UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

COGNITIVE RADIO PERFORMANCE OPTIMISATION THROUGH SPECTRUM AVAILABILITY PREDICTION

by

Simon Daniel Barnes

Submitted in partial fulfilment of the requirements for the degree

Master of Engineering (Electronic Engineering)

in the

Department of Electrical, Electronic and Computer Engineering
Faculty of Engineering, Built Environment and Information Technology
UNIVERSITY OF PRETORIA

March 2012

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

SUMMARY

COGNITIVE RADIO PERFORMANCE OPTIMISATION THROUGH SPECTRUM AVAILABILITY PREDICTION

by

Simon Daniel Barnes

Supervisor(s):	Prof. B.T.J. Maharaj
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Master of Engineering (Electronic Engineering)
Keywords:	Channel switching, cognitive radio, occupancy modelling, opportunistic spectrum allocation, prediction accuracy, secondary user performance, spectrum measurements, spectrum sensing, traffic density, training algorithm complexity

The federal communications commission (FCC) has predicted that, under the current regulatory environment, a spectrum shortage may be faced in the near future. This impending spectrum shortage is in part due to a rapidly increasing demand for wireless services and in part due to inefficient usage of currently licensed bands. A new paradigm pertaining to wireless spectrum allocation, known as cognitive radio (CR), has been proposed as a potential solution to this problem.

This dissertation seeks to contribute to research in the field of CR through an investigation into the effect that a primary user (PU) channel occupancy model will have on the performance of a secondary user (SU) in a CR network. The model assumes that PU channel occupancy can be described as a binary process and a two state Hidden Markov Model (HMM) was thus chosen for this investigation. Traditional algorithms for training the model were compared with certain evolutionary-based training algorithms in terms of their resulting prediction accuracy and computational complexity. The performance of this model is important

since it provides SUs with a basis for channel switching and future channel allocations.

A CR simulation platform was developed and the results gained illustrated the effect that the model had on channel switching and the subsequently achievable performance of a SU operating within a CR network. Performance with regard to achievable SU data throughput, PU disruption rate and SU power consumption, were examined for both theoretical test data as well as data obtained from real world spectrum measurements (taken in Pretoria, South Africa). The results show that a trade-off exists between the achievable SU throughput and the average PU disruption rate. Significant SU performance improvements were observed when prediction modelling was employed and it was found that the performance and complexity of the model were influenced by the algorithm employed to train it. SU performance was also affected by the length of the quick sensing interval employed. Results obtained from measured occupancy data were comparable with those obtained from theoretical occupancy data with an average percentage similarity score of 96% for prediction accuracy (using the Viterbi training algorithm), 90% for SU throughput, 83% for SU power consumption and 71% for PU disruption rate.

OPSOMMING

KOGNITIEWE RADIOPRESTASIEOPTIMERING DEUR SPEKTRUMBESKIKBAARHEIDVOORSPELLING

deur

Simon Daniel Barnes

Studieleier(s):	Prof. B.T.J. Maharaj
Departement:	Elektriese, Elektroniese en Rekenaar-Ingenieurswese
Universiteit:	Universiteit van Pretoria
Graad:	Magister in Ingenieurswese (Elektroniese Ingenieurswese)
Sleutelwoorde:	Besetting modellering, kanaal skakel, kognitiewe radio, opleidingalgoritmekompleksiteit, opportunistiese spektrum toekenning, sekondêre gebruiker prestasie, spektrummetings, spektrumopgelet, verkeerdigtheid, voorspellingakkuraatheid

Die Federale Kommunikasie Kommissie (FKK) het voorspel dat ’n spektrum-tekort in die nabye toekoms, onder die huidige regulatoriese omgewing, ondervind kan word. Hierdie dreigende spektrum-tekort is gedeeltelik as gevolg van die vinnig groei van aanvraag na draadlose dienste en van ondoeltreffende gebruik van die huidige gelisensieerde spektrum. ’n Nuwe paradigma met betrekking tot draadlose spektrumtoekenning, wat bekend staan as kognitiewe radio (KR), is voorgestel as ’n moontlike oplossing vir hierdie probleem.

Die doel van hierdie verhandeling is om navorsing op die gebied van KR te doen. Die navorsing sal bydra deur ’n ondersoek in te stel op die effek van ’n primêre gebruiker (PG) kanaalbesettingsmodel op die prestasie van ’n sekondêre gebruiker (SG) in ’n KR-netwerk. Die model is gebaseer op die aanname dat die PG-kanaalbesetting as ’n binêre proses beskryf kan word en ’n twee-toestand Verborge Markov-model (VMM) is dus vir hierdie ondersoek gekies. Tradisionele algoritmes vir die afrigting van die model is met sekere evolusionêre-gebaseerde opleidingalgoritmes in terme van hulle voorspelling van akkuraatheid en bere-

keningkompleksiteit vergelyk. Die prestasie van die model is belangrik omdat dit SGs ’n basis bied vir kanaalskakeling en toekomstige kanaaltoekennings.

’n Basis vir KR-simulasies is ontwikkel en die effek wat dié model op die kanaalskakeling het sowel as die bereikbare prestasie van ’n SG binne ’n KR-netwerk is deur die resultate geïllustreer. Prestasie met betrekking tot die haalbare deurvoer van SG data, PG-ontwrigtingkoers en die SG-kragverbruik, is vir beide teoretiese toetsdata asook die data wat verkry is deur werklike spektrummetings (geneem in Pretoria, Suid-Afrika) word geïllustreer. Die resultate wys dat daar ’n duidelike verband bestaan tussen die haalbare SG-deurset en die gemiddelde PU-ontwrigtingkoers. ’n Beduidende SG-prestasieverbetering is waargeneem wanneer voorspellingsmodellering gebruik word en dit is gevind dat die prestasie en kompleksiteit van die model beïnvloed word deur die algoritme wat gebruik word om dit op te lei. SG-prestasie word ook deur die lengte van die vinnige afmeting in tyd beïnvloed. Uitslae wat verkry is op grond van gemete besettingsdata is vergelykbaar met dié wat met teoretiese besettingsdata verkry is, met ’n gemiddelde gelykvormigheidstelling van 95% vir die voorspellingakkuraatheid, 90% vir die SG-deurset, 83% vir die SG-kragverbruik en 71% vir die PG-ontwrigtingskoers.

I dedicate this work to the Alpha and the Omega, the First and the Last, the Beginning and the End.

ACKNOWLEDGEMENTS

I would like to thank the following people and organisations for their assistance in making this work possible:

- My parents for their selfless support, encouragement and wisdom.
- My supervisor, Prof B.T. Maharaj, for his provision and guidance.
- The Sentech Chair in Broadband Wireless Multimedia Communication (BWMC) at the University of Pretoria and the National Research Foundation (NRF) for their financial support.
- My fellow BWMC students, K. Dhuness, M.J. Prinsloo and R.R. Thomas, for their insightful suggestions and technical support.
- The Computing Clusters of the University of Pretoria's Department of Electrical, Electronic and Computer Engineering, maintained by Mr H. Grobler.
- All those who stuck by and motivated me throughout this research project.

LIST OF ABBREVIATIONS

ADC	Analogue-to-Digital Converter
AWGN	Additive White Gaussian Noise
ASIC	Application-Specific Integrated Circuits
BWA	Baum-Welch Algorithm
CBP	Coexistence Beacon Protocol
CDMA	Code-division Multiple Access
CSD	Cyclic Spectral Density Function
CR	Cognitive Radio
DAC	Digital-to-Analogue Converter
DSA	Dynamic Spectrum Access
DSP	Digital Signal Processor
DSF	Downstream Subframe
DTT	Digital Terrestrial Television
ED	Energy detection
EM	Expectation-Maximisation
EGC	Equal Gain Combining
FAP	False Alarm Probability
FCC	Federal Communications Commission
FEC	Forward Error Correction
FLOPS	Floating-Point Operations Per Second
FPGA	Field-Programmable Gate Arrays
GA	Genetic Algorithm
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HMM	Hidden Markov Model
ICASA	Independent Communications Authority of South Africa
IEEE	Institute of Electrical and Electronics Engineers
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union Radiocommunication Sector

LNA	Low Noise Amplifier
MA	Memetic Algorithm
MAC	Media Access Control
MRC	Maximum Ratio Combining
MC	Markov Chain
MIMO	Multiple Input Multiple Output
MP	Miss-detection Probability
NLMS	Normalised Least Mean Square
OCED	Optimal Combining Energy Detection
OSA	Opportunistic Spectrum Access
PBSW	Proactive Busy Switching
PC	Personal Computer
PSSW	Proactive Short Switching
PSW	Proactive Switching
PSO	Particle Swarm Optimisation
PU	Primary User
QoS	Quality of Service
QPSK	Quadrature phase-shift keying
RLS	Recursive Least Squares
RSW	Reactive Switching
SA	Spectrum Analyser
SCH	Superframe Control Header
SNR	Signal to Noise Ratio
SDA	Software Defined Antenna
SDR	Software Defined Radio
SS	Spectrum Sensing
SU	Secondary User
UCS	Urgent Coexistence Situation Notification
USF	Upstream Subframe
UWB	Ultra Wide Band
VA	Viterbi Algorithm
WRAN	Wireless Regional Area Network

LIST OF TABLES

LIST OF FIGURES

TABLE OF CONTENTS

CHAPTER 1

RESEARCH OVERVIEW

1.1 INTRODUCTION

The past couple of decades have seen an ever-increasing number people relying on the internet for both work and entertainment. This global trend, together with a continual growth in demand for mobile voice and data services, has revolutionised the way in which many people go about their daily lives, and continues to present communications engineers with a number of fresh challenges [1]. The information and communications industry has become a prominent member of the global economy, even reaching and significantly impacting upon many of the smaller and more remote regions around the globe. Meeting the demands of providing greater user capacity while still maintaining a healthy level of service quality, is one of the challenges that wireless network planners are continually faced with. The problem of meeting capacity requirements is further complicated by another, yet related, challenge that may become increasingly prominent in future wireless networks: the issue surrounding the availability of useful radio spectrum.

The federal communications commission (FCC) of the United States of America has predicted a rapidly increasing demand for mobile data services, illustrated in Figure 1.1, that may even exceed traditional network capacity in the near future [2]. Since regulators have already allocated most of the useful radio spectrum, finding vacant bands to support further growth for wireless technologies is becoming more of a challenge [3]. However, the

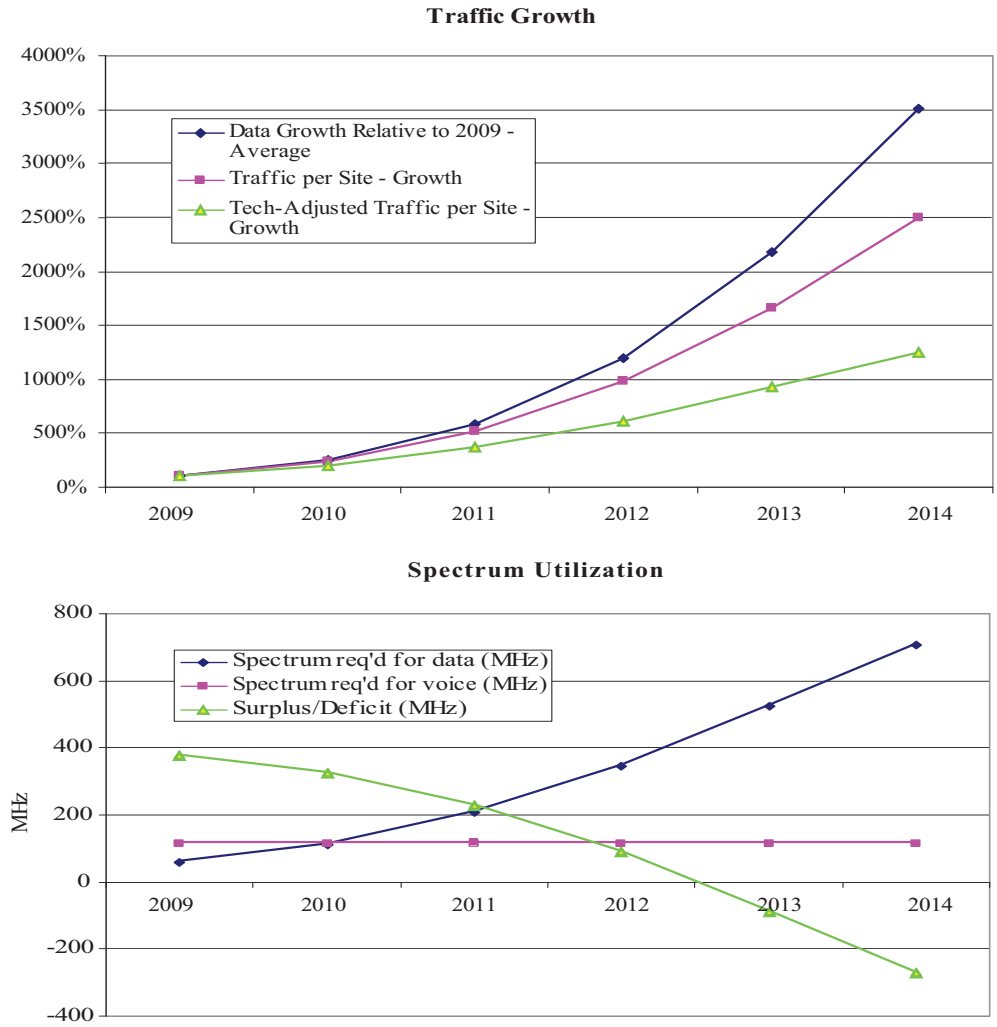

Figure 1.1: Mobile traffic growth and spectrum utilisation predictions according to the FCC, adapted from [2].

FCC has reported that although capacity appears to be running out, many licensed frequency bands are in actual fact underutilised [4]. Legacy regulation in the form of fixed spectrum assignment policy is a significant contributing factor in their observations, since the problem surrounding spectrum availability has more to do with the potential to gain spectrum access rather than simply a physical spectrum shortage [5, 6].

Radio spectrum is a finite natural resource, much like land, that needs to be efficiently managed so as to ensure capacity for the continued growth of wireless and other communication technologies. The way in which land becomes scarcer as urban areas

become more densely populated, is readily compared to the decrease in availability of useful radio spectrum with an increase in the number of wireless devices vying for usage of similar frequency bands. Given this relationship, and that operation within certain frequency bands is of greater usefulness for mobile wireless communication than others, the need to develop ways in which to make more efficient usage of the radio frequency spectrum is becoming increasingly relevant. Improving on spectral efficiency, would help to increase the overall capacity of wireless communication networks. Quantitative and qualitative benefits of such activities may be experienced both by network users and service providers alike. Service providers may potentially benefit from increased revenues, since a larger number of users would be serviceable at one time, while network users may benefit from improved network capacity and in the long run potentially lower data costs.

To this end, the concept of cognitive radio (CR) has been proposed as a means to greater spectral efficiency in future wireless networks. The concept of CR was first presented in 1999 in a paper by Joseph Mitola [7] where an outline for a more flexible approach to wireless communications, based on software defined radios (SDR) [8], was discussed. A CR is an intelligent communications system that makes decisions by being aware of and understanding its environment. An attempt at addressing the impending spectrum scarcity problem is made by CR, by allowing unlicensed secondary users (SU) to be dynamically assigned temporarily available spectrum on an opportunistic basis. This spectrum would traditionally be reserved for the exclusive use of licensed primary users (PU) who may not always make efficient usage of this spectrum. A CR seeks to exploit these temporary spectral opportunities, also referred to in the literature as 'spectrum holes' [5], and in so doing improve spectral efficiency by making use of spectrum that may otherwise have remained unoccupied [5, 6, 7].

1.2 MOTIVATION

Performance in CR networks is inherently affected by the spectrum sensing (SS) and channel selection process, which has a direct impact on parameters such as the chosen digital modulation scheme and forward error correction (FEC) coding rate. Incorrect identification of

spectrum opportunities and furthermore, subsequent suboptimal channel selections, would result in unnecessary delays, throughput degradation, disruption to PUs and an overall degradation in performance for a CR network. Information gained about PU behaviour during the SS process, is critical to the prediction of future channel availability, channel selection and the subsequent allocation of radio resources. Usage of historical PU behaviour has been shown to improve the SS process [9]. Accurate, proactive predictions of future channel usage may improve the channel allocation process, by helping to minimise the probability of incorrectly detecting spectrum holes, thus actively reducing interference delays and minimising transmission gaps due to sub-optimal channel switching [10]. Since processing and battery power are limited in mobile devices, too much complexity in the prediction process may not be desirable. A need for the development of a PU traffic prediction model that will provide SUs with the necessary basis upon which optimal channel selection may be performed, while maintaining a balance between complexity and accuracy, is thus evident and compelling.

1.3 OBJECTIVE

The primary objective of this research is to investigate an algorithm that will, in time and frequency, be able to accurately model and predict the channel occupancy behaviour of PUs within licensed frequency bands. Based on channel occupancy prediction, this work aims to contribute to the field of CR by specifically providing a platform from which SU channel allocation may accurately and intelligently be performed. Although prediction accuracy is of primary concern, convergence time and computational complexity should not be ignored.

While theoretical simulations are useful for observing initial trends, this research aims to validate its primary objective by testing the prediction algorithm and approach to channel allocation on actual PU spectrum occupancy measurements from the mobile cellular bands in South Africa. SU throughput, the potential for PUs to experience interference and the effect of the model on SU power consumption are tangible parameters that are used to test this objective.

1.4 CONTRIBUTION

The contribution that this research makes to the field of CR is through a comparison, of existing published techniques, with a new approach to the prediction of PU channel occupancy. This is achieved by investigating the effect of applying both traditional and evolutionary algorithms to the training of a hidden Markov model (HMM). This research aims to improve upon existing channel prediction methods by focusing on providing better prediction accuracy, quicker solution convergence times and lower computational complexity, as well as providing a set of results tested upon locally measured spectrum occupancy data [1].

The channel selection process employed by SUs in a CR network should benefit from this improved method through an improvement in its ability to correctly and quickly perform SU channel allocations. This may lead to a reduction in the number of unnecessary channel switches to be performed. The benefit of which may be seen in an increase in potential SU data throughput as well as a reduction in the amount of interference experienced by PUs due to SU activity. This research validates its effect on channel allocation performance by investigating the effectiveness of combining the channel availability prediction and allocation processes. Simulated results are verified using practically measured data from the commercial cellular bands in South Africa. A comparison is also made between results obtained from theoretical occupancy data and those obtained through practical measurements.

1.5 OUTLINE OF DISSERTATION

This dissertation has been organised into seven Chapters. In Chapter 2 CR is discussed and an overview of the current literature surrounding the SS and channel allocation processes is provided. The first institute of electrical and electronics engineers (IEEE) standard on CR is

[1] The spectrum occupancy data was collected from a measurement campaign that formed part of a joint post-graduate student initiative of the Sentech Broadband Wireless and Multimedia Communications research group at the University of Pretoria.

also summarised.

PU traffic modelling and prediction are discussed in Chapter 3. The HMM structure is presented and various training algorithms and their complexities are also investigated.

The process of allocating spectrum to SUs is covered in Chapter 4, where an opportunistic spectrum access (OSA) simulation platform is presented.

A spectrum measurement campaign was performed to gather input data for the simulation platform. The measurement system and campaign are described in Chapter 5.

The contribution of this work is highlighted through the presentation and discussion, in Chapter 6, of results obtained from tests run on the simulation platform. These results encompass the modelling of PU occupancy, the spectrum allocation process and the subsequently achievable performance of a SU in a CR network. The results obtained using theoretically simulated channel occupancy are compared to those obtained from data collected during the measurement campaign.

Finally, the research problem is summarised and conclusions are drawn in Chapter 7. Recommendations for possible future research relating to this work are also discussed.

1.6 PUBLICATIONS

The following peer reviewed publications have been derived from the work undertaken during this research project. These publications pertain to the topics covered in this dissertation which include: channel occupancy modelling, spectrum allocation, resulting CR network performance and the spectrum measurements campaign.

1.6.1 Conference proceedings

The following paper was presented at and published in the peer reviewed proceedings of an international conference:

1. S.D. Barnes and B.T. Maharaj, "Performance of a Hidden Markov channel occupancy model for cognitive radio," *Proceedings of IEEE AFRICON Conference*, Livingstone, Zambia, September 2011.

1.6.2 Journal publications

As part of the research activities for this degree, the following articles were submitted to peer-reviewed and ISI accredited journals:

1. S.D. Barnes and B.T. Maharaj, "Spectrum occupancy prediction and channel allocation performance for cognitive radio," *IET Communications*, submitted for publication.

2. S.D. Barnes, M.J. Prinsloo and B.T. Maharaj, "Investigation into spectrum occupancy through a South African measurement campaign," *IET Electronics Letters*, submitted for publication.

CHAPTER 2

COGNITIVE RADIO

2.1 INTRODUCTION

CR has been proposed as a method to more efficiently utilise existing spectrum and intelligently allocate radio resources in future wireless networks [7]. At present, PUs of the spectrum are exclusively allocated fixed portions of the spectrum, which may or may not be efficiently utilised. It has been noted [5] that in certain bands, spectrum may be scarce, but is also quite often underutilised or erratically utilised. Traditionally, spectrum is allocated to PUs by regulatory bodies [1] through a licensing process. This means that a PU is given exclusive access to the bands of spectrum within which it has been licensed to operate. One of the consequences of the current exclusive spectrum allocation policy, is that it often leads to the problem of artificial spectrum scarcity. Artificial spectrum scarcity occurs when the bands licensed to PUs are not fully utilised but at the same time are not available for use by other would-be users. Thus certain currently licensed regions exist, either in time, geographical location or frequency, where spectrum is actually unused or only partially used. Spectrum occupancy measurement studies [2] confirm this [4, 9, 11, 12, 13, 14, 15, 16]. These regions may be referred to as 'spectrum holes'. In its broader sense, a CR aims to address the artificial spectrum scarcity problem by exploiting these 'holes' in the spectrum.

[1] In South Africa this is the independent communications authority of South Africa (ICASA).

[2] A spectrum occupancy measurement campaign carried out in Pretoria, South Africa as part of this research work, is described in Chapter 5.

2.2 OVERVIEW OF COGNITIVE RADIO

The pursuit of achieving greater spectral efficiency is not new to wireless communication and there are a number of commonly used techniques currently employed for this purpose. Amongst others these may include: multiple access schemes (based on time, frequency, code and spatial division), spectrally efficient modulation schemes and also adaptive modulation and FEC schemes (based on local channel conditions) [1, 17]. Network planners also make use of frequency reuse schemes with adaptive power control, as well as smart and adaptive antennas [18].

These techniques focus on the way in which information is packaged before and after entering the wireless channel. Further improvements in spectral efficiency may however be achieved by the use of CR and ultra wide band (UWB) technologies. These technologies represent two opposing approaches. A CR follows an overlay approach by exploiting spectrum 'holes' through dynamic spectrum access (DSA). However, in UWB, an underlay approach is followed where, due to bandwidth increases, waveforms may be successfully transmitted at power levels low enough for SUs to coexist with PUs over the entire frequency range [19]. The difference between overlay and underlay communication is illustrated in Figure 2.1. In CR the SU (represented by the green block) avoids interference with the PU by operating in a region where no PU (denoted by the maroon triangles) is present. In the UWB case (represented by the orange region) transmission coincides with PUs, causing tolerable levels of interference at power levels that fall below the noise floor (represented by the blue region). Combining the under and overlay approaches to further improve channel capacity has also been proposed [19].

The IEEE 1900.1 standard on DSA provides the following definition for CR [20]:

1. "A type of radio in which communication systems are aware of their environment and internal state and can make decisions about their radio operating behaviour based on that information and predefined objectives."

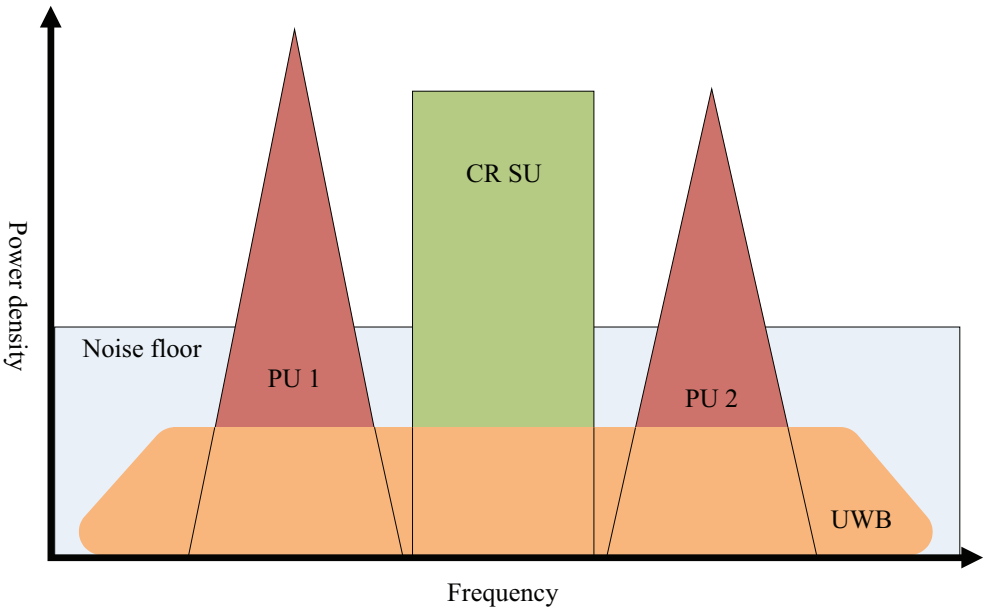

Figure 2.1: Graphical depiction of the difference between overlay and underlay communication.

2. "Cognitive radio (as defined above) that uses software-defined radio, adaptive radio, and other technologies to adjust automatically its behaviour or operations to achieve desired objectives."

Similarly, the international telecommunication union radiocommunication sector (ITU-R) provides the following definition for a CR system [21]:

> "A radio system employing technology that allows the system to obtain knowledge of its operational and geographical environment, established policies and its internal state; to dynamically and autonomously adjust its operational parameters and protocols according to its obtained knowledge in order to achieve predefined objectives; and to learn from the results obtained."

A key concept surrounding the CR is that secondary, unlicensed users are allowed to address the artificial spectrum scarcity problem by exploiting spectrum holes. This is achieved through dynamic opportunistic access to detected spectrum holes in such a way that interference to currently licensed PUs is avoided. It is important that SUs are as invisible

as possible to PUs of the spectrum, since PUs are licensed to make exclusive use of certain parts of the spectrum, for which they pay large amounts of money to regulatory bodies. Any interference from SUs would thus infringe upon the exclusive rights of the PU.

Built on the platform of SDR [22], a CR will determine and try to predict when and where spectrum-hole opportunities exist, and then intelligently select a portion of available spectrum for its own use. Radio resources may then be allocated to this SU, based on the channel conditions under which it is seeking to operate [5, 7]. This spectrum selection and resource allocation is temporary and will thus only be valid for as long as the PU does not need to make use of that portion of the spectrum. Once the PU again requires use of that portion of the spectrum, the SU must immediately vacate the band. The selection and allocation process will then have to be repeated so as to find an alternative part of the spectrum within which to operate. To maintain uninterrupted communication, a SU would thus be required to repetitively perform this process, and would also have to rely on the availability of a sufficient number of alternative spectrum holes.

