be used to model the cognitive radio network, and is valuable to examine the delay performance resulting from applying this discipline.

Moreover, the proposed spectrum handoff schemes assumes identical traffic parameters for the licensed wireless channels such as arrival rates and service times for the primary and secondary users. Alternatively, the effect of non-identical traffic parameters on the proposed spectrum handoff schemes can be explored.

Based on proposed queuing network models, other performance measures such as channel utilisation under traditional wireless networks and cognitive radio networks can be studied and compared to show the benefit from adopting dynamic spectrum access (DSA) and opportunistic spectrum access approaches (OSA) in the wireless networks.

In some application, multiple secondary users may provide with different priorities to utilise available spectrum bands. In such cases, the effect of spectrum handoff can be modelled using both preemptive queuing disciplines (between the PUs and the SUs) and non-preemptive (HOL) queuing disciplines (between the SUs themselves).

In a real wireless networks, spectrum sensing is not perfect as assumed in this thesis, thus it is worthy to incorporate the innovative prioritised principle proposed in this thesis with the effect of false alarm probability and detection probability in the proposed spectrum handoff models.

REFERENCES

[1] A. Wyglinski and M. Nekovee, "Cognitive radio communications and networks," *IEEE Communications Magazine,* vol. 46, pp. 30-31, 2008.

[2] E. FCC, "Docket No 02-135," *Spectrum Policy Task Force Report*, 2002.

[3] S. M. Mishra, D. Cabric, C. Chang, D. Willkomm, B. Van Schewick, S. Wolisz, and B. Brodersen, "A real time cognitive radio testbed for physical and link layer experiments," 2005, pp. 562-567.

[4] S. Huang, X. Liu, and Z. Ding, "Optimal transmission strategies for dynamic spectrum access in cognitive radio networks," *Mobile Computing, IEEE Transactions on,* vol. 8, pp. 1636-1648, 2009.

[5] I. F. Akyildiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, "A survey on spectrum management in cognitive radio networks," *Communications Magazine, IEEE,* vol. 46, pp. 40-48, 2008.

[6] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *Selected Areas in Communications, IEEE Journal on,* vol. 23, pp. 201-220, 2005.

[7] B. A. Fette, *Cognitive radio technology*: Academic Press, 2009.

[8] L. Kleinrock, "Queueing Systems, Vol. II: Computer Applications," ed: Wiley, New York, 1976.

[9] L. Kleinrock, "Queueing systems, vol. I: Theory," ed: Wiley, New York, 1975.

[10] Available: http://www.cse.fau.edu/~bob/publications/encyclopedia.pdf

[11] J. Walraevens, D. Fiems, and H. Bruneel, "Performance analysis of priority queueing systems in discrete time," in *Network performance engineering*, ed: Springer, 2011, pp. 203-232.

[12] M. Tanner, *Practical Queueing Analysis*: McGraw-Hill, 1995.

[13] D. Efrosinin and A. Winkler, "Queueing system with a constant retrial rate, non-reliable server and threshold-based recovery," *European journal of operational research*, vol. 210, pp. 594-605, 2011.

[14] B. Choi and Y. Chang, "Single server retrial queues with priority calls," *Mathematical and Computer modelling,* vol. 30, pp. 7-32, 1999.

[15] J. Roszik, "Retrial Queues and their Application in Performance Modelling of Communication Networks," 2010.

[16] J. Amador and J. Artalejo, "The M/G/1 retrial queue: New descriptors of the customer’s behavior," *Journal of computational and Applied Mathematics,* vol. 223, pp. 15-26, 2009.

[17] A. Kalmychkov and G. Medvedev, "Probability characteristics of Markov local-area networks with random-access protocols," *Automatic Control and Computer Sciences,* vol. 24, pp. 38-46, 1991.

[18] W. D. Kelton and A. M. Law, *Simulation modeling and analysis*: McGraw Hill Boston, MA, 2000.

[19] B. Cao, Q. Zhang, J. W. Mark, L. X. Cai, and H. V. Poor, "Toward efficient radio spectrum utilization: User cooperation in cognitive radio networking," *Network, IEEE,* vol. 26, pp. 46-52, 2012.

[20] P. Yadav and P. P. Bhattacharya, "An Effective Dynamic Spectrum Access Method for Use in Cognitive Radio," 2013.