A possible PU occupation pattern encompassing four adjacent frequency channels is illustrated in Figure 2.2. The horizontal blocks represent frequency channels over a period

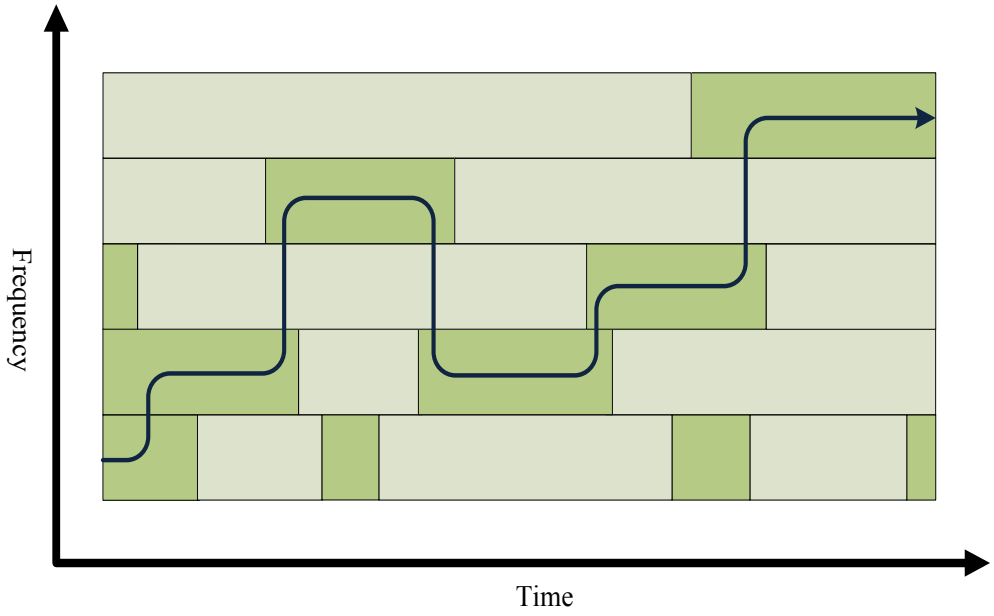

Figure 2.2: PU spectrum occupation over time and frequency.

of time, where the green blocks represent un-utilised spectrum holes and the grey blocks represent PU activity. When the PU returns to the channel currently occupied by a SU, the SU needs to jump to the next best available channel, as indicated by the arrowed line. To maintain uninterrupted communication, a SU would thus need to continuously operate within the green blocks.

2.3 PRIMARY COGNITIVE FUNCTIONS

By definition CR is a system that operates in an intellectual manner. By being acutely aware of its environment and consequently performing actions based on an understanding and processing of that knowledge, CRs possess an inherent form of intelligence. The primary functionality required of a CR, can be summarised into three broader categories [5],

1. analysis and awareness of the radio environment,

2. channel identification and prediction, as well as

3. intelligent resource allocation and spectrum management.

These three categories of operation form the basis for an intelligent feedback communication system and are repetitively performed within a CR environment. A graphical depiction of this concept is presented in Figure 2.3. The SDR [7] provides the platform upon which this process may be performed.

2.3.1 Analysis and awareness

An essential part of the CR process involves awareness about the operating environment. Sensing of the available spectrum opportunities is a major challenge in a CR network [23]. For SUs to operate in a disruptive yet unobtrusive manner, it is essential that holes in the spectrum are accurately detected and that reliable information about the interference temperature of the radio environment is obtained. A spectrum hole occurs when a specific band of frequencies, that have been allocated to a PU, become temporarily unoccupied for a specific

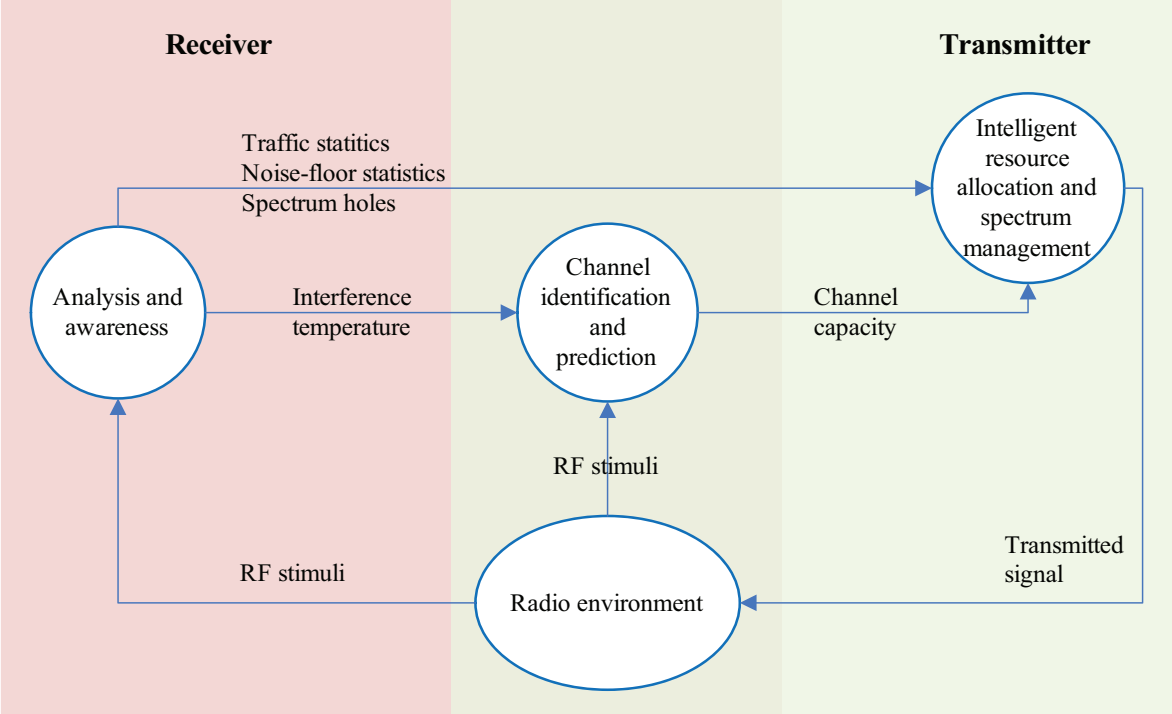

Figure 2.3: Basic description of the CR cycle, adapted from [5].

period of time at a specific geographical location [5].

2.3.2 Channel identification and prediction

Based on information gained about the environment in which a CR is to operate, channel state information needs to be estimated. Intelligent processing and interpretation of the knowledge obtained during sensing, enables the SU to decide how to further proceed under the channel conditions within which it finds itself. Estimates of potential channel capacity may be calculated and decisions made about how best the SU may be configured to operate within the current conditions.

2.3.3 Intelligent resource allocation and spectrum management

Once operating conditions have been properly determined and decisions made about how to proceed, spectrum is assigned and resources are allocated to the SU in an intelligent manner. This allocation is performed on a personal basis, taking into consideration the specific needs

of the SU. Communication with the outside environment may then commence.

2.4 ENABLING TECHNOLOGIES FOR COGNITIVE RADIO

The practical implementation of the concept of CR relies heavily on technological advances in radio communications hardware. Since a SU in a CR network needs to be able to switch quickly and dynamically between different operating channels, preferably without having to physically change its hardware configuration, the SDR and the software defined antenna (SDA) are essential enabling technologies for CR communication.

2.4.1 Software-defined radio

One of the main enablers of CR is the SDR. The SDR is a radio communication system that provides for a greater level of communication flexibility than traditionally available. Advances in technology have made it possible for a SDR to access multiple frequency bands at one time. This is made possible by the software implementation of many components previously implemented in hardware. By making use of variable-frequency filters, oscillators and mixers, together with wide-band analogue-to-digital and digital-to-analogue converters (ADCs and DACs), a SDR is able to handle all aspects of the radio air interface in software [22]. The SDR is thus able to deliver real-time and dynamically programmable communication services through the software implementation of traditional hardware components such as filters, modulators, mixers and detectors. A SDR may be implemented using a personal computer or on embedded platforms such as: digital signal processors (DSP), field-programmable gate arrays (FPGA) or application-specific integrated circuits (ASIC) [8]. A basic illustration of a SDR is illustrated in Figure 2.4.

2.4.2 Software-defined antennas

The other major enabler for CR communication is the SDA. As with the SDR, the physical characteristics of the SDA are configurable in software, allowing for dynamically adjustable operation over a wide range of frequencies. Potential implementations of the SDA

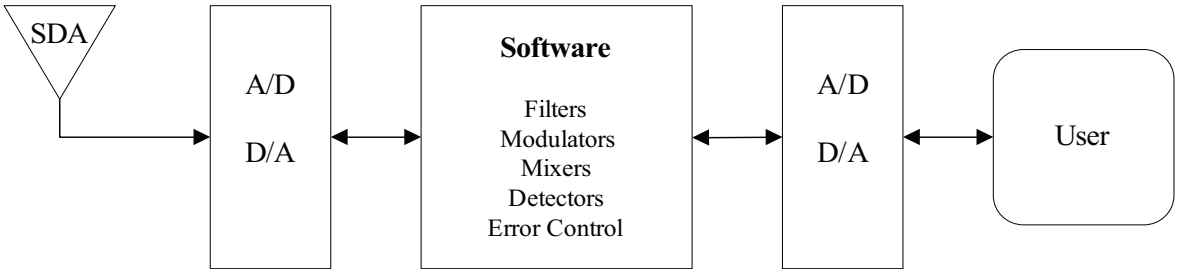

Figure 2.4: Simple illustration of a SDR using a SDA, adapted from [22].

are presented in [24, 25]. A microstrip patch antenna, made up of individually controllable pixels, is proposed to create a programmable and reconfigurable antenna array. The pixels in the array are mechanically actuated to change the physical dimensions and orientation of the antenna, which in turn changes its frequency, gain, polarisation and beam angle.

2.5 DETECTING SPECTRUM OPPORTUNITIES

Before a SU may operate in a CR network, it first needs to determine within which portion of the spectrum it may operate. An initial sensing of the spectrum must thus be performed. This process of detecting spectrum holes is one of the most important components of a CR system. Thus, measuring, sensing and learning about the radio environment within which the SU will be operating, are of great importance. It is imperative that knowledge about the radio environment, availability of spectrum, user requirements, local policies and other operating restrictions are obtained before SU communication may commence.

A number of methods, including energy detection, cyclostationary feature detection, radio identification, matched filtering, waveform-based sensing, cooperative sensing and sensing with multiple antennas, have been proposed to accomplish this task [23, 26]. Of these methods, energy detection is the most commonly applied technique presented in the literature, due to its relative ease of implementation and low computational complexity. A comparison between the accuracy and complexity of the first five of these SS techniques is shown in Figure 2.5 and a description of these techniques, as found in the literature, now follows [23, 27].

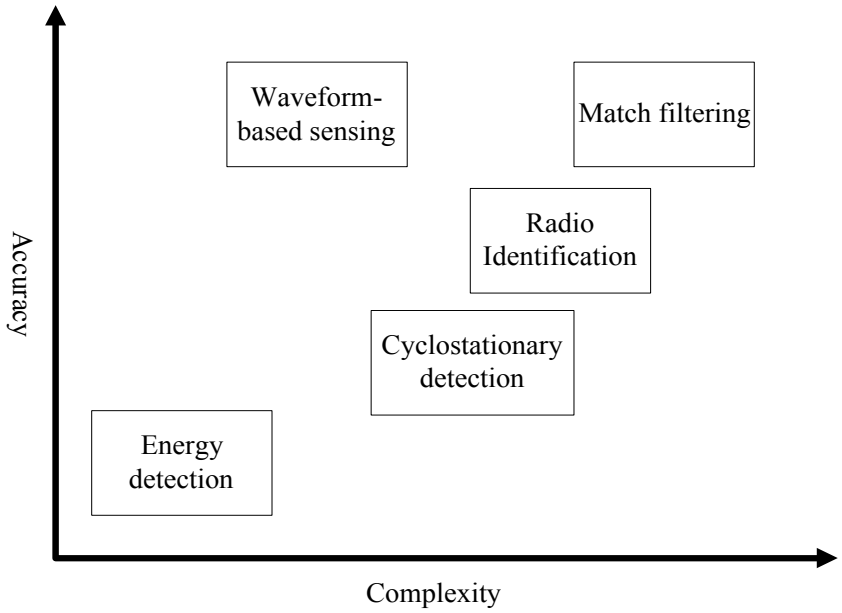

Figure 2.5: Relationship between SS techniques in terms of their accuracy and complexity, adapted from [23].

2.5.1 Energy detection

Energy detection (ED) is the most commonly implemented approach to SS for CR applications. Signal detection is performed by comparing the output of an energy detector to a predetermined threshold that is influenced by the environmental noise floor. Mathematically, the detected signal may be represented by the following signal model [23],

$$r(n) = s(n) + w(n), \tag{2.1}$$

where $s(n)$ is the original PU signal and $w(n)$ denotes additive white Gaussian noise (AWGN). If one assumes that PU channel occupancy follows a binary model of being either unoccupied H_0 or occupied H_1, then the following binary hypothesis holds,

$$\begin{aligned} H_0 : r(n) &= w(n), \\ H_1 : r(n) &= s(n) + w(n). \end{aligned} \tag{2.2}$$

Thus, in the absence of the PU, $s(n) = 0$. Alternatively $s(n) > 0$ when the PU is present.

To test the binary hypothesis presented in Equation 2.2, a signal detection threshold

Table 2.1: Binary hypothesis decision criteria.

Relationship	Hypothesis choice
$M \leq \lambda_T$	H_0
$M > \lambda_T$	H_1

λ_T needs to be calculated under noise only conditions. A decision metric M may then be compared against this threshold to determine channel occupancy, as shown in Table 2.1 [28].

If N is the observation vector size, the decision metric M may be represented as follows [23],

$$M = \sum_{n=0}^{N} |r(n)|^2. \tag{2.3}$$

Errors that may arise when testing the binary hypothesis include miss-detection and false alarm [10]. When performing SS it is desirable for the probability of either of these conditions occurring to be minimised. These erroneous conditions are defined as follows: the miss-detection probability (MP) is the probability of detecting the band of interest to free when it is actually already occupied and the false alarm probability (FAP) is the probability of detecting a state to be occupied by a PU when the PU is not actually present.

Although ED is easier to implement and provides lower computational complexity, it performs poorly under low signal to noise ratio conditions and may suffer under Rayleigh fading channel conditions. It has also been noted in [23] that sensing efficiency is compromised when detecting spread spectrum signals using ED. The basic structure of a conventional energy detector is illustrated in Figure 2.6 [29].

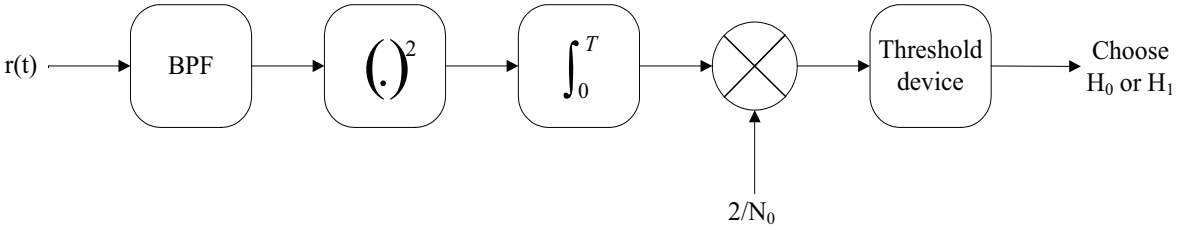

Figure 2.6: Generic block diagram of an energy detector, adapted from [29].

2.5.2 Cyclostationary feature detection

The cyclically varying features of a PU signal are exploited by the cyclostationary feature detection method. Statistical properties such as the mean and variance of the signal are used to detect the presence of a PU. Since noise may be assumed to be a wide-sense stationary process, modulated signals may be distinguished from noise. This also makes it possible to distinguish between different types of PU signals. The PU is detected by analysing the cyclic spectral density function (CSD) of a received signal, since the CSD outputs peak values when the cyclic frequency is equal to the frequency of the transmitted signal. The CSD may be expressed as follows [23],

$$S(f,\alpha) = \sum_{\tau=-\infty}^{\infty} R_r^{\alpha}(\tau)\, e^{-j2\pi f\tau}. \tag{2.4}$$

where α is the cyclic frequency and the cyclic autocorrelation function $R_r^{\alpha}(\tau)$ is given as,

$$R_r^{\alpha}(\tau) = E\left[r(n+\tau)\, r^*(n-\tau)\, e^{j2\pi\alpha n}\right]. \tag{2.5}$$

Several techniques for application in CR have been proposed in the literature [30, 31, 32, 28].

Although cyclostationary detection has the advantage of being robust in the presence of noise and channel attenuation, it has a disadvantage in that it requires a very high sampling rate, which in turn leads to high computational complexity.

2.5.3 Radio identification based sensing

PUs may also be detected by identifying the type of transmission technology that they are using. This involves two main tasks, initial mode identification (the initial search for a possible transmission) and alternative mode monitoring (the continual search for additional transmission modes during communication in an already selected mode). The most probable PU technology is selected by extracting various features from a received signal. In order to classify the type of signal, these features may be obtained by observing various parameters such as: the amount of energy detected, the centre frequency, the channel bandwidth and other statistical information. Amongst other methods, Bayesian classifiers, neural networks and

HMMs may use these features for classification [23].

2.5.4 Matched filter detection

If the characteristics of the transmitted signal are known to the receiver in advance, then matched filtering represents the most optimum scheme for detecting a PU [17]. If matched filtering is to be used for PU detection, then the receiver will need to have pre-knowledge of the PU's signalling features and be able to demodulate a number of different signal types. This presents the unwanted issue of impractical levels of receiver complexity [33].

Matched filter detection assumes the use of an optimal correlation receiver. If the received signal is denoted by $r(t)$ and the set of basis functions by $x(t)$, then the output of the correlator may be given by [17],

$$r_x = \int_{-\infty}^{\infty} r(t)x(t)dt. \tag{2.6}$$

The impulse response of the matched filter $h(t)$ is defined as,

$$h(t) = x(t-T), \tag{2.7}$$

where T is chosen so as to ensure the causality of the filter, i.e., $h(t) = 0, t < 0$. The impulse response $h(t)$ is thus matched to $x(t)$ provided that $y(t)$ is sampled at time $t = T$. The output of the filter $y(t)$ is obtained by the convolution of $r(t)$ and $h(t)$ and may thus be expressed as,

$$y(t) = \int_{-\infty}^{\infty} r(\tau)x(T-t+\tau)d\tau. \tag{2.8}$$

2.5.5 Waveform-based sensing

In systems where the waveform patterns are known, waveform-based sensing can also be used for PU detection. Known waveform patterns may be correlated with the received signals to determine if they belong to PUs. Examples of non-statistical patterns that may be compared include amongst others: transmitted pilot symbols, spreading sequences, and other known sequences such as preambles and mid-ambles. It has been noted that waveform-based

detection methods do provide the advantage of short measuring times, but also suffer from synchronisation issues [23]. Using the signal model from Equation (2.1), the decision metric M (presented in Table 2.1) for waveform-based sensing, may be described as follows,

$$M = \Re\left[\sum_{n=1}^{N} r(n)\, s^*(n)\right]. \tag{2.9}$$

2.5.6 Cooperative sensing

It has been shown in the literature [23, 34] that a collaborative approach to SS leads to an improvement in correctly detecting available spectrum. In cooperative sensing, SUs share information with each other about their individual sensing measurements. The real benefit of cooperative sensing is experienced under conditions where fading, shadowing and noise uncertainty, are prevalent. Significant reductions in the MP and FAP have been observed when cooperative sensing is employed under these conditions [34, 35]. Space, time and frequency diversity can also be exploited, when cooperative sensing is employed, to further improve SS performance [36].

However, when the cooperative sensing approach is employed issues pertaining to complexity as well as algorithm efficiency may arise. Some of the approaches employed by CRs to achieve cooperative sensing include: sharing information centrally for processing and decision making (this is also known as data fusion [27]), following a distributed sensing approach where multiple users perform processing independently before sending their outputs to a single user for final decision making (this is also known as decision fusion [27]) and gaining frequency spectrum information from external devices [23]. Energy detection based hybrid cooperative sensing algorithms have also been proposed that require less CR nodes to perform cooperative SS [37, 38]. It has also been shown in the literature [26], that soft combination for cooperative SS provides an improvement over hard combination.

2.5.7 Sensing with multiple antennas

Existing multiple input multiple output (MIMO) techniques, such as maximum ratio combining (MRC) and equal gain combining (EGC), may be employed to perform SS. In this way the time and spatial correlation between multiple versions of the received signal may be exploited [26, 39, 40].

In [27] optimal combining energy detection (OCED) is proposed for this purpose. OCED aims to maximize the signal to noise ratio (SNR) of a set of $M > 1$ combined receiver signals such that,

$$\Gamma(\boldsymbol{B}) = \frac{E\left[\left\|\boldsymbol{B}^T s(n)\right\|^2\right]}{E\left[\left\|\boldsymbol{B}^T w(n)\right\|^2\right]} \tag{2.10}$$

where $\boldsymbol{B}$ is the combining matrix and where the individual elements of Equation 2.1 change as follows for the MIMO case,

$$\begin{aligned} r(n) &= [r_1(n) \cdots r_M(n)]^T \\ s(n) &= [s_1(n) \cdots s_M(n)]^T \\ w(n) &= [w_1(n) \cdots w_M(n)]^T \end{aligned} \tag{2.11}$$

2.6 MODELLING SPECTRUM OPPORTUNITIES

The process of SS may be taken further to include predicting the future radio environment and the behaviour of PUs based on past knowledge. The purpose behind predicting future PU behaviour is to contribute towards improved CR performance [41]. Predicting future channel occupancy first requires SUs to observe/sense the frequency bands of interest to them over a fixed training interval. During this time, the SU gathers statistical information about the appearances and disappearances of other users of these bands. The statistical model generated from this information may then be used to predict how the future spectrum occupation of these bands will look. Proposed spectrum occupancy models in the literature include applying Markov chains (MC) and machine learning techniques to process channel usage statistics [42], modelling channel occupancy as an exponentially distributed process [43, 44] as well as applying regression techniques to binary time series channel representations [45].

Accurate prediction of future channel occupancy states contributes toward the SU channel selection process [43].

2.6.1 Markov modelling

The HMM has been suggested for modelling channel occupancy in the literature [10, 46, 47, 48, 49, 50]. A MC may thus be used to represent sub-band channel occupancy over a number of consecutive time periods L. In this approach, spectrum occupancy is modelled as having a particular state of being and a PU's current occupancy state will depend on its previous state. HMMs have been suggested for this purpose, since true occupancy states are not always known to SUs after the SS process. Maximum likelihood techniques have been proposed to aid in the prediction of true HMM states [10], however issues pertaining to prediction accuracy and complexity have arisen. The use of evolutionary based algorithms [51] has been proposed to further aid this process.

2.6.1.1 HMM structure

A HMM is a statistical process whereby PU channel occupancy may be modelled as a probabilistic finite state machine [52]. The basic structure of a HMM is shown in Figure 2.7.

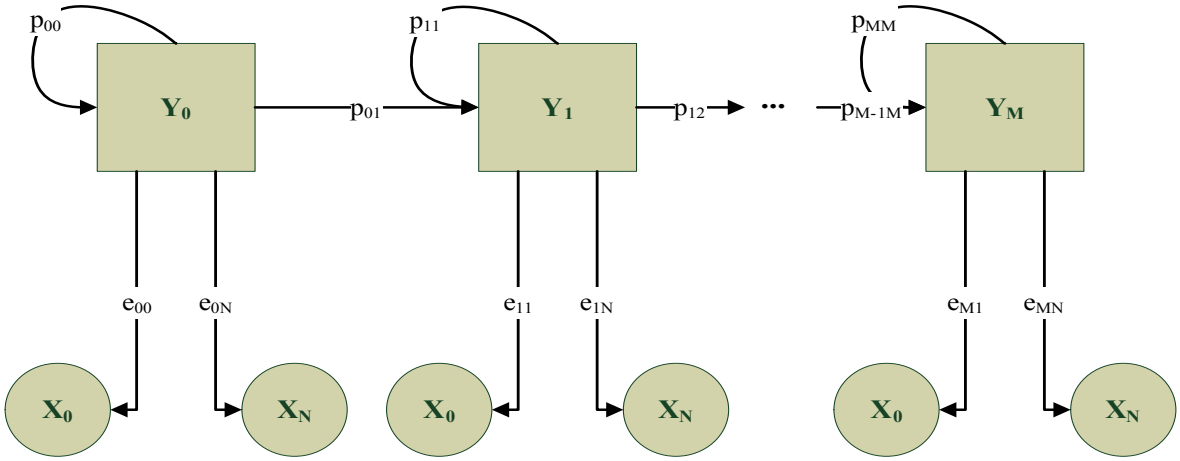

Figure 2.7: Basic structure of a HMM.

A HMM comprises a set $\boldsymbol{Y}$ of M possible states,

$$\boldsymbol{Y} = \{y_1, y_2, \cdots, y_M\}, \tag{2.12}$$

and a set $\boldsymbol{X}$ of N possible emissions,

$$\boldsymbol{X} = \{x_1, x_2, \cdots, x_N\}. \tag{2.13}$$

Two statistical parameters govern the operation of a HMM. The first parameter $\boldsymbol{P}$ is an M-by-M state transition probability matrix, representing the probabilities associated with changing from one state to another [17, 53], and is given by,

$$\boldsymbol{P} = \left(p_{ij}\right)_{M\times M}, \quad i, j \in \boldsymbol{Y}, \tag{2.14}$$

where the individual elements of Equation (2.14) are denoted by,

$$p_{ij} = P_r\left(y_n = j \mid y_{n-1} = i\right), \quad 2 \leq n \leq \tau, \tag{2.15}$$

and τ represents the length of an observation period. Similarly, the second parameter $\boldsymbol{E}$ is an M-by-N emission probability matrix that represents the probabilities associated with obtaining a certain output given that the model is currently in a true state y_n, and is given by,

$$\boldsymbol{E} = \left(e_{ij}\right)_{M\times N}, \quad i, j \in \boldsymbol{X}. \tag{2.16}$$

The individual elements of Equation (2.16) are denoted by,

$$e_{ik} = P_r\left(x_n = j \mid y_n = i\right), \quad 2 \leq n \leq \tau. \tag{2.17}$$

When attempting to predict a sequence of channel occupancy states with a HMM, it is necessary to find $P(\boldsymbol{X}|\boldsymbol{\lambda})$. The HMM parameters may be denoted by $\boldsymbol{\lambda} = (\boldsymbol{P}, \boldsymbol{E}, \pi)$, where π is the initial state distribution [54].

2.6.2 ON-OFF occupancy models

Channel occupancy is driven by human behavioural patterns and is thus not a purely random process. Methods for modelling PU occupancy that assume that PU activity follows an ON-OFF mode occupancy pattern have thus been proposed in the literature [43, 55]. For these methods, it is assumed that the PU is occupying the band of interest for an ON mode, and an OFF mode is assumed when the band of interest is unoccupied.

2.6.2.1 Alternative exponential approach

The most prominent of these methods assumes that usage by a PU of a set of frequency channels $q = 1,2\ldots,\vartheta$, follows a Poisson process with a PU arrival rate of λ_q, provided that only one PU occupies a channel over a time interval t_ψ. In this case the probability that a PU accesses the band of interest k times over this interval, may be expressed as follows [55],

$$f\left(k,\lambda_q\right) = \frac{{\lambda_q}^k e^{-\lambda_q}}{k!}, \quad q = 1,2\ldots,\vartheta. \tag{2.18}$$

This is known as the ON-OFF alternative exponential distribution. Channel occupancy is modelled as an independently exponentially distributed processes. A PU channel q, with an ON occupancy duration of length $t_{ON}\left(q\right)$ and mean ON duration of $1/\lambda_{t_{ON}(q)}$, may then be described by the following probability density function,

$$f\left(t_{ON}\left(q\right),\lambda_q\right) = \begin{cases} \lambda_q e^{-\lambda_q t_{ON}(q)}, & t_{ON}\left(q\right) \geq 0 \\ 0, & t_{ON}\left(q\right) < 0. \end{cases} \tag{2.19}$$

Similarly, a PU channel q with an OFF duration of length $t_{OFF}\left(q\right)$ and mean OFF duration $1/\lambda_{t_{OFF}(q)}$, may be described by the following probability density function,

$$f\left(t_{OFF}\left(q\right),\lambda_q\right) = \begin{cases} \lambda_q e^{-\lambda_q t_{OFF}(q)}, & t_{OFF}\left(q\right) \geq 0 \\ 0, & t_{OFF}\left(q\right) < 0. \end{cases} \tag{2.20}$$

2.6.2.2 Periodic approach

The periodic ON-OFF model is another method that follows an ON-OFF mode channel occupancy pattern. In this model it is assumed that over time, channel occupancy follows a pattern of fixed ON and OFF occupancy periods [43]. However, for it to be accurate, this method does require long term observations of PU occupancy behaviour.

2.6.3 Linear techniques

Due to their relative simplicity, various linear approaches to traffic modelling in wireless networks, have been proposed. These include regression modelling, the normalised least mean square (NLMS) algorithm and the recursive least squares (RLS) algorithm.

2.6.3.1 Regression model

A method that makes use of linear regression techniques to predict spectrum opportunities, is proposed in [45]. PU channel occupancy is modelled as a binary time series B_t with a time index t, where the presence of a PU is indicated as $B_t = 1$ and its absence as $B_t = 0$. With the model parameters given by β, the probability of success is then expressed as the sigmoid function transformation of the following expression,

$$\mu_t = \sum_{k=1}^{d} \beta_k B_{t-k}. \tag{2.21}$$

If $\boldsymbol{S}$ represents the band of interest to a CR, then a linear regression is performed on past observations and the probability of success is evaluated by logit transformation, such that,

$$P(S_{it} = 1 | S_{t-1}, S_{t-2}, \cdots, S_{t-p}) = \frac{1}{1 + e^{-\boldsymbol{M}}}, \tag{2.22}$$

where $\boldsymbol{M}$ is an auto regression function,

$$\boldsymbol{M} = a_0 + \sum_{j=1}^{d} \left(\boldsymbol{A}_j \boldsymbol{S}_{t-j} \right) + \boldsymbol{n}, \tag{2.23}$$

with the coefficient matrices $\boldsymbol{A}_j$, an error vector $\boldsymbol{n}$, an offset a_0 and a set of past observed values $\boldsymbol{S}_{t-j}$.

2.6.3.2 Recursive least squares algorithm

A method known as the RLS algorithm has been proposed for predicting traffic as an aid to proactive network management in wireless networks [56]. This method is reported to be a computationally efficient approach to predicting wireless network traffic over a relatively short time scale without having to make any assumptions about the statistics of the data being observed. Recent traffic observations are used to predict traffic load in the near future and the basis of the RLS method is rooted in adaptive filter theory.

Using the RLS algorithm, future traffic load $y(n)$ is estimated by the following expression,

$$y(n) = w^T (n-1) u(n), \tag{2.24}$$

where $u(n)$ represents a set of n recent traffic observations and $w(n)$ a corresponding set of weighting factors. If $d(n)$ represents the actual traffic load then the prediction error $e(n)$ may be described as,

$$e(n) = d(n) - y(n). \tag{2.25}$$

The RLS method tries to minimise the mean square value of Equation (2.25) by minimising the cost function $J(n)$, together with an exponential weighting factor λ, as follows,

$$J(n) = \sum_{i=1}^{n} \lambda^{n-i} [e(i)]^2. \tag{2.26}$$

The weight factor $w(n)$ is calculated so as to attain the minimum value for Equation (2.26) using the following expression,

$$w(n) = w(n-1) + K(n)\, e(n), \tag{2.27}$$

where $K(n)$ is known as the gain vector and is derived as follows,

$$K(n) = \frac{\lambda^{-1} P(n-1)\, u(n)}{1 + \lambda^{-1} u^T(n)\, P(n-1)\, u(n)}. \tag{2.28}$$

The cross correlation $P(n)$ between the observation signal and the actual traffic load signal, is calculated using the following expression,

$$P(n) = \sum_{i=1}^{n} u(i)\, d(i). \tag{2.29}$$

2.6.3.3 Normalised least mean square algorithm

Both the RLS and NLMS algorithms have the advantage that they do not rely on prior statistical traffic information. However, they have been reported to suffer performance degradation when there are rapid traffic variations [57]. The NLMS algorithm is similar to the RLS method, with the only perceivable difference being that the weight factor $w(n)$ is calculated as follows [58],

$$w(n) = w(n-1) + \frac{\mu e(n)\, u(n)}{\|u(n)\|^2}, \tag{2.30}$$

where μ is a step size constant.

2.7 CHANNEL ALLOCATION AND SWITCHING

Once SUs have determined which channels are available to them, they must then determine which of these channels will be best to operate within. Similarly, at the return of the PU to the channel, SUs must also determine which channels will be best to switch to. The main objectives here should be to contribute to achieving optimal performance in a CR network [41, 59]. These objectives include minimising interference to PUs, maintaining required quality of service (QoS) levels, facilitating disruption free communication by minimising delays and maximising throughput whilst still maximising spectral efficiency and limiting interference to PUs. SUs should not only choose channels that are simply available to them, but also look to select the channels that they will be switching to in an intelligent manner [43].

2.7.1 Switching techniques

In [41] and [43] different channel switching techniques are presented. These may be grouped into two basic switching categories, namely: reactive switching (RSW) and proactive switching (PSW).

2.7.1.1 Reactive switching

RSW is undesirable as it causes interference to the PU. This is due to the fact that when using RSW, SUs will continue to occupy a band until they detect the presence of a PU on the band, regardless of what may be happening on that band. If the arrival of a PU occurs in between a SS event (assuming that SS is performed on a periodic basis), then the SU will cause interference to the PU on that band, until the next SS and switching event is performed. Furthermore, once the SU does realise that it is causing interference, it will randomly pick and blindly switch to another channel. This new channel may already be occupied by a PU, in which case the shortcomings of RSW would then be perpetuated by causing further PU interference.

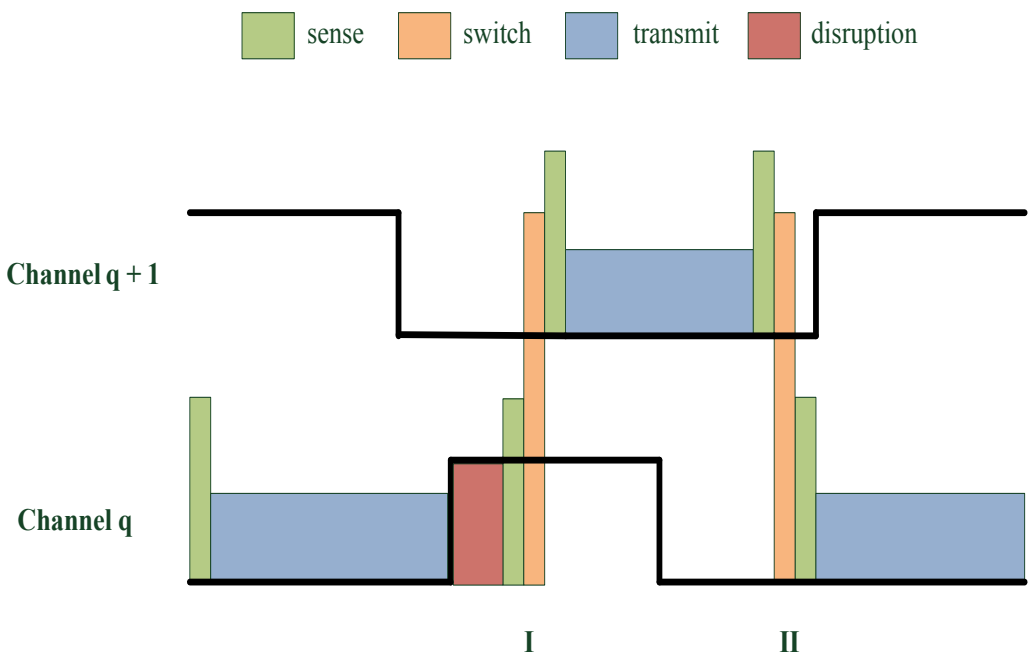

Figure 2.8: Illustration of reactive (I) and proactive (II) channel switching.

Basic RSW is illustrated by part I of Figure 2.8. The solid black line represents PU spectrum usage. When it is in the high position, a PU is occupying the channel q and the converse is true when it is in the low position. The green bars represent sensing operations and the orange bars switching operations. The blue sections represent data transmission and the red section represents a disruption to the PU, since the SU is still occupying the band even though the PU has returned.

2.7.1.2 Proactive switching

PSW attempts to address the problems associated with RSW by limiting interference to PUs and reducing the delays that may be caused by excessive channel switching. When PSW is employed, SUs are able to make informed channel switching decisions so as to avoid PU interference, since they are able to vacate the band before the PU actually arrives. This process is based on statistical knowledge gained about the past behaviour of PUs on that band. The performance of PSW is thus heavily reliant on the accuracy of the employed spectrum occupancy and prediction model. PSW is illustrated by part II of Figure 2.8. In this case the SU vacated the band before the PU returned.

PSW may be classed as either intelligent or dumb [43]. When PSW is performed in an intelligent way, the channel that the SU will switch to will not only be an available channel, but should be the best available channel for the SU to operate within. Various factors may be considered when making this decision, e.g., the channel that is least likely to be re-occupied may be deemed to be the best choice. In this way PU interference is avoided and the number of channel switches that will have to be performed may be reduced.

Dumb PSW will usually occur as a result of inaccurate channel prediction, causing the SU to either switch to a band that it perceives to be free but is in actual fact already occupied by a PU, or to switch to a channel that will not be available for as long as the band that it already occupies. The former, proactive busy switching (PBSW) will lead to interference with the PU and the latter, proactive short switching (PSSW), will inevitably introduce a delay into the system, because an extra channel switching event will then have to be performed. The above mentioned switching techniques are summarised in Table 2.2.

Table 2.2: Switching techniques for CR.

Switching technique	**Description**
Reactive switching	Switch channel after detecting PUs.
Proactive intelligent switching	SU switches to a channel that has a longer remaining idle time than the channel it currently occupies.
Proactive busy switching	SU switches to a channel that is already occupied.
Proactive short switching	SU switches to a channel that has a shorter remaining idle time than the channel it currently occupies.

2.8 IEEE STANDARDS FOR COGNITIVE RADIO

The IEEE 802.22 standard for cognitive wireless regional area networks (WRAN) (published in July 2011) and the IEEE DySPAN/P1900 standard for dynamic spectrum access networks (currently under development) are both applicable to CR. The IEEE DySPAN standard covers issues pertaining to dynamic and real-time usage of spectrum, taking into consideration the changing objectives and environmental circumstances of that spectrum. Since IEEE DySPAN is not yet available, only the IEEE 802.22 standard will be discussed in detail.

A high-level overview of the IEEE 802.22 standard for cognitive WRANs standard is presented in [60]. This document discusses the standards guidelines for how CR should be implemented to provide broadband wireless access in the WRAN context. It covers issues pertaining to both the physical and media access control (MAC) layers as well as basic cognitive functionality [60].

2.8.1 Physical layer features

The IEEE 802.22 standard specifies three major areas pertaining to CR implementation at the physical layer. These include: handling the main data communications, performing SS and providing geo-location functionality. The main physical layer features prescribed by the standard are listed in Table 2.3. The standard assumes that CR will operate within the favourable propagation environment of the current very high frequency (VHF) and ultra high frequency (UHF) licensed television bands and caters for 6 MHz, 7 MHz and 8 MHz channel bandwidths.

The standard also provides guidelines for the use of adaptive modulation and coding. Fourteen physical layer modulation and coding modes have been prescribed. These modes are tabulated in Table 2.4. Mode 1 will be used for the transmission of code-division multiple access (CDMA) ranging and bandwidth request messages and for urgent coexistence situation notification (UCS), mode 2 will be used for transmitting the coexistence beacon protocol (CBP) and modes 3 to 14 are provided for data communications purposes.

Table 2.3: CR physical layer features as prescribed by the IEEE 802.22 draft standard.

Description	**Physical layer features**
Network type	WRAN
Air interface	OFDMA
FFT	Single mode (2048)
OFDMA channel profile	6, 7, or 8 MHz
Max data rate	23, 27 or 31 Mb/s
Range	17- 30 Km
Operating frequencies	54 - 862 MHz (VHF/UHF)
Cyclic prefix	37 s (1/4, 1/8, 1/16, and 1/32)
Frame size	10 ms
Superframe size	160 ms (16 frames)
FEC	1/2 rate convolutional code, constraint length 7

2.8.1.1 Spectrum sensing

Provision is made for sensing of analogue television transmissions, digital television transmissions and other licensed low power devices, e.g., wireless microphones. The standard specifies that the sensing antenna must be mounted at least 10 m above the ground, be kept clear of any obstructions and have a reference antenna gain of 0 dBi. As listed in Table 2.5, receiver sensitivity levels for different signals types have also been specified.

2.8.1.2 Geo-location

The standard specifies the need for SUs to keep track of where they are geographically situated in relation to each other and in relation to the nearest base station. In the WRAN context, SUs need to be able to pin point the location of other SUs to within a radius of 100 m and the location the BS to within a radius of 15 m. Geo-location technologies such as global positioning system (GPS) or Gallileo are proposed for this purpose.

Table 2.4: IEEE 802.22 draft standard adaptive modulation and coding rates.

PHY mode	Modulation	Coding rate	Peak data rate at 6 MHz (Mb/s)	Peak data rate at 8 MHz (Mb/s)	Spectral efficiency
1	BPSK	none	4.54	6.05	0.76
2	QPSK	1/2 & repeat: 3	1.51	2.02	0.25
3	QPSK	1/2	4.54	6.05	0.76
4	QPSK	2/3	6.05	8.07	1.01
5	QPSK	3/4	6.81	9.08	1.13
6	QPSK	5/6	7.56	10.09	1.26
7	16-QAM	1/2	9.08	12.11	1.51
8	16-QAM	2/3	12.10	16.14	2.02
9	16-QAM	3/4	13.61	18.16	2.27
10	16-QAM	5/6	15.13	20.18	2.52
11	64-QAM	1/2	13.61	18.16	2.27
12	64-QAM	2/3	18.15	24.21	3.03
13	64-QAM	3/4	20.42	27.24	3.40
14	64-QAM	5/6	22.69	30.27	3.78

2.8.2 Frame structure

The MAC layer specification of the standard prescribes guidelines pertaining to CR frame structure. Figure 2.9 illustrates this specification. According to the standard, data should be

Table 2.5: IEEE 802.22 draft standard receiver sensitivity levels for different signal types.

Signal type	Receiver sensitivity (dBm)
Digital TV	-116
Analogue TV	-94
Wireless microphones	-107

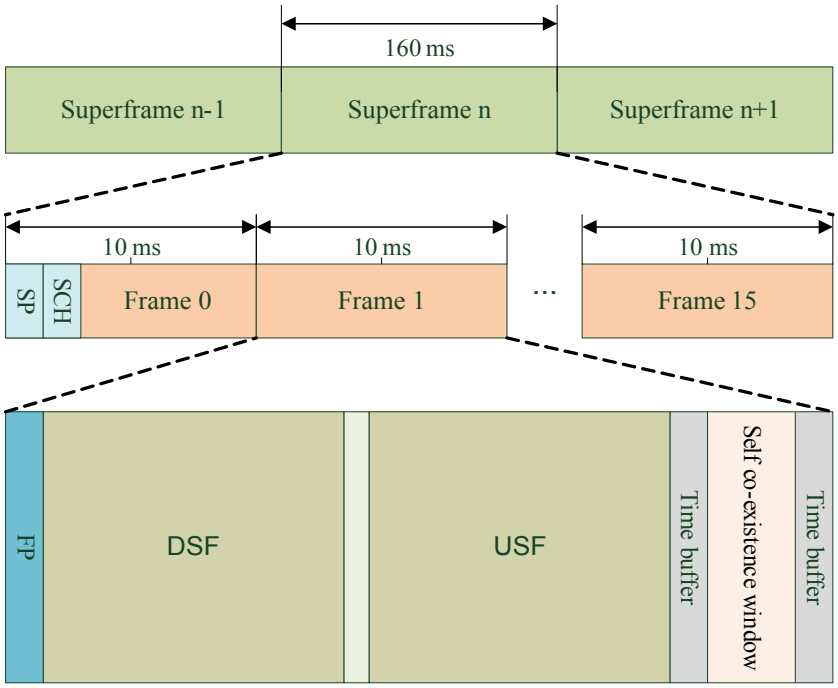

Figure 2.9: CR superframe and frame structure, adapted from [60].

transmitted in 160 ms sized superframes. These superframes are preceded by a superframe preamble and a superframe control header (SCH). They are in turn followed by a group of 16 MAC frames. Each MAC frame has a frame size of 10 ms and is itself preceded by its own frame preamble. The rest of the frame is comprised of a downstream subframe (DSF) and an upstream subframe (USF) with an adaptive boundary in between and a time buffered self co-existence window. The self coexistence window is used to limit mutual interference among SUs.

2.8.3 South African regulatory environment

The South African telecommunications and broadcasting sectors are regulated by ICASA. ICASA is an independent regulatory body that falls under the Department of Communications within the South African government. Since South Africa falls within Region 1 of the International Telecommunication Union (ITU), ICASA has allocated the VHF bands from 174 MHz to 230 MHz and 246 MHz to 254 MHz (band III), as well as the UHF bands from 470 MHz to 582 MHz (band IV) and 582 MHz to 854 MHz (band V) for terrestrial

Table 2.6: Potential bands for IEEE 802.22 WRAN implementation in South Africa.

Band	Frequency (MHz)
Band III	174 - 230; 246 - 254
Band IV	470 - 582
Band V	582 - 854

broadcasting, radio astronomy, fixed links, low power mobile radio and single frequency and trunked mobile applications [61]. The allocations are summarised in Table 2.6. These bands may be found suitable for future implementation of the IEEE 802.22 WRAN standard in South Africa, especially once the switch over from analogue to digital terrestrial television (DTT) has been completed.

2.9 CONCLUDING REMARKS

This chapter provided an overview of the concept of CR and its primary functionality. A theoretical platform was laid that covers the key areas of importance to the work described in the chapters that follow. Various SS techniques were discussed as well as methods pertaining to the modelling and prediction of PU activity in CR networks. The channel selection and switching processes required by SUs in a CR network were also briefly covered and an overview of the current IEEE standards pertaining to CR implementation was provided.

CHAPTER 3

TRAFFIC MODELLING AND PREDICTION

3.1 INTRODUCTION

Before SU communication may be initiated, and for it to successfully continue, it is imperative that the SUs in a CR network are able to distinguish between channels that are currently occupied by PUs and those that are available for SU communication. This is important since any form of interference to PUs is undesirable. Continuous communication relies on the correct identification of spectrum holes. Accurate and up-to-date information from the SS process is thus required. However, obtaining this information does introduce certain challenges to SU communication, since every time a SU engages in the SS process, valuable time that may otherwise have been spent on data communications, is lost.

Some of this time may however be redeemed, since accurate modelling and prediction of PU behaviour may be used to supplement the SS process [62]. By allowing a certain proportion of SS operations to be substituted for by occupancy predictions, a reduction in the number of SS operations required of SUs, may be achieved. A SU may experience the benefit of such a reduction in the form of increased data throughput. Since future prediction of PU activity is an essential prerequisite for proactive channel selection, a PU may benefit from fewer incidences of PU disruption. The ability of the SU to predict future PU behaviour may thus provide significant benefits to both sets of users.

In this chapter the model chosen by the author for the modelling and prediction of PU channel occupancy is described and various different algorithms for training the model are also presented. The training algorithms are also compared by performing a complexity analysis of each algorithm.

3.2 TRAFFIC OCCUPANCY MODEL

It is necessary to model and predict channel occupancy since enabling SUs to predict future PU behaviour allows for a potential improvement in the channel allocation process. One of the methods, suggested in the literature for modelling channel occupancy in CR networks, is the two-state HMM [10, 47, 48]. In this approach, sub-band spectrum occupancy is modelled as a consecutive sequence of binary states. The HMM is better suited to this problem than the less complex MC, since true occupancy states are not always known to SUs due to errors introduced through imperfections in the SS process.

3.2.1 Hidden Markov occupancy model

Due to its simplicity, a two-dimensional HMM with parameters $\boldsymbol{\lambda}$ (as described in Chapter 2.6.1), state space $\boldsymbol{Y} = \{0,1\}$ and emission state space $\boldsymbol{X} = \{0,1\}$, is proposed to model and predict PU channel occupancy [10]. The structure of the model is illustrated in Figure 3.1.

At any given time, a channel may be detected to be either occupied by a PU, $Y_i = 1$ or available for use by a SU $Y_i = 0$. Since the HMM will be used to predict a sequence of channel occupancy observations, both $P(\boldsymbol{X} \mid \boldsymbol{\lambda})$ and the HMM model parameters need to be determined. Initially, an educated guess is made for these parameters. A set of observations $\boldsymbol{O}(t) = \{o_{t+1}, o_{t+2}, \cdots, o_L\}$ is then collected, by observing the band of interest for a time interval comprising a maximum of L observations. The model parameters are then fine-tuned by feeding these observations into a model training algorithm that will attempt to maximise $P(\boldsymbol{O}|\boldsymbol{\lambda})$. After training has been performed, the model is used to predict the most likely sequence of near-future channel occupancy states $\boldsymbol{Y} = \{y_{t+1}, y_{t+2}, \cdots, y_N\}$. These

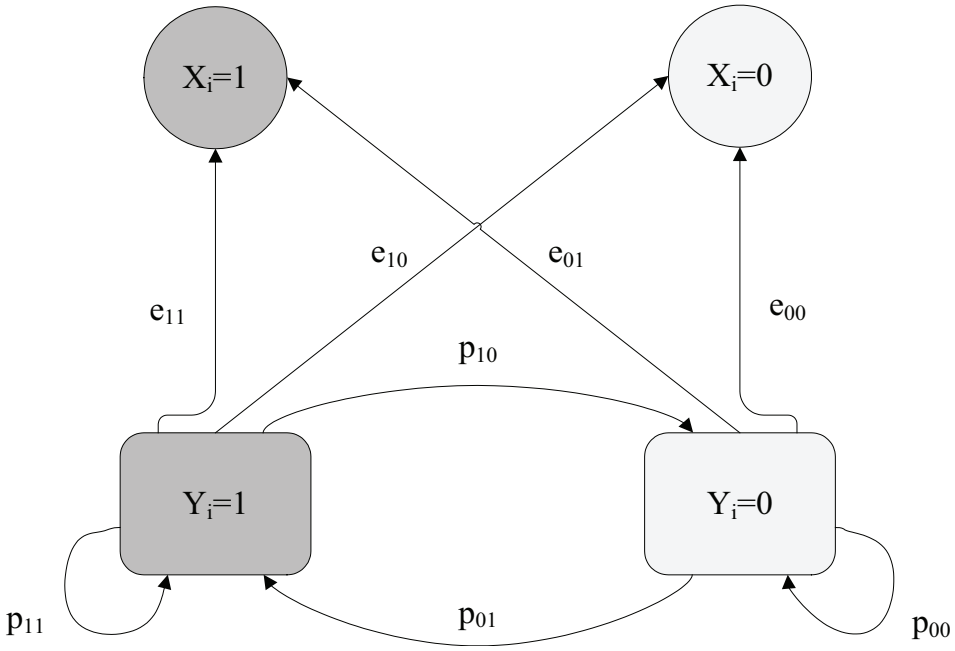

Figure 3.1: HMM representation of binary sub-band spectrum occupancy, adapted from [10].

predictions may be used to form the basis upon which future SU channel allocations may be performed. The channel allocation process will be discussed in more detail in Chapter 4.

3.3 MODEL TRAINING ALGORITHMS

The accuracy level that the model is able to achieve is influenced by the accuracy of its parameters $\boldsymbol{\lambda}$, i.e. if the model is poorly trained then it will not accurately represent the statistics of the channel within which the SU is attempting to operate. Consequently, the accuracy of predicted PU behaviour may be negatively affected. It is therefore important to select an algorithm that is able to train the model accurately. Although accuracy is the key requirement of the model, computational complexity is also important. The algorithms employed for training our model are now presented.

3.3.1 Baum-Welch algorithm

The Baum-Welch algorithm (BWA) is an instance of a generalized expectation-maximisation (EM) algorithm used for estimating HMM parameters $\boldsymbol{P}$ and $\boldsymbol{E}$. It makes use of both the forward and backward algorithms [53]. If the number of hidden and visible states are known to a SU, then it is possible for the SU to obtain a good solution for $\boldsymbol{\lambda}$ by iteratively updating the weights of the BWA until certain convergence criteria have been met.