[21] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *Signal Processing Magazine, IEEE,* vol. 24, pp. 79-89, 2007.

[22] Q. Zhao and A. Swami, "A survey of dynamic spectrum access: signal processing and networking perspectives," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on,* 2007, pp. IV-1349-IV-1352.

[23] Y. C. Liang, H. H. Chen, J. Mitola, P. Mahonen, R. Kohno, J. Reed, and L. Milstein, "Cognitive radio: theory and application," *IEEE Journal on Selected Areas in Communication*, vol. 26, pp. 1-4, 2008.

[24] C. X. Wang, H. H. Chen, X. Hong, and M. Guizani, "Cognitive radio network management," *Vehicular Technology Magazine, IEEE,* vol. 3, pp. 28-35, 2008.

[25] J. Mitola III and G. Q. Maguire Jr, "Cognitive radio: making software radios more personal," *Personal Communications, IEEE,* vol. 6, pp. 13-18, 1999.

[26] J. Mitola, "Cognitive radio: An integrated agent architecture for software defined radio," *Doctor of Technology, Royal Inst. Technol.(KTH), Stockholm, Sweden,* pp. 271-350, 2000.

[27] A. Jamalipour, "Cognitive heterogeneous mobile networks [Message from the Editor-in-Chief]," *Wireless Communications, IEEE,* vol. 15, pp. 2-3, 2008.

[28] L. BERLEMANN and S. MANGOLD, "Cognitive radio and dynamic spectrum access," 2009.

[29] *ITU page on definitions of ism bands. [online]*. Available: http: //www.itu.int/ITU-R/terrestrial/faq/index.html

[30] FCC, "Reports of commissions rules to permit unlicensed national information infrastructure (u-nii) devices in the 5 ghz band," Tech. Rep.November 2003.

[31] O. B. Abdulghfoor, M. Ismail, and R. Nordin, "Application of game theory to underlay ad-hoc cognitive radio networks: An overview," in *Space Science and Communication (IconSpace), 2013 IEEE International Conference on*, 2013, pp. 296-301.

[32] J. Jia, Q. Zhang, and X. Shen, "HC-MAC: A hardware-constrained cognitive MAC for efficient spectrum management," *Selected Areas in Communications, IEEE Journal on*, vol. 26, pp. 106-117, 2008.

[33] C. Cordeiro, K. Challapali, D. Birru, and N. Sai Shankar, "IEEE 802.22: the first worldwide wireless standard based on cognitive radios," 2005, pp. 328-337.

[34] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, "A bandwidth sharing approach to improve licensed spectrum utilization," 2005, pp. 279-288.

[35] C. W. Wang, L. C. Wang, and F. Adachi, "Performance Gains for Spectrum Utilization in Cognitive Radio Networks with Spectrum Handoff," 2009.

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

[37] R. W. Thomas, L. A. DaSilva, and A. B. MacKenzie, "Cognitive networks," 2005, pp. 352-360.

[38] L. C. Wang, C. W. Wang, Y. C. Lu, and C. M. Liu, "A concurrent transmission MAC protocol for enhancing throughout and avoiding spectrum sensing in cognitive radio," 2007, pp. 121-126.

[39] Available: http://www.cmpe.boun.edu.tr/courses/cmpe591/spring2010/Ch1-Introduction.pdf

[40] O. Ileri, D. Samardzija, and N. B. Mandayam, "Demand responsive pricing and competitive spectrum allocation via a spectrum server," in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, 2005, pp. 194-202.

[41] M. M. Buddhikot, P. Kolodzy, S. Miller, K. Ryan, and J. Evans, "DIMSUMnet: new directions in wireless networking using coordinated dynamic spectrum," in *World of Wireless Mobile and Multimedia Networks, 2005. WoWMoM 2005. Sixth IEEE International Symposium on a*, 2005, pp. 78-85.

[42] G. A. Shah and O. B. Akan, "Performance Analysis of CSMA-based Opportunistic Medium Access Protocol in Cognitive Radio Sensor Networks," *Ad Hoc Networks,* 2013.

[43] J. Tian, H. Tianfield, and G. Bi, "Non spectrum handover link maintenance performance in cognitive radio," in *Systems, Man and Cybernetics, 2009. SMC 2009. IEEE International Conference on*, 2009, pp. 2713-2717.