To begin with, a definition of the forward and backward variables and there associated algorithms is discussed. The forward variable $\alpha_t(i)$ is suggested as a means to lowering complexity when finding $P(\boldsymbol{X}|\boldsymbol{\lambda})$. This is the probability of obtaining the partial observation sequence, $\boldsymbol{O}$, when it terminates in state i. In mathematical terms, $\alpha_t(i)$ may be described as follows [63],

$$\alpha_t(i) = P(o_1, o_2, \cdots, o_t, y_t = i|\boldsymbol{\lambda}). \tag{3.1}$$

This forms part of a recursive relationship known as the forward algorithm, which may be expressed as,

$$\alpha_{t+1}(j) = e_j(o_{t+1})\gamma_t, \quad 1 \leq j \leq M, 1 \leq t \leq L-1, \tag{3.2}$$

where γ_t is given by,

$$\gamma_t = \sum_{i=1}^{M} \alpha_t(i)\, p_{ij}, \quad 1 \leq j \leq M, 1 \leq t \leq L-1. \tag{3.3}$$

The first instance of $\alpha_t(i)$, with initial state distribution π, may be expressed as follows,

$$\alpha_1(j) = \pi_j e_j(o_1), \quad 1 \leq j \leq M. \tag{3.4}$$

Therefore, since it is possible to calculate,

$$\alpha_L(i), \quad 1 \leq i \leq M, \tag{3.5}$$

the probability of obtaining the emission state space $\boldsymbol{X}$ (given the model parameters $\boldsymbol{\lambda}$) may be expressed as,

$$P(\boldsymbol{X} \mid \boldsymbol{\lambda}) = \sum_{i=1}^{M} \alpha_L(i). \tag{3.6}$$

Similarly the backward variable $\beta_t(i)$ may be defined as the probability of obtaining the partial observation sequence, $\boldsymbol{O}$, when its current state is i. In mathematical terms, this may be expressed as,

$$\beta_t(i) = P(o_{t+1}, o_{t+2}, \cdots, o_L, y_t = i|\boldsymbol{\lambda}). \tag{3.7}$$

Once again a recursive relationship follows for calculating $\beta_t(i)$ that includes the elements of both $\boldsymbol{P}$ and $\boldsymbol{E}$. This relationship is known as the backward algorithm and may be expressed,

$$\beta_t(i) = e_j(o_{t+1})\eta_{t+1}, \quad 1 \leq i \leq M, 1 \leq t \leq L-1, \tag{3.8}$$

where η_t is now given by the following expression,

$$\eta_{t+1} = \sum_{j=1}^{M} \beta_{t+1}(j)\, p_{ij}, \quad 1 \leq i \leq M, 1 \leq t \leq L-1, \tag{3.9}$$

and where initially,

$$\beta_L(i) = 1, \quad 1 \leq i \leq M. \tag{3.10}$$

Once the forward and backward algorithms have been defined, it is possible to define the probability $\gamma_t(i)$ of being in a particular state y_i at time t, given the observation sequence $\boldsymbol{O}$ and model parameters $\boldsymbol{\lambda}$. This probability may be expressed as follows,

$$\gamma_t(i) = P(q_t = y_i|\boldsymbol{O}, \boldsymbol{\lambda}), \tag{3.11}$$

which in terms of the forward and backward variables, may be expressed as,

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(\boldsymbol{O}|\boldsymbol{\lambda})}. \tag{3.12}$$

Since the values calculated for $\alpha_t(i)$ and $\beta_t(j)$ from the forward and backward algorithms are only estimates, a new parameter $\xi_{ij}(t)$ may be defined. This is the probability of the state transitions between y_i at time t and y_j at time $t+1$ given that the SU has knowledge of the complete visible sequence $\boldsymbol{O}$. This is defined as,

$$\xi_t(i,j) = P(q_t = y_i, q_{t+1} = y_j|\boldsymbol{O}, \boldsymbol{\lambda}). \tag{3.13}$$

Writing Equation 3.13 in terms of the forward and backward variables leads to the following expression,

$$\xi_t(i,j) = \frac{e_j(o_{t+1})\,\alpha_t(i)\,\beta_{t+1}(j)p_{ij}}{P(\boldsymbol{O}|\boldsymbol{\lambda})}. \tag{3.14}$$

Using this new parameter, more accurate estimates for p_{ij} and $e_j(o_{t+1})$ may now be obtained. Since the expected number of state transitions from y_i at any time is given by $\sum_{t=1}^{T-1}\gamma_t(i)$ and the expected number of state transitions from y_i to y_j is given as $\sum_{t=1}^{T-1}\xi_t(i,j)$, improved model parameter estimates may be calculated. The improved initial probability estimate may thus be given as,

$$\hat{\pi}_i = \gamma_1(i), \tag{3.15}$$

and the improved transition probability estimate as,

$$\hat{p}_{ij} = \frac{\sum_{t=1}^{T-1}\xi_t(i,j)}{\sum_{t=1}^{T-1}\gamma_t(i)}. \tag{3.16}$$

Similarly, the improved emission probability estimate may be expressed as,

$$\hat{e}_{jk} = \frac{\sum_{t=1,O_t=x_k}^{T}\gamma_t(j)}{\sum_{t=1}^{T}\gamma_t(j)}. \tag{3.17}$$

which is the ratio between the frequency that a particular symbol x_k will be emitted and the frequency that any symbol will be emitted.

3.3.2 Viterbi algorithm

The Viterbi algorithm (VA) may be applied to the training of a HMM and is well suited to the HMM decoding problem. The VA may be used to calculate the most likely sequence of states $V_{t,k}$ that may be obtained from the set of observations $\boldsymbol{O}(\boldsymbol{t})$. The algorithm is given by the following expression [64],

$$V_{t,k} = P(o_t|k)\,.\max_{y\in\boldsymbol{Y}} p_{y,k}V_{t-1,y}, \tag{3.18}$$

which may also be expressed in logarithmic form as,

$$V_{t,k} = log[P(o_t \mid k)] + \max_{y\in\boldsymbol{Y}}\left[log\left(p_{y,k}\right) + V_{t-1,y}\right]. \tag{3.19}$$

Using Equation (3.19) rather than Equation (3.18) helps us to avoid the underflow problem, which occurs when $p_{y,k}$ gets really small. Given that the final state is k, the output of Equation

(3.19) corresponds to the most likely sequence of states for the first $t+1$ observations. By saving the actual states that the model was in at each time instance in the sequence, it is possible to backtrack through the sequence and calculate the maximum likelihood Viterbi path as follows,

$$y_L = \arg\max_{y \in \boldsymbol{Y}} V_{L,y}. \tag{3.20}$$

An example of such a path, for a sample of three consecutive states, is illustrated by the green arrows in Figure 3.2. A path has been calculated where, at time $t = 0$, the band is unoccupied ($Y = 0$) but then changes state and becomes occupied ($Y = 1$) at $t = 1$ and remains that way at $t = 2$.

From initial HMM parameter estimates, Equation (3.20) may be used to obtain a sequence of predicted states from which new estimates for the HMM transition and emission parameters may be calculated. This process is repeated until the model generated by the new HMM parameter estimates matches the observed sequence to within a pre-defined error tolerance level.

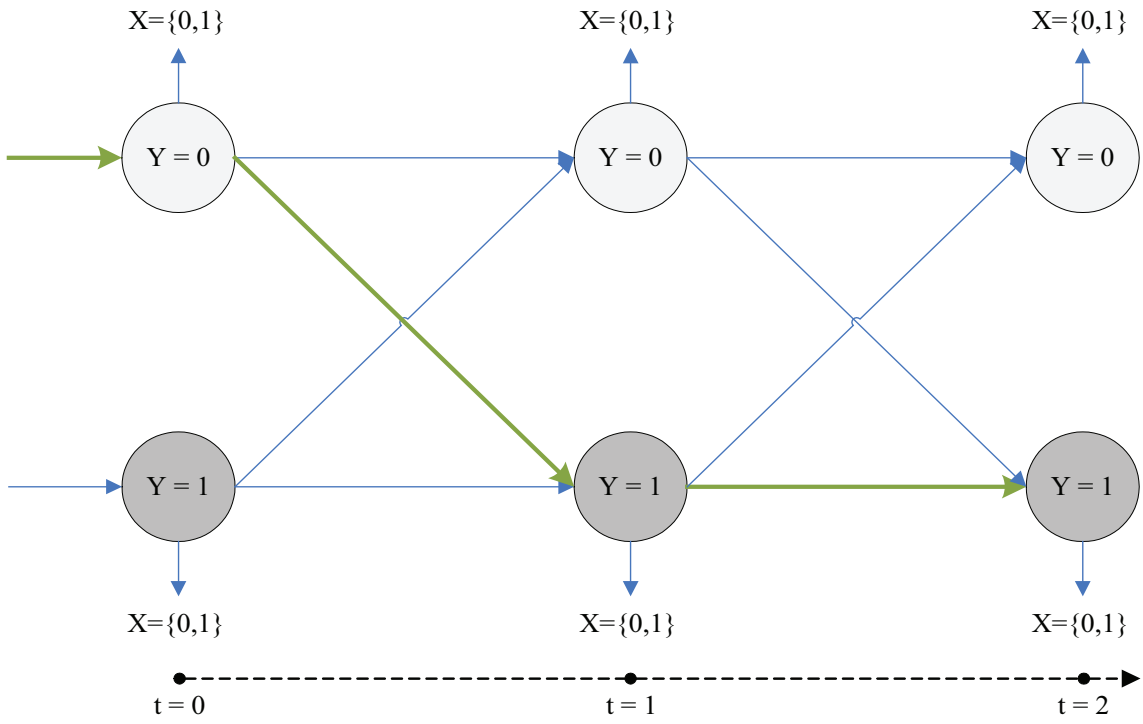

Figure 3.2: Trellis diagram used by the VA for training the HMM channel occupancy model.

3.3.3 Particle swarm optimisation

Particle swarm optimisation (PSO) is an evolutionary-based algorithm that imitates the social behaviour of a flock of migrating birds trying to reach a certain destination [65]. In contrast to traditional evolutionary computing, individuals are evolved by both competition and cooperation amongst the individuals of the flock. These birds/particles communicate with each other at fixed intervals and then adjust their flying speeds and resultant positions accordingly [66, 67]. Each bird's flying speed is a function of both its local position and the position of the best bird in the flock. Thus PSO incorporates both intelligence and social interaction, since particles learn from both their own experience (local search) and from each other (global search).

3.3.3.1 Algorithm

The process followed by PSO is as follows. In an S-dimensional space consisting of N particles, each particle i monitors three values: its own current position,

$$\boldsymbol{X}_i = (x_{i1}, x_{i2}, \cdots, x_{iS}), \tag{3.21}$$

the best position it reached in previous cycles,

$$\boldsymbol{P}_i = (p_{i1}, p_{i2}, \cdots, p_{iS}), \tag{3.22}$$

and its own flying velocity,

$$\boldsymbol{V}_i = (v_{i1}, v_{i2}, \cdots, v_{iS}). \tag{3.23}$$

At the end of every time interval t, each particle's position is evaluated according to a predefined and problem dependent fitness function. The best position of the particles $\boldsymbol{P}_i$ is then updated by comparing each particle's current position to its previously best achieved position and the particle that is deemed to have the global best position $\boldsymbol{P}_g$, is then calculated and stored. Each particle then has its velocity updated according to the following expression,

$$\boldsymbol{V}_i^{t+1} = \omega \boldsymbol{V}_i^t + c_1 r_1 \left[\boldsymbol{P}_i^t - \boldsymbol{X}_i^t\right] + c_2 r_2 \left[\boldsymbol{P}_g^t - \boldsymbol{X}_i^t\right]. \tag{3.24}$$

This is illustrated in Figure 3.3. The constants c_1 and c_2 represent the amount by which the velocity is updated and the constants r_1 and r_2 are causal uniformly distributed random variables. An inertial weight factor ω is also included to balance the local and global search ability. The optimal balance between local and global search depends on the problem being solved [66]. Through a process of trial and error, a choice was made to decrease ω linearly from a numerical value of 0.8 down to 0.6.

The maximum particle velocity increase allowed at each time interval is limited by V_{max}, in accordance with the following constraint,

$$-V_{max} \leq \boldsymbol{V}_i^{t+1} \leq V_{max}. \tag{3.25}$$

Using the new set of velocities, each particle's position X_i may be updated according to the following expression,

$$\boldsymbol{X}_i^{t+1} = \boldsymbol{X}_i^t + \boldsymbol{V}_i^{t+1}. \tag{3.26}$$

This process is then repeated until the algorithm's termination criteria are reached. This is achieved when the particles either converge upon a solution within an acceptable tolerance level or the algorithm reaches a specified maximum number of iterations.

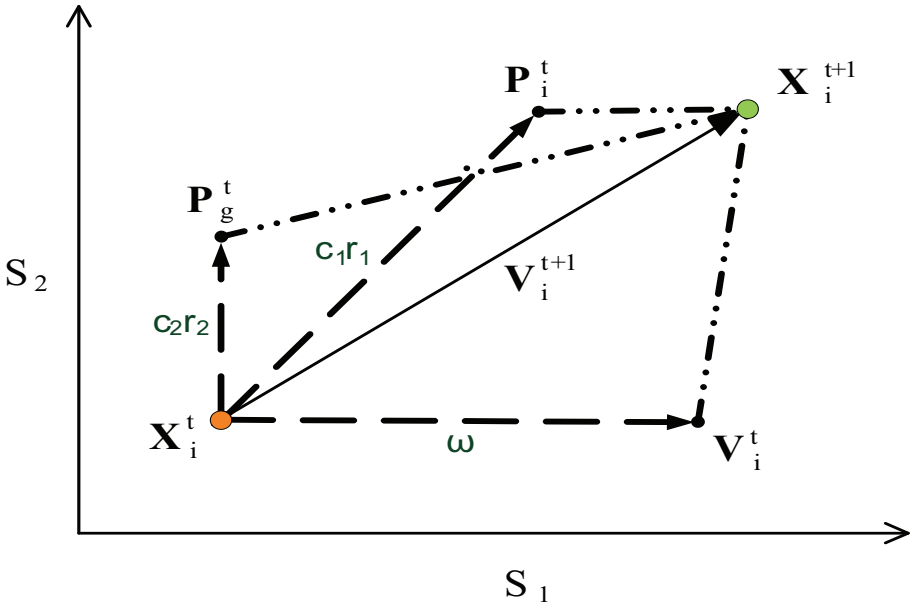

Figure 3.3: Graphical illustration of the PSO training algorithm, adapted from [66].

3.3.3.2 HMM training

PSO may be used to train a HMM. The position of each particle in the swarm is given by an $\boldsymbol{X}+\boldsymbol{Y}$ multi-dimensional set of HMM parameters $\boldsymbol{\lambda}$, i.e. the elements of both $\boldsymbol{P}$ and $\boldsymbol{E}$ are represented within the dimensions of each particle. The starting position of each particle is a random variation of the initial guess for the HMM parameters. Thereafter, each particle's position is adjusted according to the outcome of the applied fitness function.

At each iteration of the algorithm, the proposed fitness function, compares a predicted sequence of length $\boldsymbol{W}$ PU observations $\boldsymbol{K} = \{k_{w+1}, k_{w+2}, \cdots, k_W\}$ to an actual set of measured training observations $\boldsymbol{O} = \{o_{t+1}, o_{t+2}, \cdots, o_T\}$. It is assumed that $T = W$. Each particle's current position (i.e. HMM parameters) is used to generate a predicted sequence by maximising $P(\boldsymbol{K}|\boldsymbol{\lambda})$. The error between the two sets of observations may be defined to be,

$$\Psi_{Prd} = \sum_{t=0}^{N} (\boldsymbol{O} - \boldsymbol{K})/W. \tag{3.27}$$

The error calculated in Equation (3.27) is used to update each particle's best position. The global best position of all N particles may then be calculated as the particle that has the smallest value for Ψ_{Prd}. Using this information, each particle's position is updated according to Equation (3.24).

3.3.4 Memetic algorithm

The memetic algorithm (MA) is presented in the context of a genetic algorithm (GA), since the MA is quite similar to the GA. Both the GA and MA attempt to find a solution by following the evolutionary process of natural selection and survival of the fittest. The main difference between the two, however, is that in a MA the offspring gain a certain amount of experience through local searching. Thus, the MA also bases survival on the most experienced members of the population by performing additional local searching, i.e., a MA is a GA that heavily uses local search. It is thus sometimes also referred to as a cultural algorithm [65, 68].

The MA consists of a population of chromosomes. These chromosomes contain memes, which hold values for the optimisation variables (in a GA, chromosome elements are called genes). Chromosomes and offspring will first gain experience through local search before evolution is performed to bring about the next generation.

The main factors that affect the performance of the MA include the size of the population N_C, the number of generations generated, the crossover rate P_{CR}, the mutation rate P_{MUT} and the local search mechanism employed.

The key elements of the MA are now briefly introduced [65, 68, 69, 70].

1. Selection: at each successive generation, a set of parents Z_n are selected for breeding. Our approach to selection is to match the best parents, as determined by a predefined fitness function. This is based on the assumption that if both parents provide a good solution, then their offspring should do likewise.

2. Crossover: crossover is the process whereby information is exchanged between parent members of the same generation. The assumption here is that the offspring will be a combination of the best attributes of both sets of parents. The simplest method for performing crossover is known as single-point crossover. For a chromosome of length L_C a random point is selected at a location p, where $1 \leq p \leq L_C - 1$ [59]. The two parent chromosomes are then interchanged at this point to generate two child chromosomes C_1 and C_2 as follows,

$$\boldsymbol{C}_1 = \{Z_{1,0}, \cdots, Z_{1,p}\} + \{Z_{2,p+1}, \cdots, Z_{2,L_C}\}, \tag{3.28}$$

$$\boldsymbol{C}_2 = \{Z_{2,0}, \cdots, Z_{2,p}\} + \{Z_{1,p+1}, \cdots, Z_{1,L_C}\}. \tag{3.29}$$

   It is recommended that the rate at which crossover occurs falls within the range $0.6 \leq P_{CR} \leq 1.0$ [65]. A crossover rate of $P_{CR} = 0.8$ has thus been chosen. The single point crossover operation is illustrated in Figure 3.4.

3. Mutation: mutation is a rare occurrence where one or more genes within a chromosome are selected and mutated according to a mutation probability P_M. Usually,

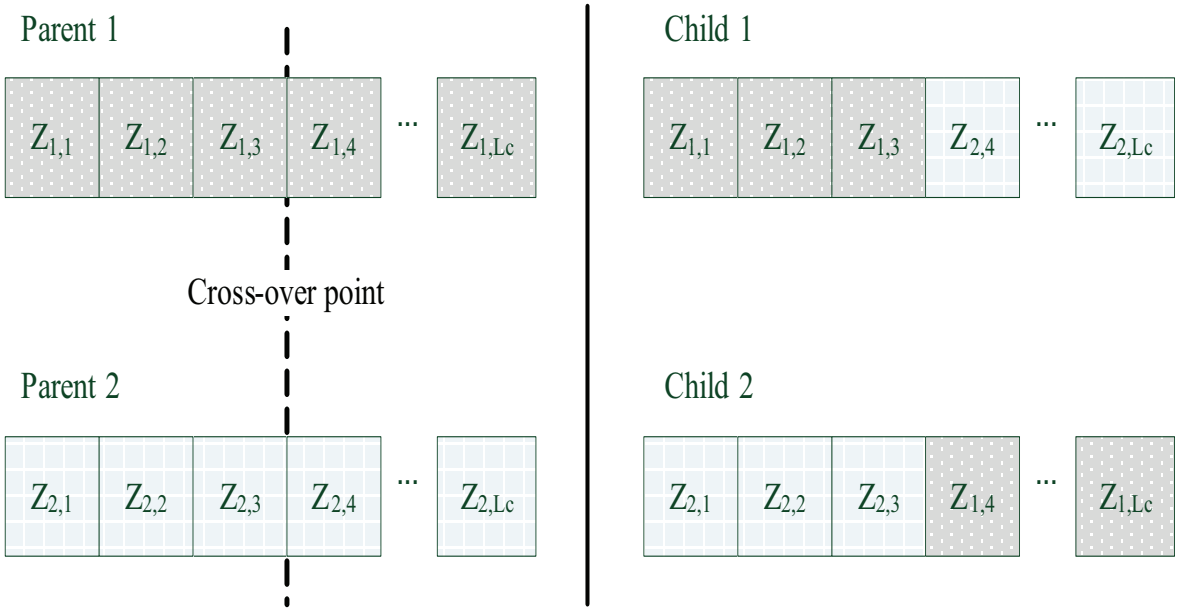

Figure 3.4: Illustration of the single-point cross-over operation.

mutation rate is kept fairly low, $P_M < 0.1$ [65]. Through this operation new genetic material is randomly introduced into the evolutionary process, helping to prevent stagnation around local minima. The simplest way of performing mutation is to swop the positions of randomly selected genes/memes within a chromosome [59], another way that this may be done is by arbitrarily selecting and randomly changing some of genes/memes contained within that chromosome [65], in this case, mutation is limited by a maximum mutation rate parameter ζ_m, such that $P_M \leq \zeta_m$. The latter method with $\zeta_m = 0.08$ has been adopted.

4. Local Search: the MA makes use of a local searching operation that provides chromosomes of the same generation with experience. For every generation, each chromosome is slightly modified before being involved in the evolutionary process. In [69] a local search through pair-wise information interchange heuristic is proposed. This local searching method can however be designed to suit the specific nature of the problem being solved. A modified approach is suggested in [65], where an incremental value x, limited by Λ_m, is added to each gene in a chromosome and then tested for improved performance. Only chromosomes that exhibit improved performance are

retained, otherwise they are ignored [65]. This process may be expressed as follows,

$$Z_{i,j} = Z_{i,j} + x, \quad -\Lambda_m \leq x \leq \Lambda_m, 1 \leq i \leq N_C, 1 \leq j \leq L_C. \tag{3.30}$$

3.4 COMPLEXITY ANALYSIS

Limiting the complexity associated with modelling and predicting PU activity is of importance to SUs for at least two reasons: namely the effect that this process has on SU data throughput and on SU power consumption. More time spent on training the HMM occupancy model means that less time is available for transmitting data. Greater training algorithm complexity places greater processing and power consumption demands on SUs. The computational complexity of the algorithms described in this chapter will thus be analysed and compared.

3.4.1 Big O notation

One approach to analysing training algorithm complexity, is big O notation. By characterising growth rate, a worst-case scenario of the amount of execution time required by each algorithm is provided [71]. The big O notation that describes a single iteration of each algorithm is summarised in Table 3.1, where N is the number of symbols in the HMM state alphabet, M is the PSO population size, W is the MA population size and L is the training sequence length. Since N, M and W may differ in size and since big O notation does not take into consideration the number of iterations i required by each algorithm to converge on a good solution, it is difficult to make an accurate comparison between the algorithms by following this approach.

3.4.2 Floating point operations

As an attempt at compensating for the perceived shortcomings of big O notation, another approach to calculating algorithm complexity is considered. This approach incorporates i into a calculation of the total number of floating-point operations per second (FLOPS) required by each algorithm to train the HMM occupancy model. In this approach algorithm

Table 3.1: Training algorithm complexity according to big O notation.

Algorithm	Complexity
BWA	$O\left(N^2L\right)$
VA	$O\left(N^2L\right)$
PSO	$O\left(ML\right)$
MA	$O\left(W^2L\right)$

complexity may be estimated according to the assumptions listed in Table 3.2. These values are based on the assumption that it only requires one FLOPS to execute an addition, subtraction, multiplication or division of two floating-point numbers [72].

The total complexity C_{Tot} of each algorithm may thus be calculated by multiplying the complexity C_s (expressed in FLOPS), required to perform a single iteration of the algorithm, by i. This may be expressed as follows,

$$C_{Tot} = i \times C_s. \tag{3.31}$$

In Table 3.3, the training algorithms have been broken down into their major functional components so as to calculate a complexity estimate for a single iteration of each algorithm. Polynomial estimates for the complexity of each component, as well as the final estimated

Table 3.2: Floating point operation assumptions.

Label	Description	Complexity
a	vector-vector scalar ($x, y \in \boldsymbol{R}^n$)	n
b	vector-vector inner product ($x, y \in \boldsymbol{R}^n$)	$2n$
c	vector-matrix product ($A \in \boldsymbol{R}^{m \times n}$)	$2n^2$
d	matrix-matrix product ($A \in \boldsymbol{R}^{m \times n}, B \in \boldsymbol{R}^{n \times p}$)	$2n^3$
e	randomisation function	n^2
f	logarithm function	$20n$
g	exponential function	$40n$

Table 3.3: Comparison of the estimated complexity required to perform a single training algorithm iteration C_s by considering the number of FLOPS required when $n = 4$.

Algorithm	Function	Polynomial	Complexity (C_s)	FLOPS
BWA	decode HMM	$5a+3b+d$	$2n^3+221n$	1012
	parameter update	a		
	calculate **P**	$5a+4f+g$		
	calculate **E**	$2a+1b+2f+g$		
VA	generate trellis	$5a$	$8n^3+12n$	560
	estimate $\boldsymbol{\lambda}$	$2a+2b+2d$		
	parameter update	$a+2d$		
PSO	calculate ω	$5a$	$18n^3+10n^2+17n$	1380
	test fitness	$3a+1b+2d+2e$		
	new position	$2a+4c+7d$		
	convergence	$5a$		
MA	test fitness	$3a+b+2d+2e$	$22n^3+26n^2+35n$	1964
	local searching	$15a+3b+6c+9d+9e$		
	crossover	$a+e$		
	mutation	$3a+c+2e$		
	convergence	$5a$		

complexity of a single iteration C_s of each algorithm, are listed. In the last column, an estimate for the number of FLOPS required by the algorithm, when $n = 4$, is presented.

These estimates indicate that the MA and the VA exhibit by far the largest and least amounts of single iteration computational complexity respectively, while the BWA and PSO algorithms would appear to fall somewhere in between. The BWA and PSO algorithms exhibit similar levels of single iteration complexity when n is small, however, PSO would become more complex than the BWA for larger values of n. In Chapter 6, the presented results together with Equation (3.31), will be employed to verify these theoretical complexity observations.

3.5 CONCLUDING REMARKS

In this chapter, a two-state HMM, for modelling PU traffic occupancy, was presented and described. This model requires a set of observations from which it may be trained. Four algorithms that could potentially be used to train the model were discussed and a theoretical complexity analysis was performed to help compare them. From this analysis the VA emerged as the least complex option (for a single iteration of the algorithm) for training the HMM occupancy model. The HMM model discussed in this chapter forms the basis for the discussion that follows in Chapter 4 on OSA. The performance of the HMM model, in combination with these training algorithms, will be further explored through the simulated and practically measured results that will be presented in Chapter 6.

CHAPTER 4

SECONDARY USER PERFORMANCE SIMULATOR

4.1 INTRODUCTION

The way in which a CR selects which channels to operate within, has a significant impact on the performance of a cognitive radio network. In Chapter 2, the concept of DSA was introduced, according to the IEEE standard definitions and concepts for DSA [20] a CR should follow the DSA approach to assign spectrum to SUs on a real-time basis as dictated by CR requirements and circumstantial changes. In this chapter a model is presented that allows SUs to employ a form of DSA known as OSA. This model assumes that SUs are able to exploit spectrum opportunities in a non-interfering way and without any form of negotiation between SUs and PUs [20]. In a CR network, OSA is an on-going process that needs to be continually performed so as to keep up with changes in PU behaviour.

This chapter will take a closer look at some of the relevant aspects pertaining to OSA, specifically with the goal of investigating how the channel occupancy modelling and prediction method, discussed in Chapter 3, may affect CR radio performance. To do this, a HMM prediction model based simulation platform is presented. This platform allows for an investigation into the effects that the occupancy model may have on CR performance measures such as achievable SU data throughput, PU disruption and basic SU power consumption.

4.2 OPPORTUNISTIC SPECTRUM ALLOCATION MODEL

A basic model for performing OSA in a CR network is described in this section. The model forms the basis of the simulation platform that will be discussed in the Section 4.3. The model is illustrated in Figure 4.1.

When a SU first enters a CR network, it needs to gather information about the environment within which it wishes to operate. Thus, the first step in the model, is for the SU to sense its current operating channel for a fixed time period so as to obtain a sequence of channel occupancy observations $\boldsymbol{O} = \{o_{t+1}, o_{t+2}, \cdots, o_T\}$. These observations are used together with an initial guess for $\boldsymbol{\lambda}$ to train the HMM occupancy model and obtain values for $\boldsymbol{\lambda}$ that accurately describe $\boldsymbol{O}$. The model is then used to gather channel availability information from a prediction of near future PU occupancy states. Once future channel availability is known to the SU channel allocation may be performed, based on these predictions, and data transmission may commence. This process is repeated at fixed re-training intervals with the trained model parameters from each previous set of observations taking the place of the current initial guess for $\boldsymbol{\lambda}$.

Channel allocations are based on a future prediction length interval N_{RT}, which is a measure of how many samples of PU activity will be considered when determining which channel would be best for SU communication. This is further discussed in the description of the simulation platform.

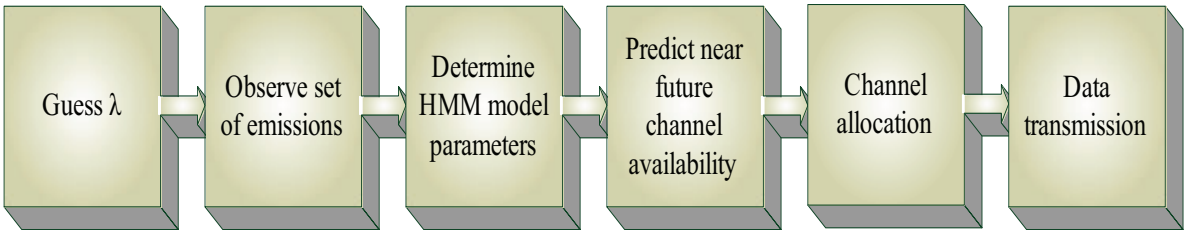

Figure 4.1: OSA system model.

4.2.1 Proactive switching and prediction length

So as to adhere to the requirements that OSA places on PU interference, it is essential that the method employed to perform SU channel switching adheres to the description of intelligent PSW described in Chapter 2. By allowing a SU to simultaneously perform SS and apply the prediction model across multiple adjacent PU channels, a list of predicted future PU channel occupancy states may be generated. From this list, the channels within which the SU should be operating at the next time instance, may be intelligently chosen. If a SU is currently operating at time t, then it should not only select any available channel at time $t+1$, but rather intelligently select the best channel within which to operate. In [41] it is suggested that the channel with the longest expected remaining idle time be selected for this purpose.

4.2.2 Effect of primary user traffic density

Channel switching performance is heavily influenced by the traffic density of the SU's operating environment. Intuitively one may assume that the higher the density of users vying for access to a particular channel, the less likely it becomes that a SU will readily find a good channel within which to operate. Consequently, a shortage of good channels within which to operate could have a negative effect on SU performance. In this investigation it was assumed that PU traffic density follows either a heavy or a light traffic density profile. In essence, for any given point in time a heavy traffic density profile may be defined to mean that it is more probable for the PU to be occupying the band of interest, than the probability that it is not. The converse is true for light traffic density. If D_L is used to denote a light traffic density and D_H a heavy traffic density, the following hypothesis may thus be defined,

$$\begin{aligned} D_L &: P_r(\boldsymbol{Y}=0) \geq P_r(\boldsymbol{Y}=1), \\ D_H &: P_r(\boldsymbol{Y}=1) > P_r(\boldsymbol{Y}=0). \end{aligned} \tag{4.1}$$

4.3 SIMULATION PLATFORM

A software simulation platform was developed to investigate the effects that channel switching may have on the performance of SUs while operating within a CR network. The channel

occupancy model discussed in Chapter 3 and the OSA model described in the first part of this chapter, form the basis of the simulation platform.

4.3.1 Physical layer considerations

The simulation platform was set up to operate under various physical layer parameters that are based on the IEEE 802.22 wireless regional network standard for CR [60]. The standard specifies a number of adaptive modulation Θ and coding rate r modes. Mode 3, described in Table 2.4, was chosen for the results gained from this simulation platform (presented in Chapter 6). This mode specifies a spectral efficiency of 0.76 b/s/hz and incorporates quadrature phase-shift keying (QPSK) with a forward error correction (FEC) code of rate $d = 1/2$. Guidelines pertaining to CR frame structure are also provided for by the media access control (MAC) layer specification of the standard. For a $B = 6$ MHz channel bandwidth, the standard suggests that data should be transmitted in $t_{sf} = 16t_f$ sized superframes, where $t_f = 10$ ms is the frame period. For the purpose of channel switching analysis, another MAC structure was added: the superblock. A superblock consists of $t_{sb} = 32t_{sf}$ superframes. The new MAC structure is illustrated in Figure 4.2 (modified from [60]).

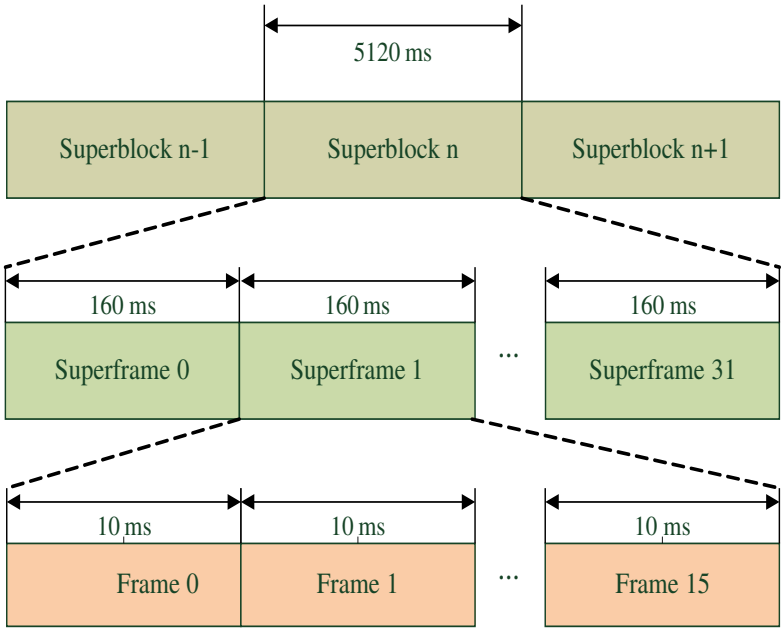

Figure 4.2: Frame, superframe and superblock MAC structure, adapted from [60].

4.3.2 Channel switching algorithm

A software flow diagram summarising the channel switching algorithm central to the simulation platform, is presented in Figure 4.3. This algorithm takes into consideration all the OSA requirements outlined in Chapter 4.2. The simulation platform was built around this algorithm and provides a set of logical instructions that govern the proposed way in which a SU will be allocated channels within in a CR radio network. A step-by-step description of the algorithm now follows.

To begin with, a SU must perform wide-band SS to determine the current channel occupancies of a set of potential narrow-band operating channels $\boldsymbol{I}_q$. During this process, occupancy data covering L frame periods is collected such that,

$$\boldsymbol{I}_q = \boldsymbol{O}_q(t), \quad 1 \leq q \leq \vartheta, 1 \leq t \leq L, \tag{4.2}$$

where ϑ denotes the size of the actual channel set. This set of channel occupancy data $\boldsymbol{I}_q$ is then used to train a HMM for each channel q in the set (as described in Chapter 3).

Once a model has been generated for each channel of interest, near future channel occupancy predictions for the entire set of channels $\boldsymbol{F}_q$ may be calculated in accordance with the following expression,

$$\boldsymbol{F}_q = \boldsymbol{X}(t), \quad t_L \leq t \leq \Psi. \tag{4.3}$$

The frame period commencing immediately after model training has been performed is denoted by t_L and Ψ represents the maximum number of states that will be predicted by the model. The foundation upon which a SU will determine when and within which channels to operate is provided by $\boldsymbol{F}_q$.

A SU should strive to avoid PU interference, therefore before SU data transmission actually commences, the SU must again sense the currently selected operating channel q_c for a time $t_{sn} \leq t_f$. This sensing operation will henceforth be referred to as a quick sensing operation. Quick sensing is performed to obtain immediate and up-to-date information

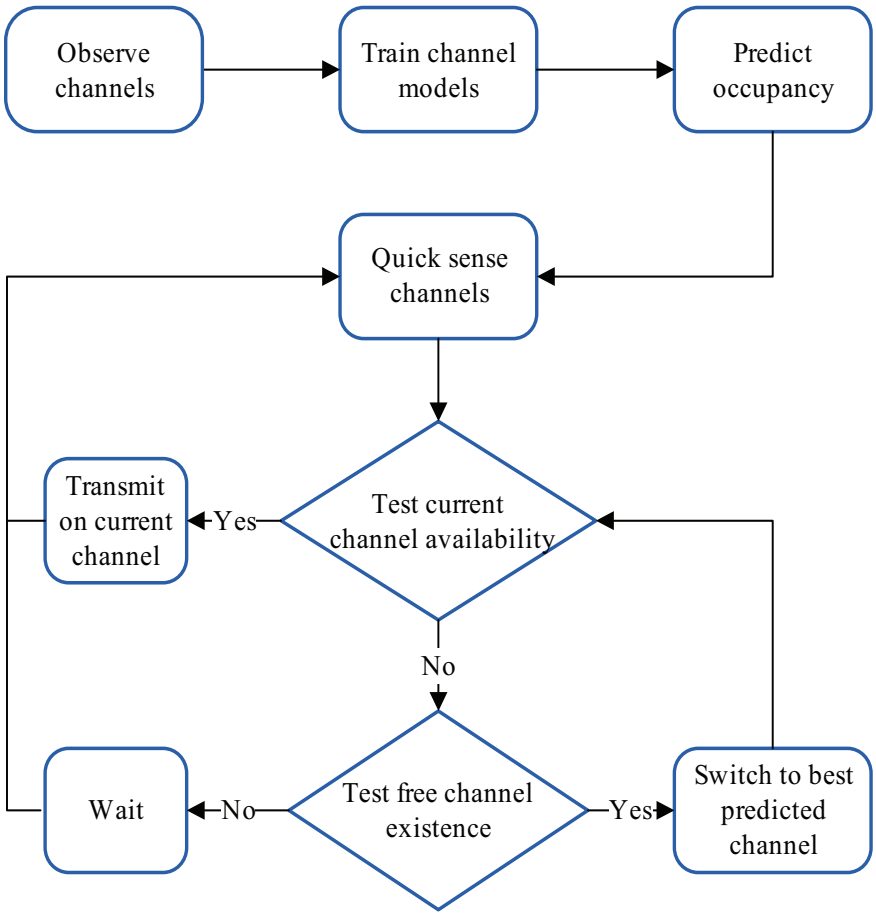

Figure 4.3: Software flow diagram illustrating the algorithm central to the simulation platform.

about PU activity. This information allows the SU to adapt to unexpected changes to the occupancy profile of the band. It is important to take note of the difference between this operation, where only the band of interest q_c is sensed, and the initial wide-band sensing operation that covers the entire range of potential operating channels $\boldsymbol{I}_q$. Initially, q_c is picked at random, but thereafter $\boldsymbol{F}_q$ will be employed for this purpose.

The next step is to use the information gained from the quick sensing operation to determine the occupancy status of q_c at time t_c. If q_c is found to be currently available at time t_c, $\boldsymbol{O}_q(t_c) = 0$, then the SU may commence data transmission on q_c for a time interval t_I. The length of this interval is governed by the selected SS approach (this will be discussed in section 4.3.3).

However, if q_c is found to be currently occupied, $\boldsymbol{O}_q(t_c) = 1$, the SU will try to find an alternative channel to which it can pro-actively switch. This channel is chosen by selecting what the SU deems to be the channel q that will remain unoccupied for the longest expected number of frame periods Ω_q (based on the set of future channel occupancy predictions $\boldsymbol{F}_q$) such that,

$$\arg\max_q \Omega_q, \quad t_c \leq \Omega_q t_f \leq t_c + \rho t_f, \tag{4.4}$$

where ρ is the maximum number of future frame periods that the SU will consider when calculating Ω_q. However, if all the channels in the set are currently occupied, $\boldsymbol{I}_q = 1, 1 \leq q \leq \vartheta$, then the SU will wait until $t = t_c + 1$ before performing another quick sensing operation and testing for an alternative operating channel.

The SU will continue the cycle of quick sensing, testing channel availability, channel switching and data transmission until the prediction period is exceeded, i.e. until $t = t_L + \Psi$. At this point the set of channel occupancy models $\boldsymbol{I}_q$ will be updated and $\boldsymbol{F}_q$ will be repopulated with a set of new channel occupancy predictions, thus continuing the SU channel switching and data transmission cycle.

4.3.3 Spectrum sensing approaches

SS is performed to verify the occupancy status of the channel within which a SU is currently operating. This is meant to limit PU interference, however, every sensing operation does introduce added delays that could affect SU performance. The SS approach determines how regularly the quick sensing operation is performed. To illustrate the effect that the quick sensing operation has on CR performance, four different approaches to the sensing of generic PU activity, based on the periodic sensing concept presented in [73], are proposed. These sensing approaches are categorised as follows: aggressive, bi-frame, quad-frame and prediction dependent. These four approaches are now briefly described and a graphical illustration is provided in Figure 4.4.

1. Aggressive: quick sensing is performed after every single frame period and should

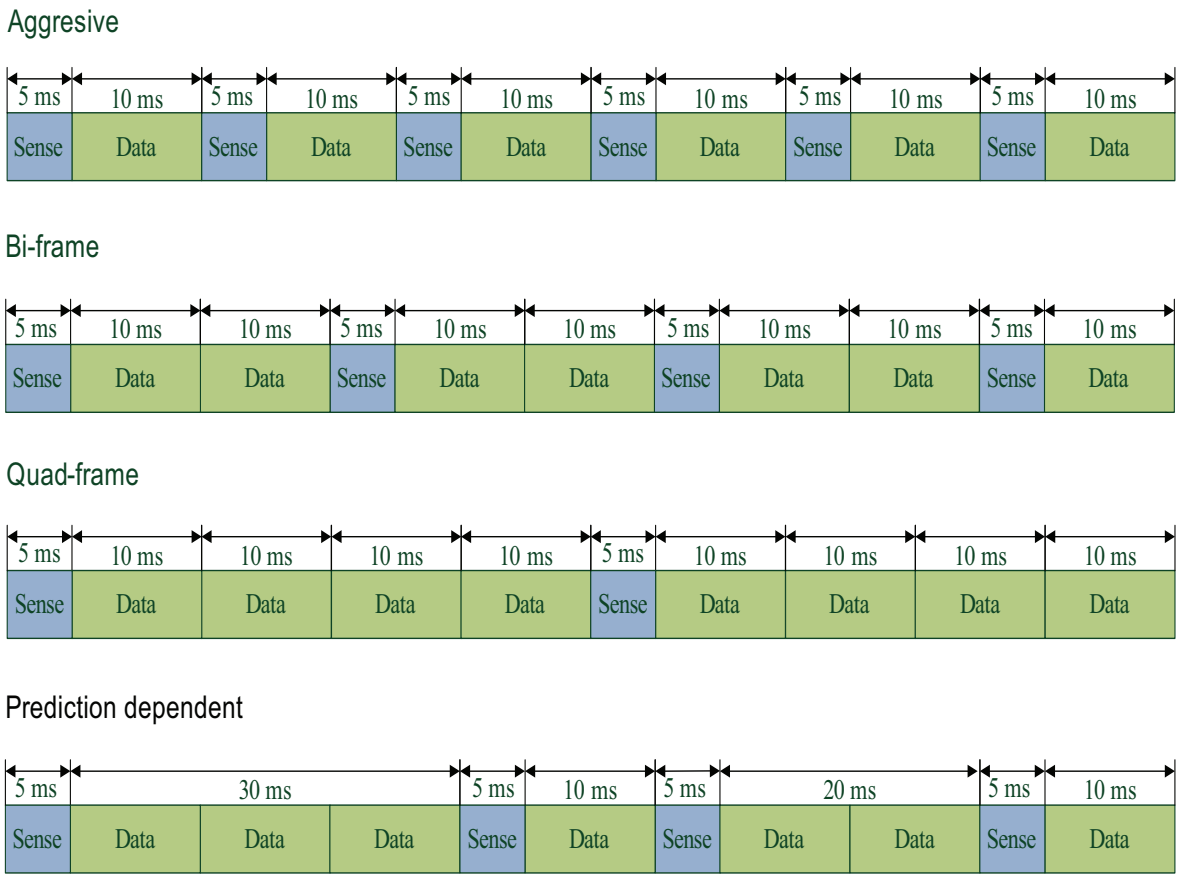

Figure 4.4: Graphical illustration of the differences between the quick sensing approaches, assuming $t_{sn} = t_f/2$.

provide maximum protection against PU disruption. The transmission interval t_I is thus given as,

$$t_I = t_f. \tag{4.5}$$

2. Bi-frame: quick sensing is performed after every second frame period. Therefore the transmission interval t_I is given as,

$$t_I = 2t_f. \tag{4.6}$$

3. Quad-frame: quick sensing is performed after every fourth frame period. While the quad-frame sensing approach provides more time for SU communication than the previous methods, it is likely to provide the least amount of protection against PU disrup-

tion. The transmission interval t_I is given as,

$$t_I = 4t_f. \tag{4.7}$$

4. Prediction dependent: quick sensing is only performed after Ω_q frame periods and relies on a high level of channel occupancy prediction accuracy. The prediction dependent approach has the potential to provide a balance between allowing for SU communication and providing protection against PU disruption. The transmission interval t_I for the prediction dependent approach is thus given as,

$$t_I = \Omega_q t_f. \tag{4.8}$$

4.4 SECONDARY USER PERFORMANCE

In this section certain criteria are discussed for evaluating the performance of a SU when the channel occupancy and OSA models are combined. For the simulation platform, occupancy model performance, SU data throughput, PU disruption and SU power consumption are considered.

4.4.1 Channel occupancy model performance

The only way that a SU will be able to perform channel switching is if it has up to date PU occupancy knowledge of all the channels that it may potentially be operating within. Channel occupancy prediction is thus an essential prerequisite to proactive channel switching. Amongst other factors, the accuracy with which future PU behaviour may be predicted, will have a significant effect on the achievable performance of a SU in a CR network. The more accurately that channel occupancy states are able to be predicted, the more accurately channel switching may be performed. Poor prediction accuracy may cause a SU to experience extra delays due to additional sensing and channel switching requirements and PUs may also be more likely to experience SU interference.

4.4.2 Secondary user data throughput

To help measure the impact that channel modelling may have on SU performance, the effect that channel switching accuracy has on the throughput of a single SU, is calculated. Certain factors that are specific to a cognitive radio network, will affect the actual throughput achieved. These are over and above traditional factors such as modulation scheme, coding rate and imperfect channel conditions.

Any delays introduced into a communications network, including those that are a result of CR-specific functions, will have a negative impact on the achievable throughput of the network. Unfortunately, due to the inherent nature of CR, various CR-specific delays are inevitable. Amongst others, these delays may include,

- sensing delays d_{sn},
- waiting delays d_{wt},
- switching delays d_{sw},
- observation delays d_{ob} and
- training delays d_{tr}.

These delays have been incorporated into the simulation platform described in the previous section, the cause of which may be attributed to the following CR specific phenomena,

- the number of SU quick sensing operations υ_{sn} performed,
- the number of times a SU has to wait for any free channel to become available υ_{wt},
- the number of times a SU has to switch channels υ_{sw},
- the HMM training algorithm observation length L and

- the number of iterations performed during the HMM training process υ_{tr}.

CR-specific delays may be calculated using these phenomena. The delay incurred by performing the quick sensing operation d_{sn}, with sensing time t_{sn}, may thus be given as,

$$d_{sn} = \upsilon_{sn} t_{sn}, \tag{4.9}$$

the delay incurred when waiting for an available channel d_{wt} may be given as,

$$d_{wt} = \upsilon_{wt} t_I t_f, \tag{4.10}$$

and the delay incurred by performing channel switching d_{sw} may be given as,

$$d_{sw} = \upsilon_{sw} t_{sw}, \tag{4.11}$$

where $t_{sw} \leq t_f$ denotes the time required to perform a channel switch. The delay incurred when gathering channel occupancy observations d_{ob} may be given as,

$$d_{ob} = t_f L, \tag{4.12}$$

and the the delay incurred by training the HMM occupancy models d_{tr} may be given as,

$$d_{tr} = \upsilon_{tr} t_{tr}, \tag{4.13}$$

where $t_{tr} << t_f$ is the average time required to perform a single iteration of the HMM training algorithm.

If data throughput R_b may be defined to be the rate at which data is received, then it may be calculated as a function of the number of bits k received over a fixed period of time T_s. Data throughput R_b may thus be given as follows[17],

$$R_b = \frac{k}{T_s}. \tag{4.14}$$

However, the theoretically achievable throughput based on CR factors only R_{su}, i.e., the effects of modulation, coding and channel gain are ignored, may thus be expressed as,

$$R_{su} = \frac{k}{T_{cr}}, \tag{4.15}$$

where T_{cr} represents the transmission time period with the extra time delay introduced by the CR process included. It is thus defined as follows,

$$T_{cr} = T_s + d_{sn} + d_{wt} + d_{sw} + d_{ob} + d_{tr}. \tag{4.16}$$

4.4.3 Primary user disruption rate

When a SU erroneously switches to a channel that is already occupied by a PU, it may cause disruption and interference to the PU. This will occur when a SU switches to a channel that has either been incorrectly sensed and/or predicted to be unoccupied, at a time when it is in actual fact currently occupied by a PU. If the occurrence of such an event over a period of time t, is denoted by δ_t, then the number of such events I_{cr} that occur during the time period T_{cr}, calculated as follows,

$$I_{cr} = \sum_{t=0}^{T_{cr}} \delta_t \tag{4.17}$$

may be used to calculate the rate D_{pu} at which a SU causes disruptions to PUs. The following expression may be employed for this purpose,

$$D_{pu} = \frac{I_{cr}}{T_{cr}}. \tag{4.18}$$

4.4.4 Secondary user power consumption

Intuitively, it would seem that the potential exists for CR-specific functionality to increase the power requirements experienced by a SU. This may become a problem if the SU is a mobile device, since the battery life of the device would inevitably be affected. The same CR-specific factors that negatively impact upon SU throughput will inevitably also affect SU power consumption. To illustrate this concept, fixed power requirements for performing SS, channel switching and channel occupancy prediction, as well as the power wasted in an idle state due to unnecessary waiting when no free channels exist, have been assigned for use in the simulation platform. These power requirements will be referred to as CR power penalty rates and may include,

- a sensing power penalty rate p_{sn},
- a waiting power penalty rate p_{wt},
- a switching power penalty rate p_{sw}, and

- a training and prediction power penalty rate p_{tp}.

If P_{TF} is defined to be the power required to transmit a single uninterrupted frame of data under standard operating conditions, then the power required to transmit the same amount of data CR-specific conditions P_{su}, may be calculated from the following expression,

$$P_{su} = P_{TF} + (p_{sn}d_{sn}) + (p_{wt}d_{wt}) + (p_{sw}d_{sw}) + p_{tp}(d_{ob} + d_{tr}). \tag{4.19}$$

4.5 CONCLUDING REMARKS

In this chapter, a stand-alone simulation platform, based on the concept of OSA, was presented. The simulation platform made use of the channel occupancy model described in Chapter 3 to investigate its effect on SU performance in a CR radio network. A detailed description of the simulation platform, including its physical layer properties, the channel switching algorithm and the approach to SS, was presented. A description of the CR performance measures investigated by the platform, which included SU throughput, PU disruption rate and SU power consumption was also provided. Performance results obtained from simulations run on the simulation platform, will be presented in Chapter 6.

CHAPTER 5

SPECTRUM OCCUPANCY MEASUREMENTS

5.1 INTRODUCTION

While theoretical simulations are useful for analysing the preliminary performance of the PU spectrum occupancy prediction model, a more complete analysis may be obtained by testing the model on actual spectrum measurements. Typical channel occupancy data was thus collected for this purpose as part of a spectrum measurement campaign carried out on the main campus of the University of Pretoria. Certain portions of the spectrum measurements, in the commercial cellular bands, were selected for model testing.

However, the purpose of the measurement campaign was to also gather information about typical local spectrum occupancy. Although, various measurement campaigns have been undertaken in other countries [9, 11, 16, 74, 75, 76, 77, 2], there is still a general lack of knowledge regarding spectrum occupancy in South Africa. A mobile system, which can be used to gather spectrum occupancy information from a range of frequencies over long periods of time, was thus developed. This system included a modular hardware system and software environment aimed at delivering detailed information about the occupancy of various South African commercial frequency bands. Data was collected during a six-week-long measurement campaign at the University of Pretoria, South Africa.

Firstly, the proposed method for calculating spectrum occupancy from raw spectrum measurements, is discussed. The frequency bands selected for analysis are then presented, followed by a description of the major functional and physical aspects of the measurement system. The measurements of the bands selected for model testing, as well as their calculated spectrum occupancies, are then presented and discussed.

5.2 OCCUPANCY THRESHOLD

Since the measured data is in the form of actual instantaneous signal power, a certain amount of post processing needs to be performed before spectrum occupancy may be determined. The following method for calculating the signal detection threshold λ_T, required for testing the binary hypothesis described in Equation 2.1, is thus proposed.

To begin with, the raw received signal $r(n)$ needs to be transformed into a more detector friendly form. This may be accomplished by raising $r(n)$ to a power δ as follows,

$$Y = |r(n)|^{\delta}, \quad 0 \leq n \leq N. \tag{5.1}$$

The value chosen for δ determines the extent to which the noise $w(n)$ and smaller signal components of $r(n)$ are suppressed (N is the observation vector size). Once the noise components $w(n)$ have been suppressed, Y is clipped by the sum of its mean μ_s and standard deviation σ_s. The clipped signal may thus be expressed as follows,

$$Y_c = \begin{cases} Y, & |Y| \leq \mu_s + \sigma_s \\ \mu_s + \sigma_s, & |Y| > \mu_s + \sigma_s. \end{cases} \tag{5.2}$$

The signal detection threshold λ_T may now be calculated by adding the mean of the clipped signal μ_c to a preselected constant γ, as follows,

$$\lambda_T = \mu_c + \gamma. \tag{5.3}$$

Once λ_T has been calculated, a measured channel occupancy vector $C(n)$ may then be populated in accordance with the following expression,

$$C(n) = \begin{cases} H_0, & |Y| < \lambda_T \\ H_1, & |Y| \geq \lambda_T. \end{cases} \tag{5.4}$$

The contents of $C(n)$ have been used as practically measured values for $\boldsymbol{O}_q$ in the simulation platform described in Chapter 4.

The design choices for δ and γ will have an effect on the calculated channel occupancy accuracy, since these parameters have a direct influence on λ_T. They thus need to be selected appropriately, so as to the minimize the probabilities associated with incorrectly detecting PU activity, namely the probabilities of miss-detection P_{md} and false alarm P_{fa}. The former is defined to be the probability of detecting a band to be free when in reality it is actually occupied and may thus be described as,

$$P_{md} = P_r\{Y > K_{Th}|H_0\}, \tag{5.5}$$

while the latter is defined to be the probability of detecting PU activity when the channel is actually free and may be described as,

$$P_{fa} = P_r\{Y < K_{Th}|H_1|\}. \tag{5.6}$$

The importance of correct threshold calculation and selection is illustrated in Figure 5.1, where P_{fa} , P_{md} and the resulting percentage accuracy of PU detection are compared. A clear trade-off exists between these two probabilities and the highest level of PU detection accuracy is achieved when these two probabilities intersect. The trends illustrated in Figure 5.1, have been generated from a set of simulated test data. The average channel occupancy of the simulated test data is 41.6% (calculated from known transmitter information as the percentage of discrete measurement points, within the band of interest, that occur above the detection threshold λ_T) and P_{fa} and P_{md} have been estimated by counting the rate at which these events occur over a test sequence of 2000 states. The calculated detection threshold, when $\delta = 1.73$ (the intersection of P_{fa} and P_{md} in this scenario) and $\gamma = 0$, is also visually illustrated.