[44] I. F. Akyildiz, W. Y. Lee, and K. R. Chowdhury, "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Networks,* vol. 7, pp. 810-836, 2009.

[45] A. Ghasemi and E. S. Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *Communications Magazine, IEEE,* vol. 46, pp. 32-39, 2008.

[46] M. A. Kalil, A. Puschmann, and A. Mitschele-Thiel, "SWITCH: A Multichannel MAC Protocol for Cognitive Radio Ad Hoc Networks," in *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, 2012, pp. 1-5.

[47] W. Li-Chun and W. Chung-Wei, "Spectrum Handoff for Cognitive Radio Networks: Reactive-Sensing or Proactive-Sensins?," in *Performance, Computing and Communications Conference, 2008. IPCCC 2008. IEEE International*, 2008, pp. 343-348.

[48] L. C. Wang, C. W. Wang, and K. T. Feng, "A queueing-theoretical framework for QoS-enhanced spectrum management in cognitive radio networks," *Wireless Communications, IEEE,* vol. 18, pp. 18-26, 2011.

[49] Q. Shi, D. Taubenheim, S. Kyperountas, P. Gorday, and N. Correal, "Link maintenance protocol for cognitive radio system with OFDM PHY," 2007, pp. 440-443.

[50] S. Srinivasa and S. A. Jafar, "The throughput potential of cognitive radio: a theoretical perspective," 2006, pp. 221-225.

[51] H. Su and X. Zhang, "Channel-hopping based single transceiver MAC for cognitive radio networks," 2008, pp. 197-202.

[52] W. Chung-Wang and W. Li-Chung, "Modeling and Analysis for Proactive-Decision Spectrum Handoff in Cognitive Radio Networks," in *Communications, 2009. ICC '09. IEEE International Conference on*, 2009, pp. 1-6.

[53] S. Zahed, I. Awan, and M. Geyong, "Prioritized Proactive Scheme for Spectrum Handoff Decision in Cognitive Radio Networks," presented at the 7th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA 2012), University of Victoria, Victoria, Canada, 2012.

[54] J. Gambini, O. Simeone, U. Spagnolini, Y. Bar-Ness, and Y. Kim, "Cognitive radio with secondary packet-by-packet vertical handover," in *Communications, 2008. ICC'08. IEEE International Conference on*, 2008, pp. 1050-1054.

[55] H. Kim and K. G. Shin, "Efficient discovery of spectrum opportunities with MAC-layer sensing in cognitive radio networks," *Mobile Computing, IEEE Transactions on,* vol. 7, pp. 533-545, 2008.

[56] S. Zahed, I. Awan, and A. Cullen, "Analytical modeling for spectrum handoff decision in cognitive radio networks," *Simulation Modelling Practice and Theory,* vol. 38, pp. 98-114, 2013.

[57] W. Chung-Wei, W. Li-Chun, and F. Adachi, "Modeling and Analysis for Reactive-Decision Spectrum Handoff in Cognitive Radio Networks," in *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE,* 2010, pp. 1-6.

[58] S. Zahed, I. Awan, and M. Geyong, "Performance Evaluation of Secondary Users in Cognitive Radio Networks," in *27 th UKPEW,* Bradford, UK, 2011, pp. 340-343.

[59] C.-W. Wang and L.-C. Wang, "Analysis of Reactive Spectrum Handoff in cognitive radio networks," *Selected Areas in Communications, IEEE Journal on,* vol. 30, pp. 2016-2028, 2012.

[60] W. Hu, D. Willkomm, M. Abusubaih, J. Gross, G. Vlantis, M. Gerla, and A. Wolisz, "Cognitive radios for dynamic spectrum access-dynamic frequency hopping communities for efficient ieee 802.22 operation," *Communications Magazine, IEEE,* vol. 45, pp. 80-87, 2007.

[61] D. Willkomm, J. Gross, and A. Wolisz, "Reliable link maintenance in cognitive radio systems," 2005, pp. 371-378.

[62] J. Tian and G. Bi, "A new link maintenance and compensation model for cognitive uwb radio systems," 2006, pp. 254-257.

[63] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *Selected Areas in Communications, IEEE Journal on,* vol. 25, pp. 589-600, 2007.