5.3 BANDS OF INTEREST

In Chapter 6, SU performance will be investigated using selected measured data from the South African GSM cellular bands. Measured channel occupancy vectors $C(n)$ from two

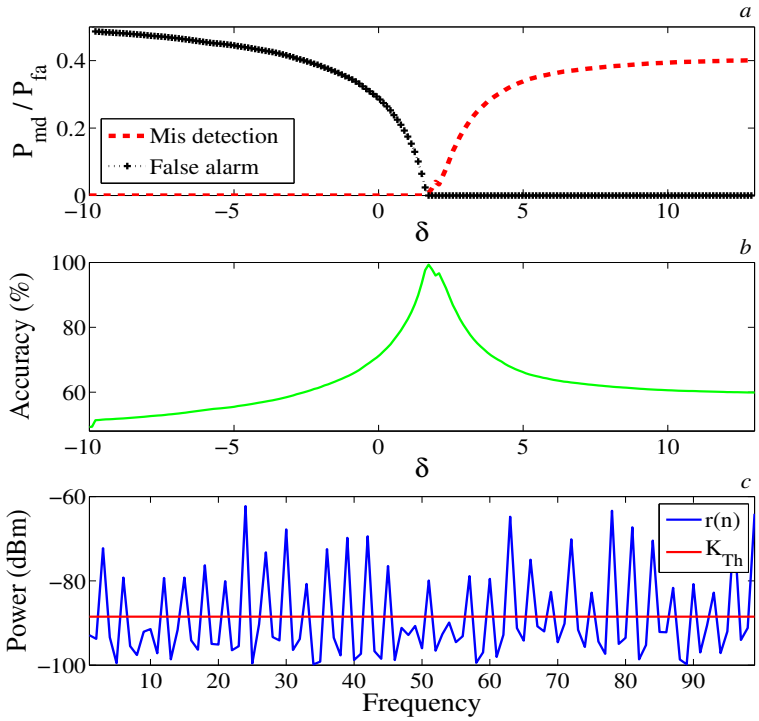

Figure 5.1: Simulated effect of δ on P_{fa}, P_{md} (part *a*) and the resulting accuracy of channel occupancy detection (part *b*). An illustration of K_{Th} when $\delta = 1.73$ is presented in part *c*.

specific bands have been selected for analysis. The bands, described in Table 5.2, are denoted as: frequency band A (890.1 to 895.1 MHz) and frequency band B (1848.5 to 1853.5 MHz). Portions of the GSM cellular bands were selected since they exhibit conditions where PU occupancy fluctuates more rapidly than in the other bands measured, e.g., additional measurements in the South African UHF bands showed very low occupancy and little variation in PU activity. Frequency bands A and B were thus specifically chosen as an attempt at representing both high and low traffic density conditions that would be appropriate for CR usage. Data from these bands will be used to test the occupancy model and also provide inputs to the simulation platform (described in Chapter 4.3). Measurements were thus taken from 5 MHz portions of the GSM 900 up-link (band A) as well as the GSM 1800 down-link (band B). A total of 150 s worth of data was collected for each measurement.

5.4 MEASUREMENT SYSTEM

The system used to take spectrum measurements [1], based on the design in [15], employs a wide band antenna that is connected to a spectrum analyser (SA) via a low noise amplifier (LNA). Operation of the system is controlled by an automated software application that interfaces with the SA over a remote Ethernet connection. The antenna, LNA and SA are housed within an air-conditioned metal cabinet at the measurement site. A more detailed description is now provided.

5.4.1 Functional description

A functional description of the measurement system is illustrated in Figure 5.2 in the form of a functional block diagram. FU H1 is a wide-band Super-M Ultra Base Station antenna [78] with a frequency range of 25 MHz to 6 GHz, capable of receiving both vertically and horizontally polarized signals.

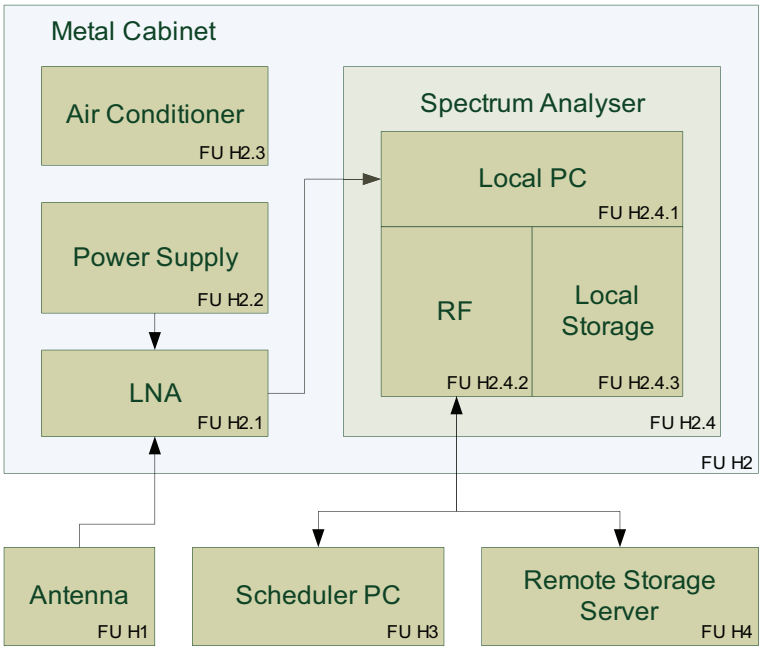

Figure 5.2: Hardware component functional block diagram of the SS system.

[1] The spectrum measurement system was developed as part of a joint post-graduate student initiative of the Sentech BWMC research group at the University of Pretoria.

The main hardware components of the receiver system were housed within a metal cabinet, FU H2, located on the roof of the Engineering I building of the University of Pretoria. This cabinet housed FU H2.1, a low-noise amplifier (LNA) with an operating range of 50 MHz to 3 GHz [79], FU H2.2, a regulated 5V DC power supply for powering the LNA, FU H2.3, a custom-built air conditioning unit (since the SA had to operate within a closed environment, the cabinet needed to be sufficiently air conditioned to protect against overheating) and a SA, FU H2.4.

The SA, FU H2.4, although physically housed within the same enclosure, was split into two functional hardware units: a local personal computer (PC) that houses device-controlling software, FU H2.4.1, and the radio frequency (RF) component of the device, FU H2.4.2. Temporary data storage space was available on the local PC, FU2.4.3. FU H1 was connected to FU H2.1 via 10 metres of rugged low-loss LMR 600 coaxial cable. FU H2.1 was in turn connected to FU H2.4.2 via a Sucoflex 100 coaxial cable. An Ethernet connection provided the interface between FU H2.4 and two other functional hardware components: FU H3, a remote personal computer (PC) that ran the remote scheduling software and FU H4, a backup and storage server for remote storage and backup of the spectrum measurement data. FU H3 and FU H4 were remotely located in an office in the Engineering II building of the University of Pretoria.

The hardware components of the measurement system were controlled by three separate software applications. These included applications to,

- interface with and locally control the SA, manage data files and provide operational status reports (installed on FU H2.4.1),
- provide secure storage of measurement result files on the backup server (installed on FU H2.4.1) and
- remotely schedule measurements and allow remote access to the system configuration function (installed on FU H3).

5.4.2 System calibration and sensitivity

The major functional components of the measurement system were all calibrated using an Agilent E5071C ENA series network analyser. The complete system link budget is illustrated in Figure 5.3, and includes the calibrated gains for the antenna G_a, the coaxial cables (LMR 600, G_{cl}, and Sucoflex 100, G_{cs}) and the LNA G_{lna}. The total calibrated gain of the system G_{tot}, illustrated by the red line in Figure 5.3, is a linear combination of the gain curves of all the major physical elements of the system and may thus be calculated as follows,

$$G_{tot} = G_a + G_{cl} + G_{cs} + G_{lna}. \tag{5.7}$$

The receiver sensitivity of the system S_r was also calculated to assist in determining the detection threshold λ_T. This was done by subtracting G_{tot} from the received signal $r(n)$ as follows,

$$S_r = r(n) - G_{tot}. \tag{5.8}$$

In Table 5.1, the measured sensitivity of the system is compared for both the GSM 900 (-103.4 dBm) and GSM 1800 (-102.3 dBm) bands respectively.

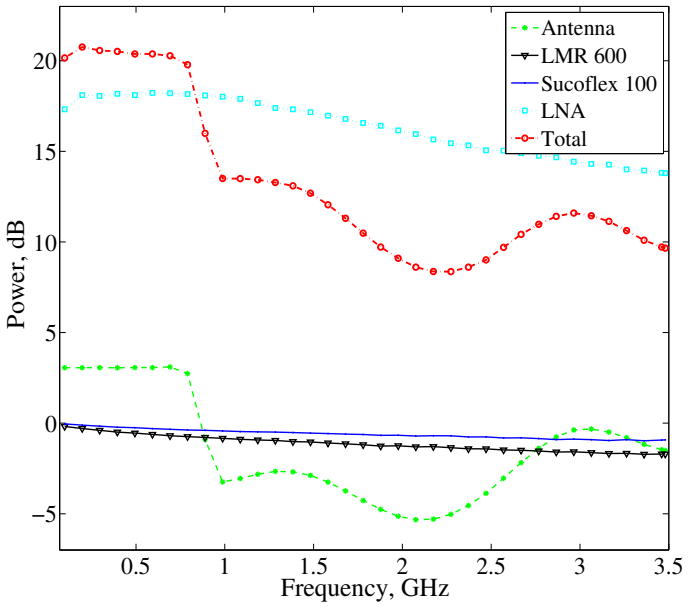

Figure 5.3: Link budget of the hardware setup.

Table 5.1: Spectrum measurement system sensitivity after calibration.

	GSM 900	**GSM 1800**
Sensitivity (dBm)	-103.4	-102.3

5.4.3 Measurement description

Measurements, spaced at a channel frequency interval of 100 kHz, were taken from a 5 MHz block of spectrum using the energy detection method. Each measured channel consisted of a sweep of 1500 discrete time samples that were spaced 100 ms apart from each other. However, these measurements were truncated to 512 time samples for the sake of consistency between the results derived from theoretical and practical occupancy data. Thus, each band was represented by a frequency-time matrix of 50 frequency bins of 512 time samples each (5 MHz x 51.2 s). Measurements were taken every two hours during a measurement campaign that covered a period of six weeks during May and June 2011 (the length and regularity of these measurements were restricted due to hardware and storage capacity limitations).

5.4.4 Physical installation and measurement site

The physical installation of the spectrum measurement system is illustrated in Figure 5.4 and includes the Super-M Ultra base antenna FU H1 (top left), the metal cabinet, FU H2 (bottom), the scheduler PC FU H3 (top middle) and the remote storage server FU H4 (top far right). The antenna was mounted towards the northern end of the roof of the Engineering I building at the University of Pretoria. The metal cabinet was also located on the roof, a few meters away from the antenna. The scheduler PC and remote storage server were both housed in an office in the Engineering II building.

An aerial photograph indicating the exact measurement site and its immediate surroundings [80], denoted by a yellow arrow, is provided in Figure 5.5. The measurement site, located on the Hatfield campus of the University of Pretoria, South Africa, is the second highest point in the immediate surrounding area (after the Human Sciences building).

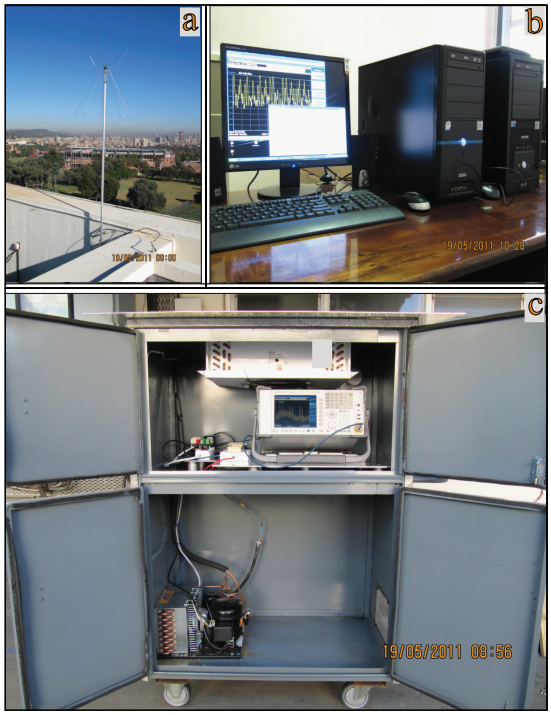

Figure 5.4: Photograph of the physical installation of the spectrum measurement system. Part *a* is the antenna, part *b* the storage and scheduler PCs and part *c* is the metal cabinet.

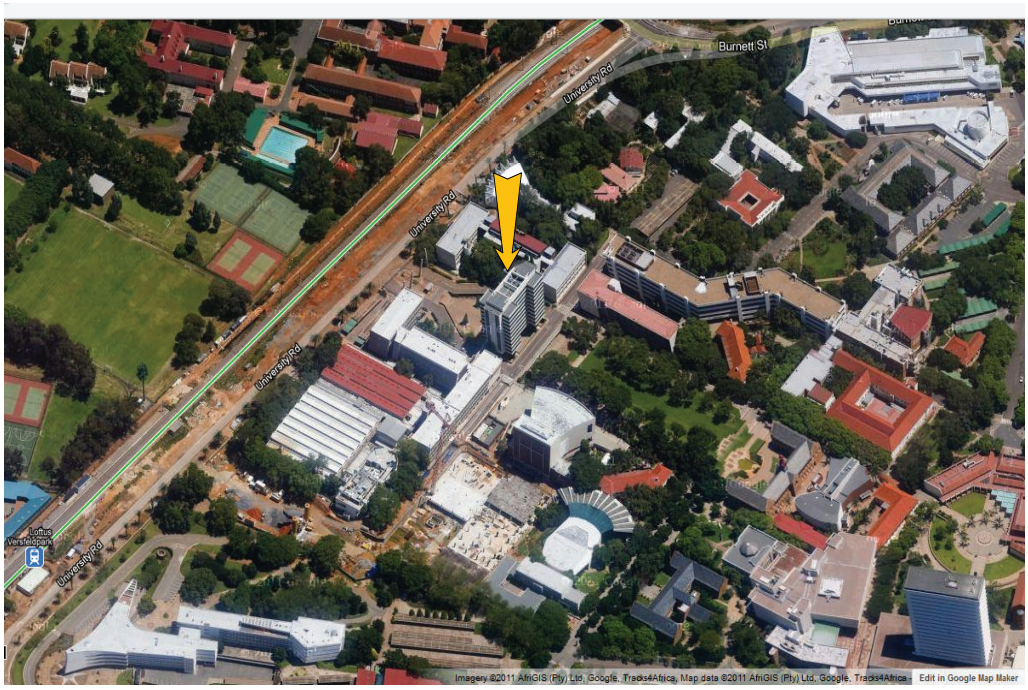

Figure 5.5: Aerial photograph of the spectrum measurement site, adapted from [79].

5.5 MEASURED CHANNEL OCCUPANCY

The measured average frequency-time power spectra of the bands presented in Table 5.2, are illustrated in Figure 5.6 and Figure 5.7 respectively. These plots represent the raw signal powers of the measured bands, averaged over the six week measurement period, that were used to validate the theoretical results obtained from the simulation platform.

Detection thresholds for these bands were calculated according to the procedure described in Section 5.2. The values chosen for δ in Equation 5.1 were selected by a process of visual inspection. Separate values for δ were selected for each band due to differences in the frequency-power profile of each band (this was necessitated by the reliance of the threshold detection method on calculating mean received power levels). Threshold exponent values of $\delta = 5.0$ and $\delta = 0.1$ were chosen for frequency band A and B respectively.

The chosen exponent values and calculated detection thresholds, together with the resulting percentage channel occupancies of each band, are listed in Table 5.2. The calculated channel occupancy for both bands is also visually illustrated in Figure 5.8 in the form of a black and white frequency-time binary occupancy map. The white areas represent PU occupancy and the black areas the absence thereof.

The selected bands clearly exhibit rapidly changing PU activity across both time and frequency. The calculated average percentage occupancy of frequency band A is 20.47% and for band B 84.38%. It is thus assumed that data from band A may be used to represent practically measured low density traffic conditions and the data from band B practically measured high density traffic conditions.

Table 5.2: Table of measured frequency bands and their associated properties.

Freq. band	Freq. (MHz)	δ	Thres. (dBm)	Occup. (%)
A	890.1 - 895.1	5.0	-91.49	20.47
B	1848.5 - 1853.5	0.1	-93.50	84.38

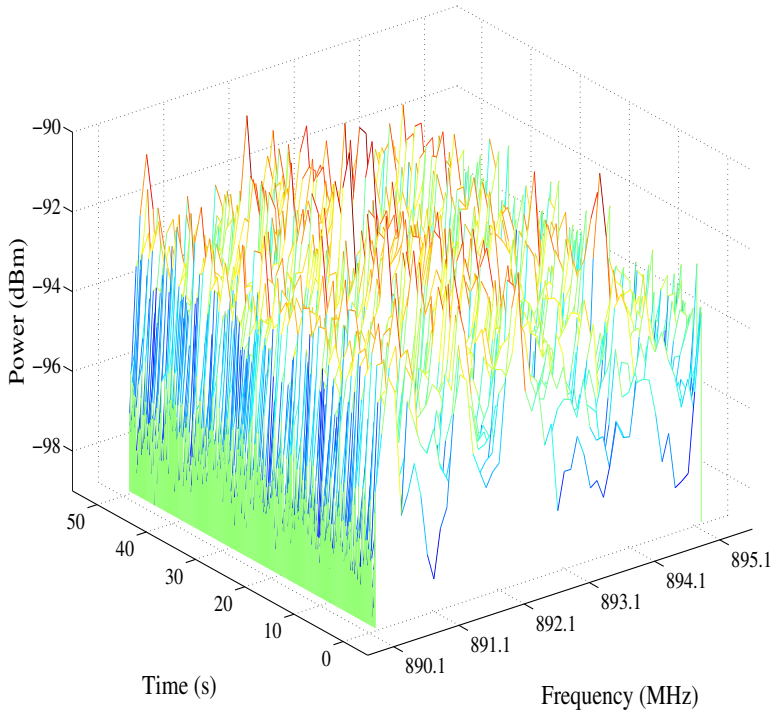

Figure 5.6: Measured power spectra of frequency band A.

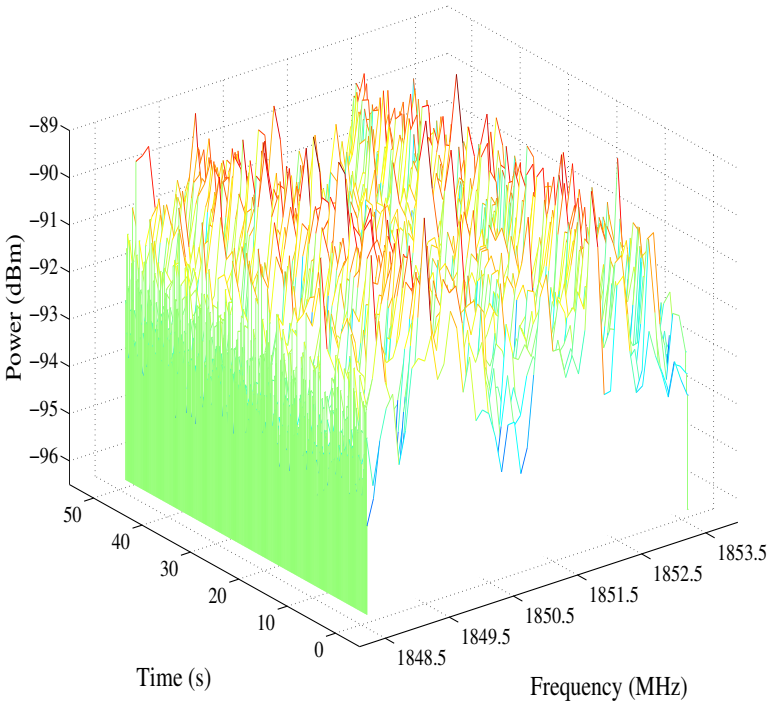

Figure 5.7: Measured power spectra of frequency band B.

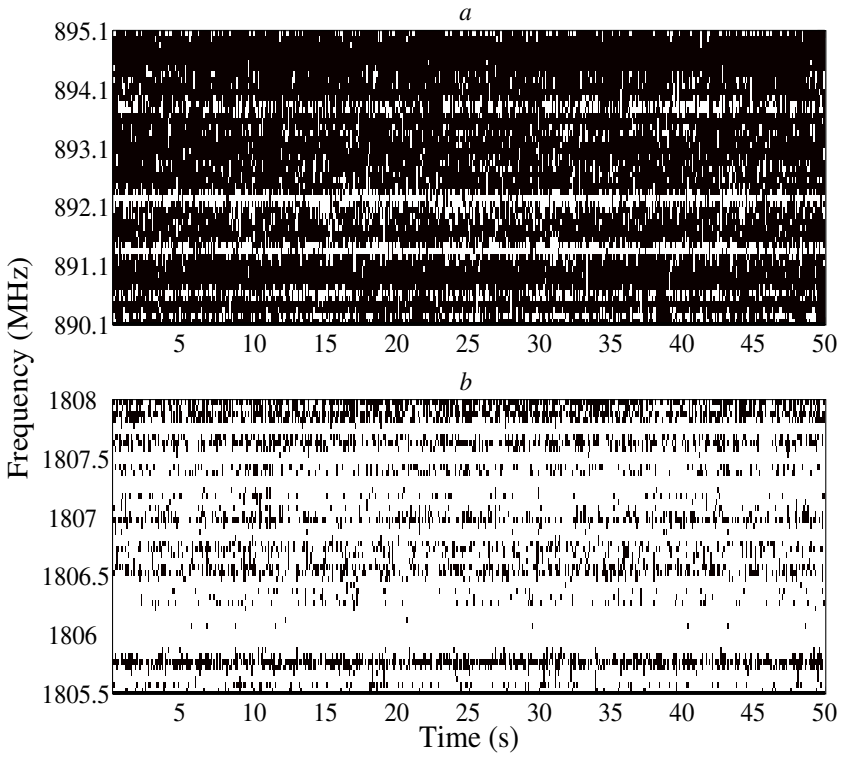

Figure 5.8: Binary occupancy plots for frequency bands A (part *a*) and B (part *b*).

5.6 CONCLUDING REMARKS

Collecting real spectrum measurements makes it possible for a more complete analysis to be performed of the spectrum occupancy, prediction and OSA models presented in Chapter 3 and Chapter 4. A description of the chosen frequency bands and the measurement system used to gather data from them, was provided in this chapter. A technique for determining spectrum occupancy from raw signal power measurements, was also presented.

The results obtained when using the measured spectrum occupancy data as an input to the simulation platform, will be presented and analysed in Chapter 6.

CHAPTER 6

RESULTS

6.1 INTRODUCTION

In this chapter a performance evaluation is presented for both the channel occupancy model and the OSA model described in Chapter 3 and Chapter 4 respectively. The simulation platform presented in Chapter 4, was used to investigate the effect that the traffic occupancy model might have on the performance of a single SU within a CR network. Simulation results are compared for both simulated and practically measured channel occupancy under both light and heavy density traffic conditions.

In Section 6.2, the effects of using different training algorithms to train the HMM based occupancy model, described in Chapter 3, are compared. The relationship between occupancy model performance and training sequence length L, is also illustrated.

In Section 6.3, spectrum allocation performance results, obtained from the simulation platform described in Chapter 4, are presented. The effect of the channel occupancy model on SU channel switching is investigated both in terms of the model training algorithm and the training sequence length L.

The effect that the spectrum allocation results have on SU performance are then presented in Section 6.4. These results have been split into two categories: results obtained by varying

the training sequence length L and those obtained by varying the quick sensing interval t_I. Performance measures presented include: SU data throughput, PU disruption and SU power consumption. The effect of traffic density on SU performance is also considered.

A similarity comparison between the SU performance results obtained from theoretical occupancy data and those obtained from practically measured data is provided in Section 6.5.

6.2 OCCUPANCY MODEL PERFORMANCE

Performance results for the HMM based occupancy model are presented in Table 6.1 (trained by the VA), Table 6.2 (trained by the BWA), Table 6.3 (trained using PSO) and Table 6.4 (trained by the MA). Results obtained from both theoretical occupancy data and measured occupancy data are included and a measure of the difference between them is also provided (the similarity scores were calculated using Equation 6.3, which will be discussed in Section 6.5). The effect that the different training algorithms have on the performance of the occupancy model has been quantified in terms of the mean and standard deviation of prediction accuracy as well as algorithm complexity (as described by Table 3.2 and Table 3.3 in Chapter 3).

The performance of the model has been investigated for a range of training observation lengths L (200, 300 and 400) and it is evident that prediction accuracy proportionately improves as L is increased. Results generated from theoretical occupancy data indicate that using different training algorithms does not result in any significant differences in the accuracy of the channel occupancy and prediction model. However, results generated from practically measured occupancy data in frequency band A, indicate that the VA results in a lower mean prediction accuracy than the other three algorithms (particularly for $L = 200$ and $L = 300$). Nevertheless, for both frequency bands A and B, the evolutionary based training algorithms seem to exhibit a higher prediction accuracy standard deviation than the VA and BWA algorithms.

Table 6.1: HMM based occupancy prediction performance using the VA training algorithm.

	L	**Theoretical**	**Freq. band A**			**Freq. band B**		
Mean (%)	200	75.00	72.45	-2.56	96.59	82.54	7.54	89.95
	300	84.57	81.81	-2.76	96.73	86.60	2.03	97.60
	400	92.70	89.85	-2.85	96.92	92.79	0.09	99.91
Std (%)	200	11.31	1.56	-9.75	13.78	1.69	-9.62	14.93
	300	6.43	1.35	-5.08	20.95	1.96	-4.47	30.53
	400	3.82	0.52	-3.30	13.55	0.01	-3.81	0.21
KFLOPS	200	2.24	3.07	0.83	62.82	2.69	0.45	80.06
	300	2.24	8.48	6.24	-178.50	2.79	0.55	75.50
	400	2.24	3.16	0.92	58.83	2.89	0.65	71.00
			Meas.	**Diff.**	**Similar.**	**Meas.**	**Diff.**	**Similar.**

Table 6.2: HMM based occupancy prediction performance using the BWA training algorithm.

	L	**Theoretical**	**Freq. band A**			**Freq. band B**		
Mean (%)	200	76.49	81.99	5.50	92.81	85.72	9.23	87.94
	300	84.06	87.44	3.37	95.99	89.78	5.72	93.20
	400	92.51	92.57	0.07	99.93	92.59	0.08	99.91
Std (%)	200	8.62	1.78	-6.84	20.63	2.15	-6.47	24.90
	300	6.28	1.19	-5.09	18.93	1.01	-5.26	16.20
	400	3.33	0.57	-2.76	17.12	0.99	-2.34	29.71
KFLOPS	200	108.08	78.28	-29.80	72.43	102.41	-5.67	94.76
	300	118.00	80.32	-37.68	68.07	92.94	-2.51	78.77
	400	128.02	91.52	-36.50	71.49	91.57	-36.45	71.53
			Meas.	**Diff.**	**Similar.**	**Meas.**	**Diff.**	**Similar.**

Table 6.3: HMM based occupancy prediction performance using the PSO training algorithm.

	L	Theoretical	Freq. band A			Freq. band B		
Mean (%)	200	75.16	81.10	5.94	92.09	83.07	7.91	89.48
	300	83.13	86.50	3.38	95.93	90.50	7.37	91.13
	400	91.70	93.70	1.99	97.82	94.96	3.26	96.44
Std (%)	200	9.79	2.77	-7.02	28.31	3.49	-6.30	35.68
	300	6.76	1.40	-5.36	20.76	1.22	-5.54	18.07
	400	3.64	0.50	-3.14	13.78	1.06	-2.58	29.07
KFLOPS	200	364.18	363.19	-0.99	99.73	363.69	-0.50	99.86
	300	369.98	368.76	-1.22	99.67	369.00	-0.97	99.74
	400	372.05	369.62	-2.42	99.35	373.35	1.31	99.65
			Meas.	**Diff.**	**Similar.**	**Meas.**	**Diff.**	**Similar.**

Comparing the two evolutionary based algorithms, it would appear that the overall difference between them lies in the standard deviation of prediction accuracy, with PSO

Table 6.4: HMM based occupancy prediction performance using the MA training algorithm.

	L	Theoretical	Freq. band A			Freq. band B		
Mean (%)	200	75.62	94.40	18.78	75.17	83.50	7.88	89.58
	300	83.86	97.57	13.71	83.66	89.77	5.90	92.96
	400	91.82	99.00	7.18	92.18	94.65	2.84	96.91
Std (%)	200	10.54	10.15	-0.39	96.29	6.50	-4.04	61.71
	300	7.43	6.11	-1.33	82.14	3.15	-4.28	42.43
	400	3.77	2.87	-0.90	76.03	2.02	-1.75	53.54
KFLOPS	200	647.65	670.51	22.86	96.47	661.48	13.83	97.87
	300	656.08	698.40	42.32	93.55	673.65	17.57	97.32
	400	655.46	712 54	57.08	91.29	690.15	34.69	94.71
			Meas.	**Diff.**	**Similar.**	**Meas.**	**Diff.**	**Similar.**

clearly having a better standard deviation that the MA. The difference between the VA and the BWA is less pronounced, but the BWA does provide a marginal improvement in standard deviation.

However, the major distinguishing factor, between the results obtained from the different algorithms, lies in the complexity and number of iterations required by the prediction model to converge upon a good solution. While the VA may have a marginally worse prediction accuracy, it is by far the least complex of the algorithms, requiring approximately only three KFLOPS to converge upon a good solution. This is significantly lower than the roughly 100, 370 and 660 KFLOPS required of the BWA, PSO and MA algorithms respectively. Further results, indicating the number of iterations required for each training algorithm, are presented in Table A.1 in Appendix A.

From the tables it is evident that the results for all four training algorithms, both in terms of measured prediction accuracy and algorithm complexity, are either comparable with or show an improvement over the simulated results (especially for shorter lengths of L). The similarity between these results will be dealt with in more detail in Section 6.5.

6.3 SPECTRUM ALLOCATION PERFORMANCE

Results illustrating the effect that PU traffic prediction modelling has on the channel allocation process, are presented in this section. The number of channel switches υ_{sw} required by a SU, to communicate within a CR network, is compared against the size of the channel set ϑ available to them. The trends observed in this section form the basis for the CR performance results that will be presented in Section 6.4.

6.3.1 Simulation parameters and assumptions

Certain parameters and assumptions were adhered to when calculating the SU performance results described in the following sections. To begin with, two different traffic density scenarios were defined as described in Section 4.2.2. Chosen model parameter values, used

Table 6.5: Model parameters for simulating traffic density.

Density	$\boldsymbol{P}_{tr}$	$\boldsymbol{E}_{tr}$	**Occup. (%)**
Light traffic	$\begin{pmatrix} 0.70 & 0.30 \\ 0.80 & 0.20 \end{pmatrix}$	$\begin{pmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{pmatrix}$	75.78
Heavy traffic	$\begin{pmatrix} 0.30 & 0.70 \\ 0.20 & 0.80 \end{pmatrix}$	$\begin{pmatrix} 0.95 & 0.05 \\ 0.05 & 0.95 \end{pmatrix}$	31.84

to generate theoretical channel occupancy data $\boldsymbol{P}_{tr}$ and $\boldsymbol{E}_{tr}$, are presented in Table 6.5. The resulting percentage channel occupancies, for both traffic density scenarios, are also included.

The physical layer parameters listed in Table 6.6 were also adhered to and the observation and training window was limited to $L+\Psi = 512$ consecutive states. This means that as L was increased, Ψ decreased by a proportionate amount. However, due to the physical limitations of the SA on the speed at which spectrum measurements were able to be taken, the frame period had to be adjusted to be ten times longer than that specified in the IEEE 802.22 standard. A frame period of $t_f = 100$ ms was thus utilised. A channel bandwidth of $B = 100$ KHz was assumed and simulations were run for a length of one superblock t_{sb} for $S_x = 1024 t_{sb}$ iterations.

For determining SU channel switching and performance results, the VA was chosen to train the channel occupancy model. This decision was made since the VA exhibits significantly lower computational complexity than the other algorithms (as reported in Section 6.2).

Table 6.6: Physical layer parameters for investigating SU performance.

Param.	Θ	d	B	t_{sb}	t_{sf}	t_f	ρ
Value	QPSK	$1/2$	100 kHz	$32t_{sf}$	$16t_f$	100 ms	5

Since a-priori information about PU activity is usually not known to SUs, channel occupancy was calculated by selecting a detection threshold δ at acceptable SNR conditions. Therefore, only CR-specific effects were considered, i.e. channel gain was ignored.

A summary of the assumptions made, regarding the length of the CR-specific delays and power penalty rates employed in the simulation platform, is presented in Table 6.7. The power penalty rate assumptions were partially based on the typical power consumption of a GSM handset, as described in [81].

6.3.2 Channel switching evaluation by training sequence variation

In Figure 6.1, theoretical SU channel switching requirements are presented for varying occupancy model training sequence lengths L. Both simulated traffic densities, as described in Table 6.5, are considered (the results in part *a* are for light traffic density and those in part *b* are for high traffic density). The channel switching requirements that are based on occupancy modelling and prediction, are compared against those where channel switching is performed on a purely random basis.

Table 6.7: CR-specific simulation parameter assumptions for investigating SU performance.

	Parameter	Value	Units
Delays	d_{sn}	$t_f/2$	ms
	d_{wt}	t_f	ms
	d_{sw}	$t_f/2$	ms
	d_{tr}	$t_f/10$	ms
Power penalty rate	p_{sn}	500	mW/s
	p_{wt}	50	mW/s
	p_{sw}	80	mW/s
	p_{tp}	100	mW/s
	P_{TF}	600	mW

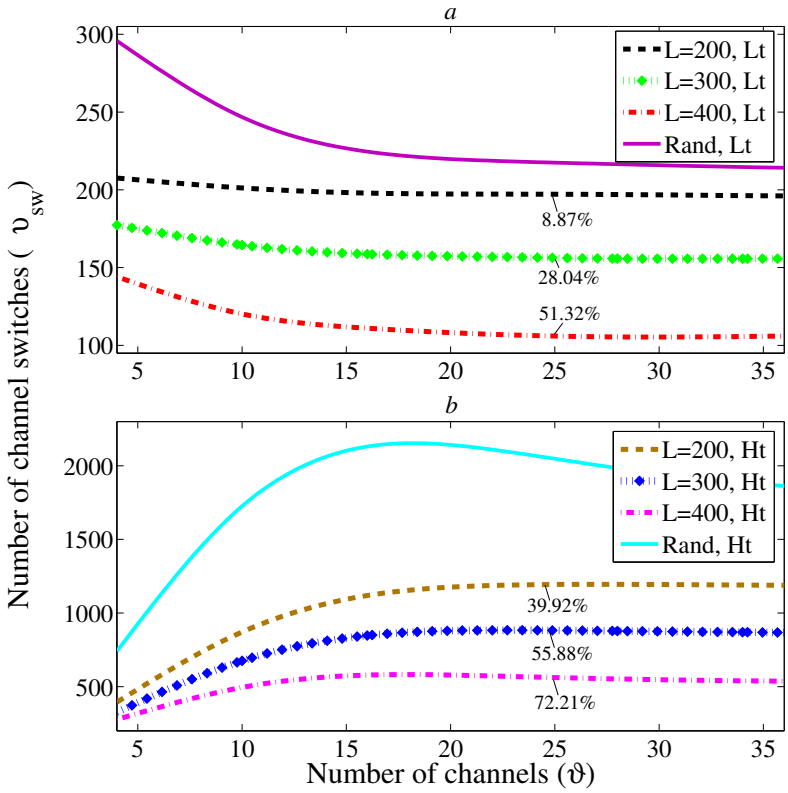

Figure 6.1: Simulated number of required channel switches, for both light (part *a*) and heavy (part *b*) traffic densities, for varying lengths of L.

Since prediction accuracy improves as L increases, a longer L should mean that fewer channel switching operations will be required of the SU. It is thus not surprising that the results shown in Figure 6.1 indicate that this is indeed the case. The results at $\vartheta = 25$, indicate a significant reduction in the required number of SU channel switches than in the random case when prediction is employed and $L = 400$ (51.32% and 72.21% reduction for light and heavy density traffic conditions respectively). However, when L is decreased, the prediction model's benefit to channel switching also decreases by roughly 20% per 100 samples of L.

It is noticeable that under heavy traffic density conditions, the prediction model provides a greater reduction in the required number of SU channel switches than under light traffic density conditions. The adverse effects that higher traffic density has on channel switching are however evident. When $L = 400$ the required number of SU channel switches

under heavy traffic density conditions, is clearly greater than under light traffic density conditions, e.g., when $\vartheta = 22$, SUs under the simulated heavy traffic density conditions would be required to make 80.98% more channel switches than under light traffic conditions.

An asymptotic relationship is evident as ϑ is increased. For light traffic conditions, there is a gradual asymptotic decrease in υ_{sw} as ϑ is increased. However, for heavy traffic conditions, there is an initial increase in υ_{sw} up until $\vartheta = 20$, after which υ_{sw} begins to gradually decrease. The lack of available channels for the SU to operate within under heavy traffic density conditions, for smaller values of ϑ, may explain the increasing shape of these results as opposed to a decreasing shape under light traffic density conditions.

6.3.3 Channel switching evaluation by training algorithm

To illustrate the effect that the different training algorithms may have on channel switching, theoretical SU channel switching requirements are presented in Figure 6.2 for these algorithms. For these plots, the training sequence length has been fixed at $L = 400$.

As shown in Section 6.2, there is not a large difference in the accuracy obtained when using the different training algorithms. It is evident that the BWA does however, provide some improvement over the other training algorithms. At $\vartheta = 25$, the BWA provides a 5.93% and 9.56% improvement over the MA (which seems to be the worst performing algorithm) under light and heavy traffic density conditions respectively. It is also interesting to note, that while the VA performs similarly to the MA under light traffic conditions, it performs 6.90% better than the MA under heavy traffic conditions. However, these improvements are relatively small compared with the differences in complexity required of each algorithm.

6.4 SECONDARY USER PERFORMANCE EVALUATION

Simulations, based on both theoretical and measured occupancy data, have been run on the simulation platform to investigate the effect that prediction model based channel allocation has on certain CR network performance parameters. In this section the channel switching

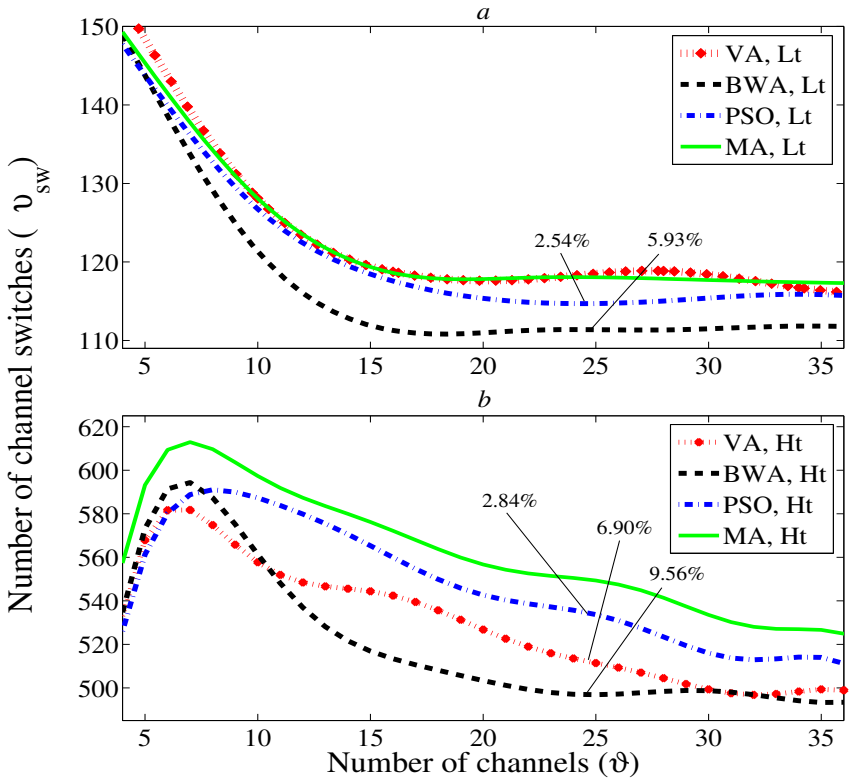

Figure 6.2: Simulated number of required channel switches, for both light (part *a*) and heavy (part *b*) traffic densities, for various occupancy model training algorithms when $L = 400$.

trends observed in Section 6.3 are further applied to calculating their effect on potentially achievable SU throughput, PU disruption rate and SU power consumption. The relationship between prediction model accuracy, PU traffic density, the quick sensing interval and these SU performance measures, is investigated. Due its relatively lower level of computational complexity, the VA was chosen to train the occupancy models in this section.

6.4.1 Performance evaluation by training sequence variation

In previous sections it was shown that prediction model accuracy is a function of the training sequence length L. The effect that this relationship has on the previously mentioned SU performance measures is now investigated. The prediction dependant approach to quick sensing was adopted when generating these results.

6.4.1.1 Secondary user data throughput

SU throughput results based on theoretically simulated PU channel occupancy are presented in Figure 6.3 and results based on practically measured data are presented in Figure 6.4. These results compare the potentially achievable throughput of a SU in a CR network for both heavy and light traffic densities and also for different levels of prediction model accuracy. The throughput measured when prediction is employed is compared against the throughput measured when channel allocation is performed on a purely random basis.

A reduction in υ_{SW} should free up more time for data transmission and ultimately result in an increase throughput. However, although it was shown in the previous section that prediction modelling results in a decrease in υ_{SW}, the trend for throughput under light traffic density with prediction seems to be reversed. For high traffic density though, the results are as expected and throughput does improve as prediction accuracy is increased and is significantly higher than without prediction. A possible explanation for this phenomenon is that under light traffic density conditions there are many more channels available

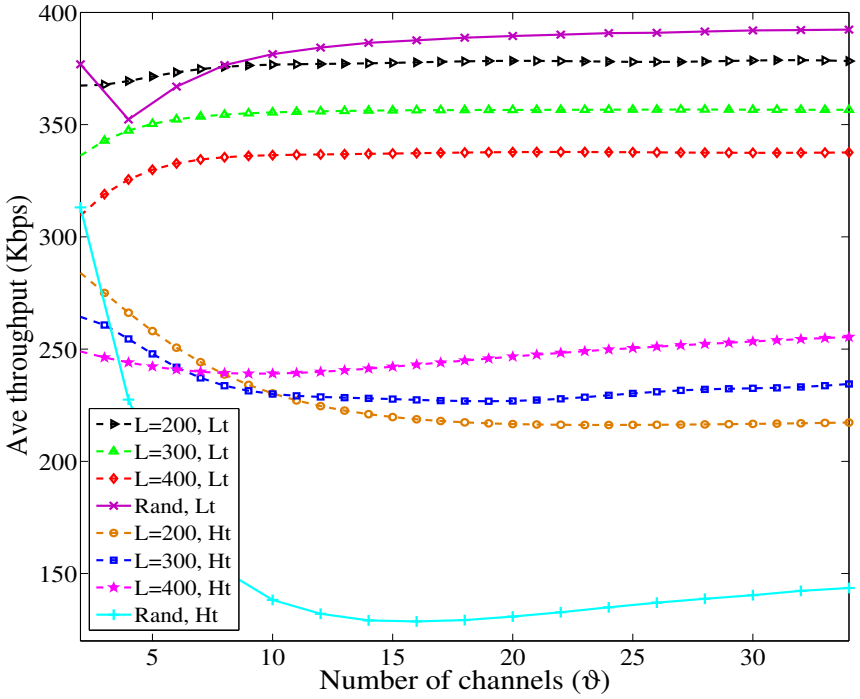

Figure 6.3: Simulated SU throughput for different traffic densities and varying lengths of L.

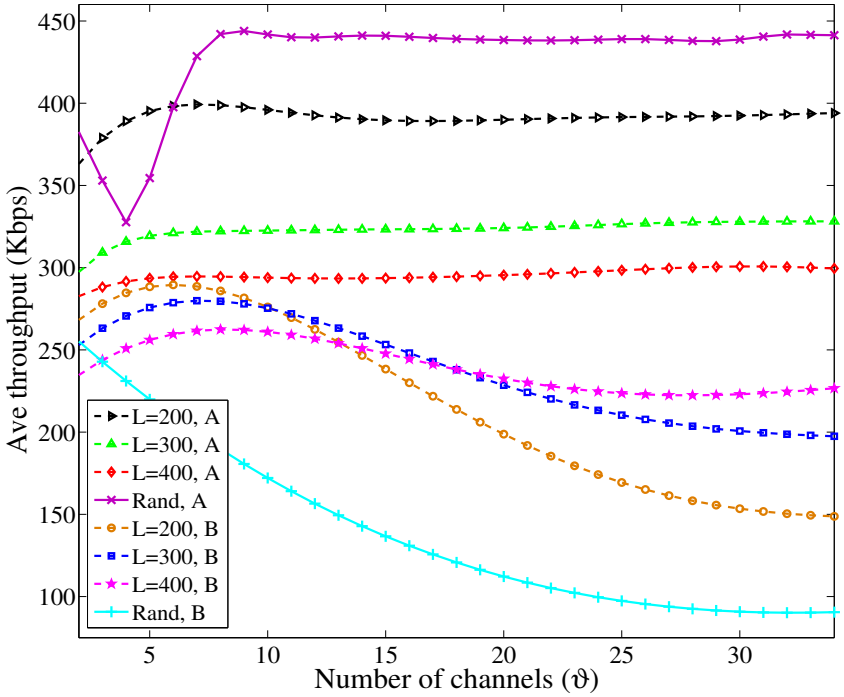

Figure 6.4: Measured SU throughput for different traffic densities and varying lengths of L.

for the SU to switch to and thus in terms of throughput, the reduction gained for υ_{SW} is too small to justify the time spent on gathering training sequence observations, i.e., more time is spent on training the model than is saved by the achieved reduction in υ_{SW}. This trend is observable for both the theoretically simulated and practically measured results.

Similar to spectrum allocation performance, an asymptotic trend is observable for SU throughput. A gradual increase in throughput appears to flatten out as ϑ is increased. For high traffic density there is an initial drop in throughput before this behaviour is observed.

As expected, heavy traffic density has a noticeably negative effect on SU throughput when compared with light traffic density.

When random channel switching is employed, a turning point is observable at $\vartheta = 5$ under light traffic density conditions. This point may indicate the threshold at which the capacity of the channel set $\boldsymbol{F}_q$ begins to exceed the SU input traffic rate.

6.4.1.2 Primary user disruption rate

A plot showing the measured theoretical primary user disruption rate is shown in Figure 6.5. The results for practically measured data are illustrated in Figure 6.6. These plots correspond to the SU throughput results shown in Figure 6.3 and Figure 6.4 and have been generated under the same conditions. The effects experience by PUs, due to SU activity on the same bands in which they are licensed to operate, are illustrated.

While it may seem that prediction is not of benefit to the throughput of SUs under light traffic conditions, there is a clear reduction in the amount of interference experienced by PUs when prediction is employed. Heavy density conditions still seem to experience the greatest reduction in PU disruption, but disruption also decreases under light density conditions as prediction accuracy and the length of L are increased.

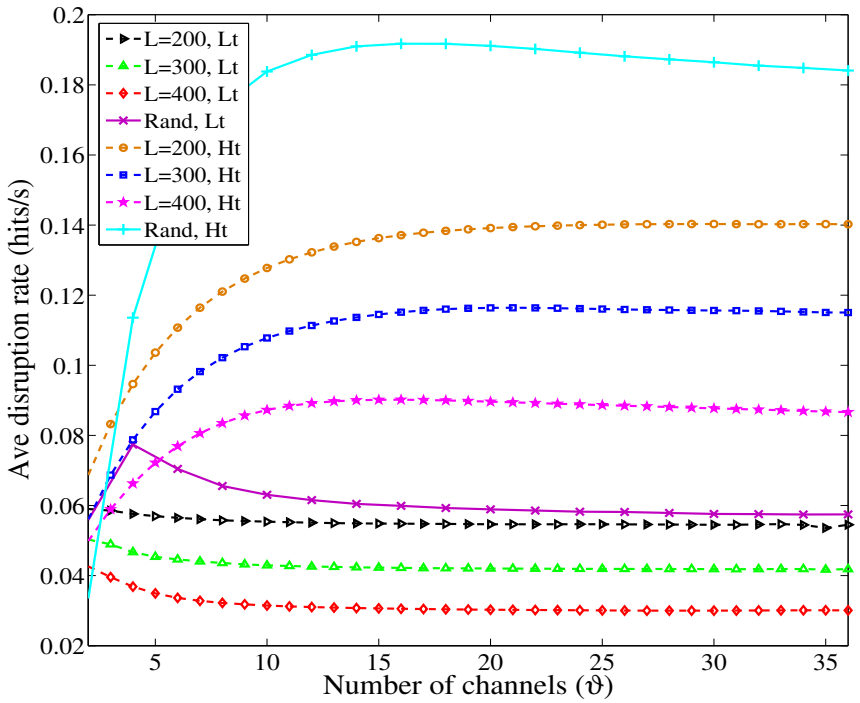

Figure 6.5: Simulated PU disruption rate for different traffic densities and lengths of L.

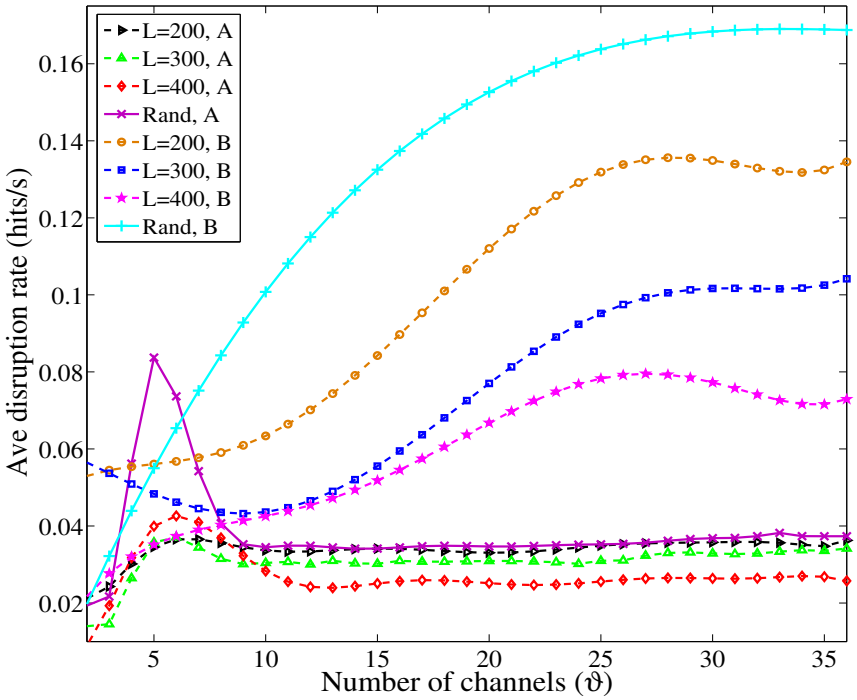

Figure 6.6: Measured PU disruption rate for different traffic densities and lengths of L.

6.4.1.3 Secondary user power consumption

The results for SU power consumption follow a similar trend to the that observed for PU disruption, where prediction provides a significant improvement to SUs as compared to the case where prediction is absent. These improvements also increase as L is increased. Although both traffic densities benefit from prediction, the greatest benefit is again experienced under heavy traffic density conditions. Plots indicating the trends in SU power consumption are provided in Figure 6.7 and Figure 6.8 and illustrate the results derived from theoretical occupancy data and those derived from practically measured data respectively.

The results obtained from practically measured data under high density traffic conditions, seem to require a larger channel set size ϑ before the asymptotic behaviour, seen in the plots based on theoretical occupancy, is observed. This may be attributed to the high power value assigned to the sensing power penalty rate p_{sn} in Table 6.7, together with the higher occupancy of frequency band B when compared to simulated high density traffic conditions. When traffic density is high, the lack of channels available to SUs results in

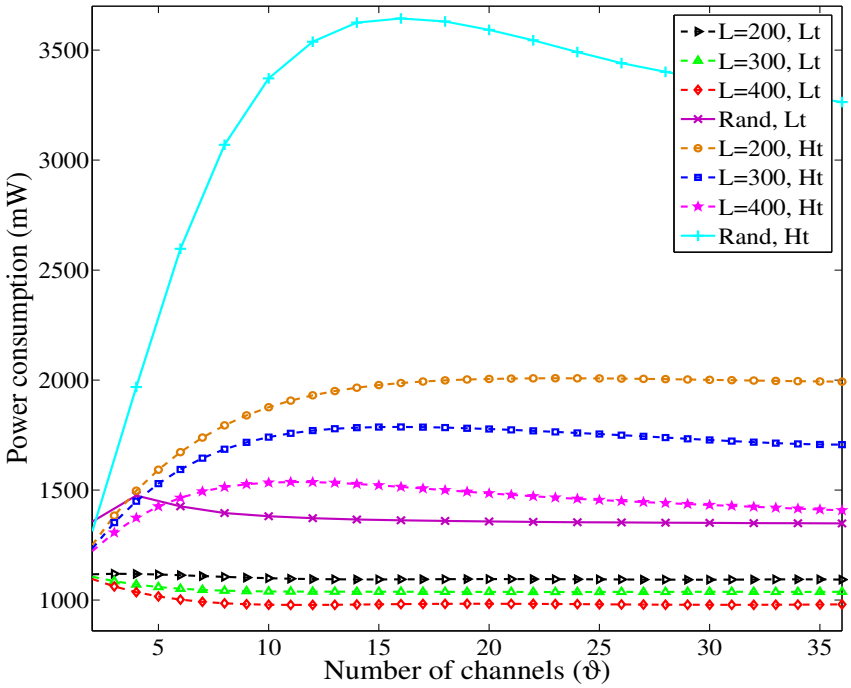

Figure 6.7: Simulated SU power consumption for different traffic densities and lengths of L.

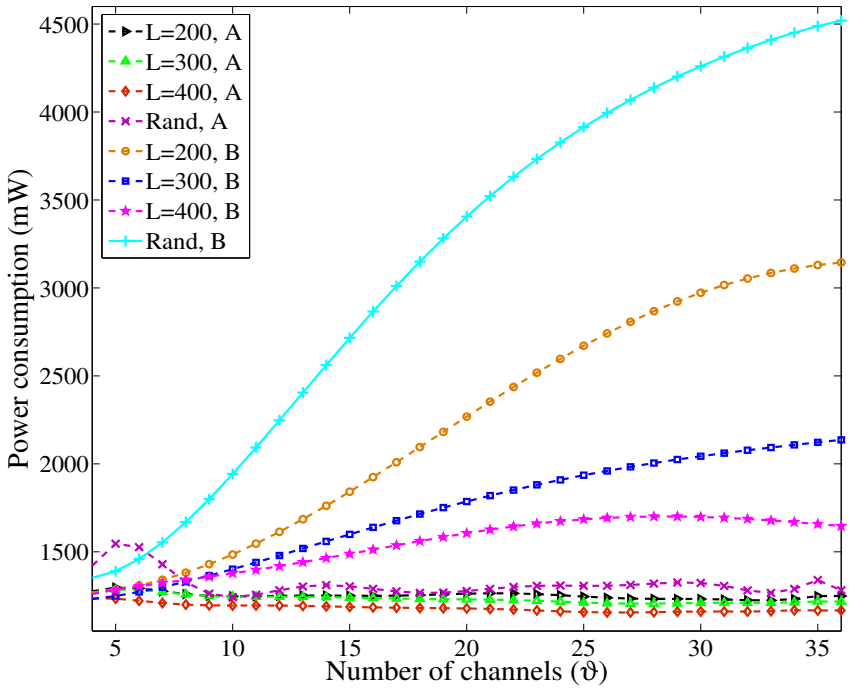

Figure 6.8: Measured SU power consumption for different traffic densities and lengths of L.

a large number of channel switches being required of the SU. The subsequent number of sensing operations needed to keep track of the entire channel set thus increases.