[64] L.-C. Wang, C.-W. Wang, and C.-J. Chang, "Modeling and analysis for spectrum handoffs in cognitive radio networks," *Mobile Computing, IEEE Transactions on,* vol. 11, pp. 1499-1513, 2012.

[65] S. U. Yoon and E. Ekici, "Voluntary spectrum handoff: a novel approach to spectrum management in CRNs," 2010, pp. 1-5.

[66] D. Aman, S. Mahfooz, and M. Waheed Ur Rehman, "A Handoff using Guard Channels Scheme (HGCS) for Cognitive Radio Networks," *Global Journal of Computer Science and Technology,* vol. 11, 2011.

[67] Y. Zhang, "Spectrum handoff in cognitive radio networks: Opportunistic and negotiated situations," in *Communications, 2009. ICC'09. IEEE International Conference on*, 2009, pp. 1-6.

[68] Y. Song and J. Xie, "Proactive spectrum handoff in cognitive radio ad hoc networks based on common hopping coordination," in *INFOCOM IEEE Conference on Computer Communications Workshops, 2010,* 2010, pp. 1-2.

[69] S. Zheng, X. Yang, S. Chen, and C. Lou, "Target channel sequence selection scheme for proactive-decision spectrum handoff," *Communications Letters, IEEE,* vol. 15, pp. 1332-1334, 2011.

[70] S. M. Potdar and K. P. Patil, "Efficient spectrum handoff in CR network based on mobility, QoS and priority using fuzzy logic and neural network," in *Contemporary Computing (IC3), 2013 Sixth International Conference on,* 2013, pp. 53-58.

[71] N. A. Lala, M. Uddin, and N. Sheikh, "Novel Spectrum Handoff in Cognitive Radio Networks Using Fuzzy Logic," *International Journal of Information Technology and Computer Science (IJITCS),* vol. 5, p. 103, 2013.

[72] J. Duan and Y. Li, "An optimal spectrum handoff scheme for cognitive radio mobile ad hoc networks," *Advances in Electrical and Computer Engineering,* vol. 11, pp. 11-16, 2011.

[73] M. A. Kalil, H. Al-Mahdi, and A. Mitschele-Thiel, "Analysis of opportunistic spectrum access in cognitive ad hoc networks," in *Analytical and Stochastic Modeling Techniques and Applications*, ed: Springer, 2009, pp. 16-28.

[74] H. Al-Mahdi, M. Kalil, F. Liers, and A. Mitschele-Thiel, "Increasing spectrum capacity for ad hoc networks using cognitive radios: an analytical model," *Communications Letters, IEEE,* vol. 13, pp. 676-678, 2009.

[75] M. A. Kalil, H. Al-Mahdi, and A. Mitschele-Thiel, "Spectrum handoff reduction for cognitive radio ad hoc networks," in *Wireless Communication Systems (ISWCS), 2010 7th International Symposium on*, 2010, pp. 1036-1040.

[76] D. B. Zhu and B. D. Choi, "Performance analysis of CSMA in an unslotted cognitive radio network with licensed channels and unlicensed channels," *EURASIP Journal on Wireless Communications and Networking,* vol. 2012, pp. 1-7, 2012.

[77] Y. Xing, R. Chandramouli, and S. Mangold, "Dynamic spectrum access in open spectrum wireless networks," *Selected Areas in Communications, IEEE Journal on,* vol. 24, pp. 626-637, 2006.

[78] L.-C. Wang, A. Chen, and D. S. Wei, "A cognitive MAC protocol for QoS provisioning in overlaying ad hoc networks," in *Consumer Communications and Networking Conference, 2007. CCNC 2007. 4th IEEE,* 2007, pp. 1139-1143.

[79] C. T. Chou and K. G. Shin, "Analysis of combined adaptive bandwidth allocation and admission control in wireless networks," in *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE,* 2002, pp. 676-684.

[80] H. G. Ebersman and O. K. Tonguz, "Handoff ordering using signal prediction priority queuing in personal communication systems," *Vehicular Technology, IEEE Transactions on,* vol. 48, pp. 20-35, 1999.

[81] M. Naghshineh and M. Schwartz, "Distributed call admission control in mobile/wireless networks," *Selected Areas in Communications, IEEE Journal on,* vol. 14, pp. 711-717, 1996.

[82] D. Hong and S. S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures," *Vehicular Technology, IEEE Transactions on,* vol. 35, pp. 77-92, 1986.

[83] A. E. Xhafa and O. K. Tonguz, "Dynamic priority queueing of handover calls in wireless networks: An analytical framework," *Selected Areas in Communications, IEEE Journal on,* vol. 22, pp. 904-916, 2004.

[84] L. Wang, C. Wang, and C. Chang, "Optimal Target Channel Sequence Design for Multiple Spectrum Handoffs in Cognitive Radio Networks," 2012.

[85] P. Zhu, J. Li, and X. Wang, "Scheduling Model for Cognitive Radio," in *2008 3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2008),* 2008, pp. 1-6.

[86] P. Zhu, J. Li, K. Han, and X. Wang, "A new channel parameter for cognitive radio," in *2007 2nd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom 2007),* 2007, pp. 482-486.

[87] L.-C. Wang, C.-W. Wang, and F. Adachi, "Load-balancing spectrum decision for cognitive radio networks," *Selected Areas in Communications, IEEE Journal on*, vol. 29, pp. 757-769, 2011.

[88] L.-C. Wang and C.-W. Wang, "Spectrum management techniques with QoS provisioning in cognitive radio networks," in *Wireless Pervasive Computing (ISWPC), 2010 5th IEEE International Symposium on*, 2010, pp. 116-121.

[89] B. Wang, Z. Ji, K. R. Liu, and T. C. Clancy, "Primary-prioritized Markov approach for dynamic spectrum allocation," *Wireless Communications, IEEE Transactions on*, vol. 8, pp. 1854-1865, 2009.

[90] M. Huang, R. Yu, and Y. Zhang, "Call admission control with soft-QoS based spectrum handoff in cognitive radio networks," in *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, 2009, pp. 1067-1072.

[91] A. Banaei and C. N. Georghiades, "Throughput analysis of a randomized sensing scheme in cell-based ad-hoc cognitive networks," in *Communications, 2009. ICC'09. IEEE International Conference on*, 2009, pp. 1-6.

[92] T. Jiang, D. Grace, and Y. Liu, "Cognitive radio spectrum sharing schemes with reduced spectrum sensing requirements," in *Cognitive Radio and Software Defined Radios: Technologies and Techniques, 2008 IET Seminar on*, 2008, pp. 1-5.

[93] I. A. M. Balapuwaduge, "Performance evaluation of channel aggregation strategies in cognitive radio networks with queues," 2012.

[94] G. Salami, O. Durowoju, A. Attar, O. Holland, R. Tafazolli, and H. Aghvami, "A comparison between the centralized and distributed approaches for spectrum management," *Communications Surveys & Tutorials, IEEE*, vol. 13, pp. 274-290, 2011.

[95] A. Lertsinsrubtavee, N. Malouch, and S. Fdida, "Spectrum handoff strategy using cumulative probability in cognitive radio networks," in *Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 2011 3rd International Congress on*, 2011, pp. 1-7.

[96] *Facebook At 13 Million Queries Per Second Recommends: Minimize Request Variance*. Available: http://highscalability.com/blog/2010/11/4/facebook-at-13-million-queries-per-second-recommends-minimiz.html

[97] M. E. Ahmed, J. S. Kim, R. Mao, J. B. Song, and H. Li, "Distributed Channel Allocation Using Kernel Density Estimation in Cognitive Radio Networks," *ETRI Journal*, vol. 34, 2012.

[98] F. Commission, "Spectrum policy task force report, FCC 02-155," ed: Nov, 2002.

[99] V. K. Tumuluru, P. Wang, D. Niyato, and W. Song, "Performance analysis of cognitive radio spectrum access with prioritized traffic," *Vehicular Technology, IEEE Transactions on*, vol. 61, pp. 1895-1906, 2012.

[100] P. Pawelczak, R. Venkatesha Prasad, L. Xia, and I. G. Niemegeers, "Cognitive radio emergency networks-requirements and design," in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, 2005, pp. 601-606.

[101] D. Gross, *Fundamentals of queueing theory*: John Wiley & Sons, 2008.

[102] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi, *Queueing networks and Markov chains: modeling and performance evaluation with computer science applications*: Wiley-Interscience, 2006.