6.4.2 Performance evaluation by quick sensing approach

In this section, the consequences of employing the different quick sensing operations, described in Section 4.3.3, are explored. A further investigation into prediction-model-based SU performance is provide by comparing the aggressive, bi-frame, quad-frame and prediction dependent approaches to quick sensing in terms of the same SU performance measures described in Section 6.4.1.

6.4.2.1 Secondary user data throughput

Plots illustrating the effects that the different approaches to quick sensing have on SU throughput, are provide in Figure 6.9 and Figure 6.10.

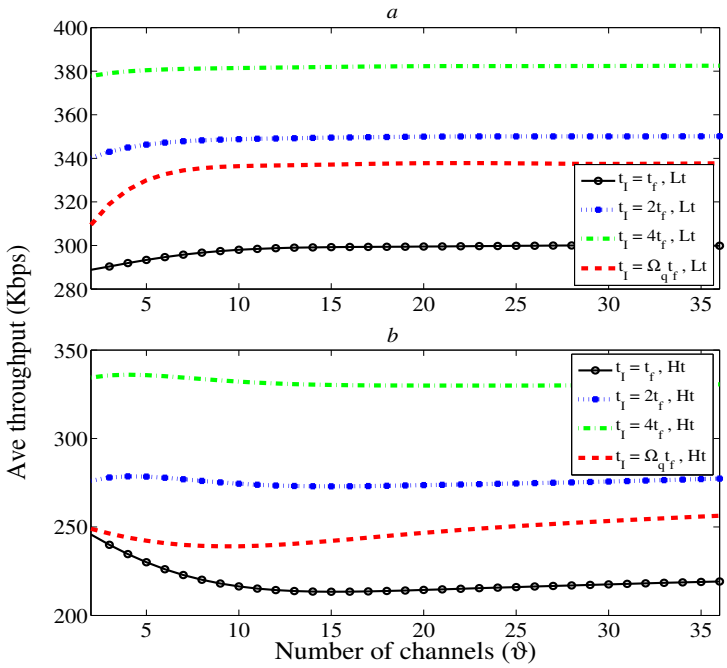

Figure 6.9: Simulated SU throughput for varying lengths of t_I under light (part *a*) and heavy (part *b*) traffic densities.

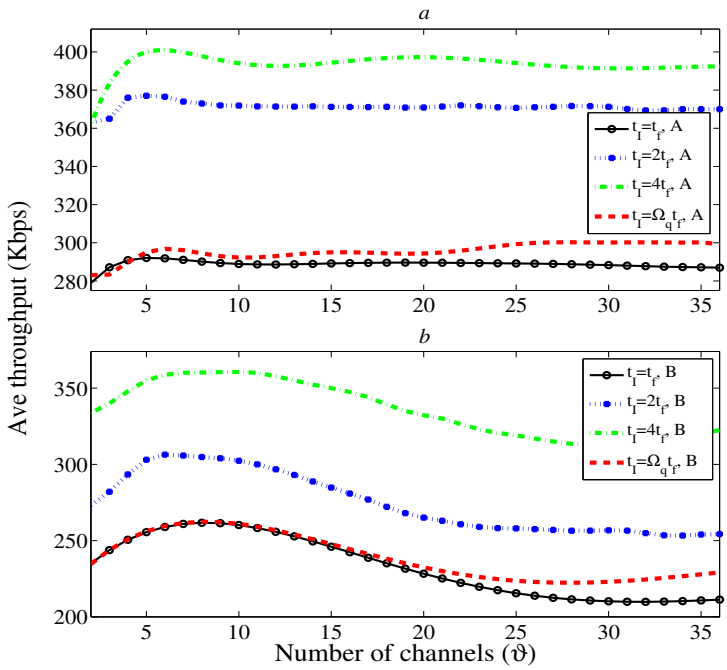

Figure 6.10: Measured SU throughput for varying lengths of t_I for frequency band A (part *a*) and frequency band B (part *b*).

It is evident from the plots, as would be intuitively expected, that SU throughput increases when less time is spent on performing spectrum sensing. The quad-frame approach shows the highest throughput and the aggressive approach the lowest. Although the prediction dependant approach provides more throughput gain than the aggressive approach, it is lower than the bi-frame approach (particularly for the results based on practically measured data). It does however seem to improve as ϑ increases. Similar trends are observable for both traffic densities as well as between the results based on theoretical and practically measured data.

6.4.2.2 Primary user disruption rate

For PU disruption, increasing the length of t_I during a quick sensing operation, has different consequences under light traffic conditions as compared with heavy traffic conditions. As shown in Figure 6.11 and Figure 6.12, PU disruption is negatively influenced by an increase in t_I and follows the same trend by which SU throughput is positively influenced. As t_I is increased, it is thus evident that a slight trade-off may exist between the gains experienced in

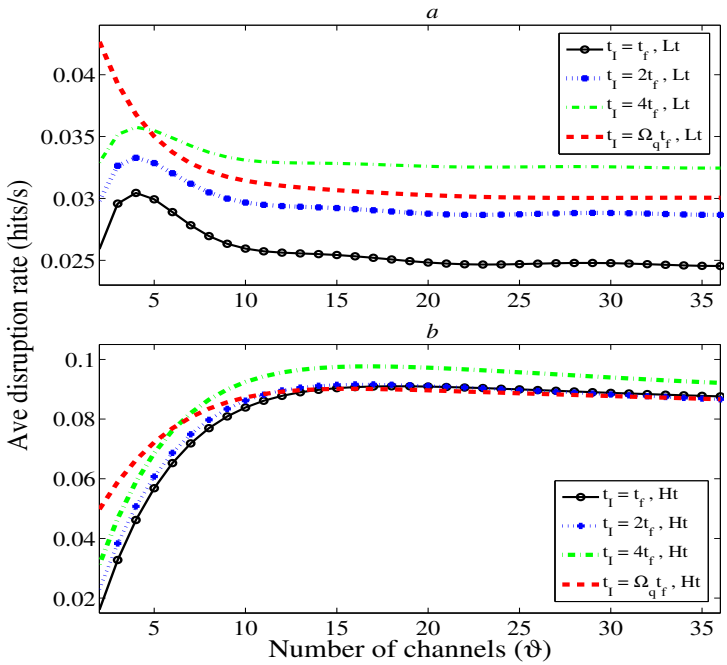

Figure 6.11: Simulated PU disruption rate for varying lengths of t_I under light (part *a*) and heavy (part *b*) traffic densities.

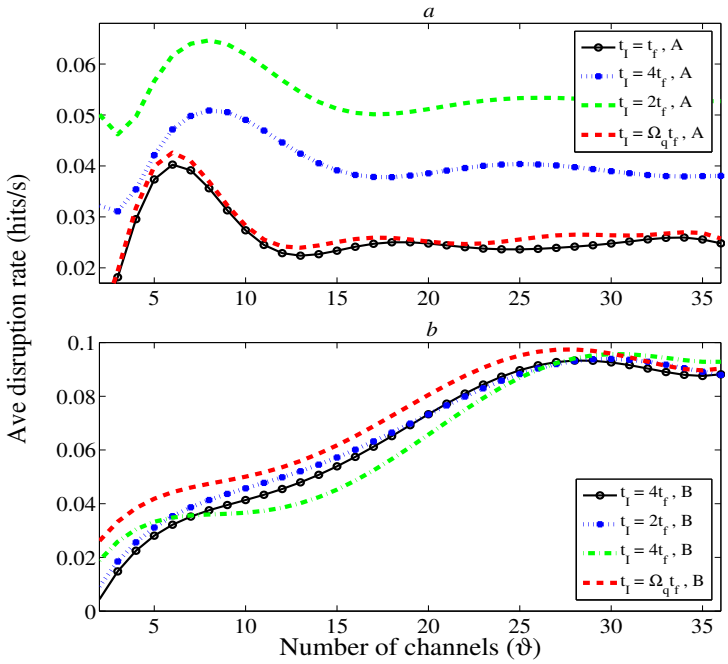

Figure 6.12: Measured PU disruption rate for varying lengths of t_I for frequency band A (part *a*) and frequency band B (part *b*).

SU throughput and the losses incurred for PU disruption. However, this seems to be true under light traffic density conditions only, as there is very little difference between the results obtained from the different quick sensing approaches under heavy traffic conditions. The same reasons, provided in Section 6.4.1.1, for the observed differences between SU throughput under light and heavy density traffic conditions, may provide a possible explanation for this observation.

6.4.2.3 Secondary user power consumption

As described in Section 4.4.4, SU power consumption (due to CR effects) is influenced by the same major parameters as the SU throughput and PU disruption rate, i.e., the required number of channel switches and sensing operations, as well as the required waiting and training time lengths. It is thus not surprising that the same trends observed regarding the length of the quick sensing operation for SU throughput, are observable in Figure 6.13 and Figure 6.14. The notable difference being that the trends surrounding the results for the prediction dependant approach, based on theoretical occupancy data, are more comparable

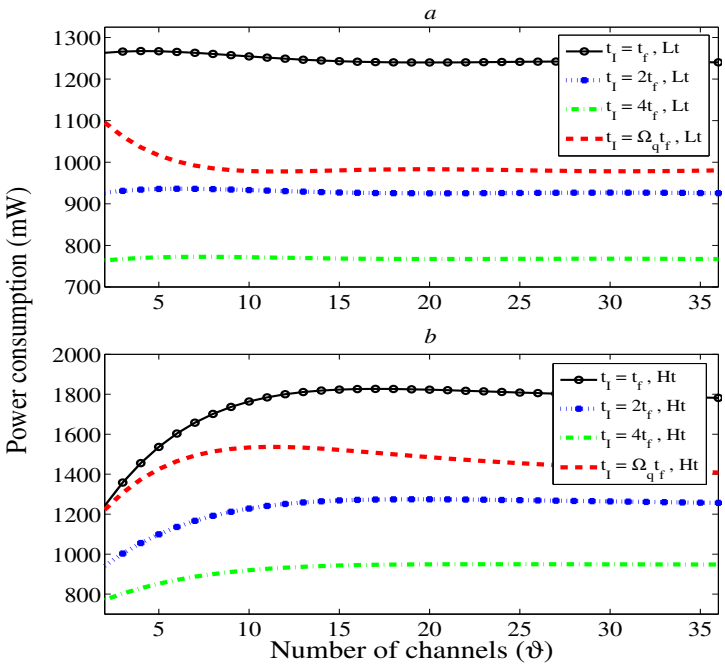

Figure 6.13: Simulated SU power consumption for varying lengths of t_I under light (part *a*) and heavy (part *b*) traffic densities.

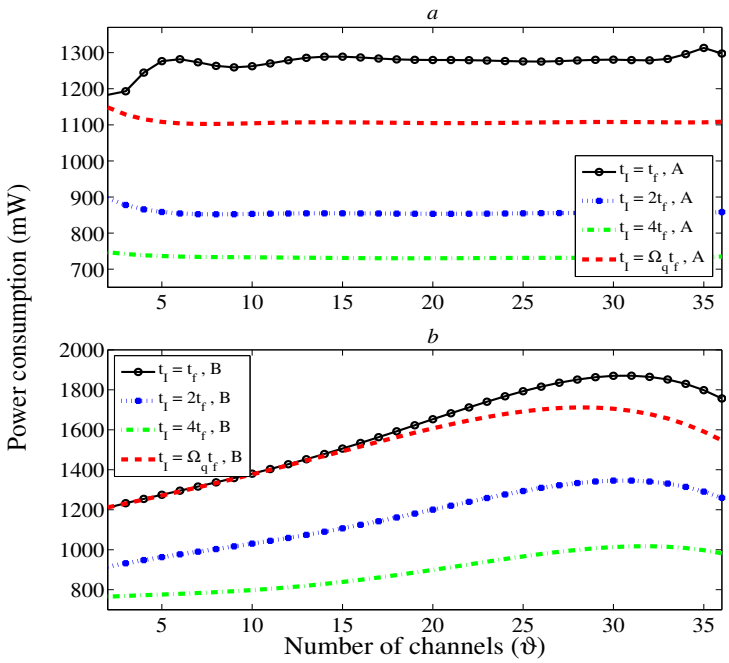

Figure 6.14: Measured SU power consumption for varying lengths of t_I for frequency band A (part *a*) and frequency band B (part *b*).

to those based on practically measured occupancy data, i.e., in these plots the prediction dependent approach shows a larger degree of improvement over the aggressive approach for practically measured data.

6.4.3 Quantification of performance improvement

Further quantification of the graphical results presented for SU performance thus far, are provided in Table 6.8 as a calculation of the SU performance improvement percentage χ. These results were obtained when performing simulations to calculate SU throughput, PU disruption and SU power consumption for varying training sequence lengths L, spectrum sensing approaches t_I and channel set sizes ϑ. To calculate improvement, Θ_x is defined as a general representation of these performance measures. Results from frequency bands A and B are compared with those obtained from theoretical occupancy data for light and heavy density traffic conditions.

When L is varied, χ_L denotes the improvements gained over the case where channel

switching is performed on a purely random basis and has been calculated as follows,

$$\chi_L = 100\left(\frac{\Theta_{Rand} - \Theta_L}{\Theta_{Rand}}\right), \quad L = 200, 300, 400. \tag{6.1}$$

When t_I is varied, χ_{t_I} denotes the differences in performance gained over the use of the aggressive approach to quick sensing. In this case χ_{t_I} has been calculated according to the following expression,

$$\chi_{t_I} = 100\left(\frac{\Theta_{t_I=t_f} - \Theta_{t_I=ut_f}}{\Theta_{t_I=t_f}}\right), \quad u = 2, 4, \Omega_q. \tag{6.2}$$

For both Equation 6.1 and Equation 6.2, values for χ_L and χ_{t_I} have been calculated at $\vartheta = 10$, $\vartheta = 20$ and $\vartheta = 30$.

The negative values listed in Table 6.8 indicate a degradation in SU performance (rather than an improvement) when prediction modelling is employed, as compared with the results obtained from either random channel switching or employing the aggressive approach to the quick sensing operation. This may be explained by the same argument as provided in Section 6.4.1.1.

6.4.3.1 Quantification by training sequence variation

In Table 6.8, the degradation in SU throughput is evident under light traffic density conditions (approximately 14% degradation for the theoretical based results and 32% for the practically based results when $L = 400$ and $\vartheta = 30$) when compared with the significant improvements in the results under heavy traffic density conditions (approximately 80% improvement for the theoretical based results and 145% for the practically based results when $L = 400$ and $\vartheta = 30$).

The decrease in PU disruption rate is also evident, where there is an improvement of roughly 50% and 53% for heavy and light density conditions and an improvement of 28% and 53% for frequency bands A and B respectively ($L = 400$ and $\vartheta = 30$).

The benefit that prediction modelling provides to the power consumption of a SU,

Table 6.8: Percentage performance improvements χ_L, gained through the use PU prediction as opposed to employing random channel switching, and χ_{t_I}, gained by employing less aggressive approaches to the quick spectrum sensing operation.

	Simulated data						**Practical data**					
L/t_I	**Light traffic**			**Heavy traffic**			**Freq. band A**			**Freq. band B**		
$\vartheta = 10$												
200	-1.2	12.2	20.4	66.6	30.5	44.3	-10.4	2.4	0.1	60.4	37.1	23.6
300	-6.8	31.9	24.7	66.4	41.4	48.4	-27.0	11.8	0.2	60.0	56.7	27.9
400	-11.8	50.1	29.1	73.0	52.5	54.5	-33.5	18.0	3.9	51.6	57.8	29.1
$2t_f$	17.0	-14.3	25.6	26.8	-2.7	30.4	28.7	-79.2	32.4	16.2	-10.5	25.4
$4t_f$	28.0	-27.6	38.5	53.5	-9.9	47.8	35.8	-126	41.9	38.6	11.4	42.2
$\Omega_q t_f$	12.9	-21.2	22.0	10.5	-4.0	13.0	1.2	-3.3	12.5	0.3	-21.0	0.2
$\vartheta = 20$												
200	-2.9	7.1	19.3	65.8	27.4	44.2	-11.1	4.3	7.5	83.7	30.1	35.9
300	-8.5	28.6	23.6	73.4	39.2	50.4	-26.1	11.1	3.6	108	52.4	48.6
400	-13.3	48.5	27.6	87.9	53.0	58.5	-32.7	26.3	7.6	110	58.3	53.5
$2t_f$	16.8	-15.7	25.4	27.7	-0.4	30.2	28.1	-52.3	33.3	15.7	-0.8	27.2
$4t_f$	27.7	-30.9	38.1	54.1	-7.1	48.0	35.4	-102	43.0	44.6	11.8	45.4
$\Omega_q t_f$	12.8	-21.7	20.7	14.8	1.3	18.2	1.6	-2.2	13.7	1.5	-10.7	2.3
$\vartheta = 30$												
200	-3.5	5.3	19.1	54.2	24.8	40.4	-10.6	3.3	12.7	74.3	19.5	32.7
300	-9.0	27.3	23.2	65.3	37.9	48.3	-25.3	10.4	8.9	124	40.3	52.9
400	-13.9	47.9	27.5	79.7	52.8	57.2	-31.6	28.0	12.5	145	52.9	60.1
$2t_f$	16.7	-16.3	25.4	26.8	0.5	29.7	28.7	-64.6	33.1	21.3	0.2	28.0
$4t_f$	27.5	-31.4	38.2	52.1	-6.0	47.3	35.7	-120	42.8	48.3	-0.9	46.0
$\Omega_q t_f$	12.5	-21.2	21.2	16.3	1.3	20.0	4.0	-9.8	13.3	5.2	-4.4	7.5
	R_{su}	D_{pu}	P_{su}	R_{su}	D_{pu}	P_{su}	R_{su}	D_{pu}	P_{su}	R_{su}	D_{pu}	P_{su}

when $L = 400$ and $\vartheta = 30$, is seen as a 28% and 57% reduction for light and heavy density conditions and 13% and 60% reduction for frequency bands A and B respectively.

6.4.3.2 Quantification by quick sensing approach

The influence that the less aggressive quick sensing approaches have on SU throughput, is provided in greater detail in Table 6.8, e.g., when $t_I = 4t_f$ and $\vartheta = 30$, there is approximately a 28% and 52% increase for light and heavy density traffic and a 36% and 48% increase for frequency bands A and B respectively.

For the PU disruption rate, the difference between the traffic densities is clear: a 31% and 120% increase (for light density traffic conditions and frequency band A respectively) compared with a 6% and 1% increase (for heavy density traffic conditions and frequency band B respectively) when $t_I = 4t_f$ and $\vartheta = 30$.

As seen from the graphical results, there is a greater correlation between SU power consumption and SU throughput with an approximate decrease of 38% and 47%, for light density traffic and frequency band A respectively, and an approximate decrease of 43% and 46% for heavy density traffic and frequency band B, when $t_I = 4t_f$ and $\vartheta = 30$.

6.5 THEORETICAL AND PRACTICALLY MEASURED RESULT SIMILARITY

To test how results based on theoretical occupancy $\boldsymbol{H}$ compare to those based on measured occupancy $\boldsymbol{\Gamma}$, an average percentage similarity score η has been calculated as follows,

$$\eta = 100\left(1 - \frac{1}{\vartheta}\sum_{i=0}^{\vartheta}\frac{|\boldsymbol{\Gamma} - \boldsymbol{H}|}{\boldsymbol{H}}\right). \tag{6.3}$$

The values calculated for η in Table 6.9 and Table 6.10 indicate that there is a good overall correlation between the results generated from theoretical occupancy data and those generated from practically measured data. Average percentage similarity scores when training sequence variation is considered, as listed in Table 6.1, Table 6.2, Table 6.3, Table

Table 6.9: CR performance similarity scores η for frequency bands A and B for varying training sequence lengths L.

L	**Freq. band A**			**Freq. band B**		
200	96.0	61.4	86.3	82.4	75.6	73.7
300	91.1	72.1	82.0	88.7	67.9	86.2
400	88.3	81.9	80.5	92.2	68.6	88.8
Ave.	91.8	71.8	82.9	87.8	70.7	82.9
	$\boldsymbol{R}_{su}$	$\boldsymbol{D}_{pu}$	$\boldsymbol{P}_{su}$	$\boldsymbol{R}_{su}$	$\boldsymbol{D}_{pu}$	$\boldsymbol{P}_{su}$

6.4 and Table 6.9, indicate a high correlation between theoretical and measured results. This is evident for both occupancy model prediction accuracy ($\eta \approx 96\%$ for VA, $\eta \approx 94\%$ for BWA, $\eta \approx 92\%$ for PSO and $\eta \approx 93\%$ for MA) and SU throughput ($\eta \approx 90\%$). There is also a good correlation for SU power consumption ($\eta \approx 83\%$) and a lower but still fair correlation ($\eta \approx 71\%$) for PU disruption rate.

The effect that the various approaches to quick sensing has on SU performance similarity, is summarised in Table 6.10. For both frequency band A and B, the results for both theoretically based SU throughput and SU power consumption, correlate well with the

Table 6.10: CR performance similarity scores η for frequency bands A and B for varying approaches to the quick sensing operation.

$\boldsymbol{t_I}$	**Freq. band A**			**Freq. band B**		
$t_I = t_f$	96.8	89.4	97.1	91.4	73.8	89.7
$t_I = 2t_f$	93.6	62.3	92.3	93.7	74.7	91.8
$t_I = \Omega_q t_f$	97.3	36.5	95.3	95.5	69.2	92.7
$t_I = 4t_f$	88.2	81.9	88.0	92.2	77.9	89.1
Ave.	94.0	67.5	93.2	93.2	73.9	90.8
	$\boldsymbol{R}_{su}$	$\boldsymbol{D}_{pu}$	$\boldsymbol{P}_{su}$	$\boldsymbol{R}_{su}$	$\boldsymbol{D}_{pu}$	$\boldsymbol{P}_{su}$

results based on practical measurements ($\eta \approx 94\%$ for R_{SU} and $\eta \approx 92\%$ for P_{SU}). The correlation for PU disruption rate is again lower but fair ($\eta \approx 71\%$). The lower η values obtained for PU disruption rate when the prediction dependant quick sensing approach was employed, may be attributed to significantly improved performance when simulations were based on the measured occupancy data as opposed to the theoretical occupancy data. These results compare well with the results listed in Table 6.9, where training sequence variation is considered.

6.6 SUMMARY OF RESULTS

In this chapter, an evaluation of the performance of the HMM based channel occupancy model was presented. Four different model training algorithms were compared for a range of training sequence lengths L. The accuracy of the model was not significantly altered by the use of the different algorithms. However, it was shown that the VA introduced by far the least amount of complexity into the modelling and prediction process.

Results indicating the effect that channel modelling and prediction may have on SU channel allocation were also presented. It was shown that increasing L provided better PU prediction accuracy and consequently resulted in improved proactive channel allocations. Prediction modelling was found to provide an improvement over the random channel switching approach under both sets of traffic densities. The use of the BWA was also found to slightly reduce SU channel switching requirements when compared with the other model training algorithms.

The SU performance simulator was employed to investigate SU performance measures, as influenced by the prediction model based channel allocation process, that included: SU throughput, PU disruption rate and SU power consumption. These performance measures were evaluated for varying lengths of L as well as for different approaches to the quick sensing operation. It was found that increasing L generally leads to an improvement in the throughput and power consumption requirements of SUs and a degradation in the amount of disruption that may be experienced by PUs. However, SU throughput performed

more poorly under light traffic density conditions.

It was also found that there is a slight trade-off between SU throughput and the PU disruption rate when t_I is increased. However, the quick sensing approach did not seem to have much of an effect on the PU disruption rate under light traffic density conditions. However, increasing t_I always led to a reduction in SU power consumption.

Overall, it was verified that predication modelling improves the performance of a SU in a CR network and that the results generated from theoretical occupancy data correlated well with those generated from practically measured data. Heavier traffic densities were also found to have a negative effect on all of the above mentioned CR performance measures.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

7.1 CONCLUDING REMARKS

In this dissertation potential CR performance benefits, due to channel occupancy and prediction modelling, were investigated.

In Chapter 2, the concept of CR was introduced and a summary of its basic functionality was presented. Topics of importance to the theoretical foundation of this thesis were discussed. These included methods employed in the literature for detecting PU activity, modelling PU occupancy and performing SU channel allocation. Some of the IEEE draft standards pertaining to CR were also summarised.

In Chapter 3, an HMM based model for modelling PU activity in a CR network was described. Various model training algorithms were also presented, namely: the BWA, VA, PSO and the MA. A complexity analysis was performed for all four of these algorithms and it was determined that the VA would require the least number of FLOPS to perform a single iteration of the training algorithm.

A simulation platform for investigating SU performance in a CR network was presented in Chapter 4. A channel switching algorithm for performing OSA (based on the two-state HMM described in Chapter 3) was proposed and various approaches to per-

forming the quick sensing operation described in the algorithm, were presented. The CR performance measures investigated on the simulation platform were also presented, namely: SU throughput, PU disruption rate and SU power consumption.

In Chapter 5 a spectrum measurement campaign was briefly described, from which two frequency bands were selected for analysis on the simulation platform. A threshold based method for determining PU occupancy was discussed and the measured power spectra and channel occupancy of the selected bands were presented.

Finally, results pertaining to the PU channel occupancy model and the CR performance measures investigated with the aid of the simulation platform, were presented and discussed in Chapter 6. It was found that all of the investigated training algorithms provided similar prediction accuracies but that the VA converged on a solution far more quickly than the other algorithms did. Results from the simulation platform indicated that channel occupancy prediction significantly reduces the number of channel switches required of SUs in a CR network. Significant performance improvements were clearly visible when PU activity was accurately modelled and intelligent proactive channel switching techniques employed. Overall, significant throughput gains, a reduction in the PU disruption rate and a reduction in SU power consumption were evident. Results obtained from measured occupancy data were comparable with those obtained from simulated occupancy data and the model was shown to be of general benefit to SUs under both light and heavy traffic density conditions.

7.2 FUTURE RESEARCH

Further research areas that could be explored may include the following:

- A further investigation into the optimal combination and values for ρ, L and t_I. Various combinations of these parameters may be optimal for different PU traffic profiles.

- A comparison with other occupancy modelling techniques. The HMM based occupancy model may be compared with other channel occupancy models in the literature

to determine which model would be most suited to solving this problem. This could also be extended to the development of a new PU occupancy model.

- Further verification of the performance of the results based on theoretical PU activity, by testing the model on additional sets of measured occupancy data. This may include different commercial frequency bands and varying channel bandwidths.

- Incorporating additional constraints into the channel allocation process, e.g., the quality of the channel within which the SU wishes to operate. A cross layered approach may also be considered for this purpose.

- Expanding the current SU simulator to include multiple SUs. An investigation into how to best manage the channel allocation process when more than one SU is vying for usage of the same set of PU channels, would be a logical continuation of the work presented in this dissertation. In this case, SUs may need to employ some form of conflict resolution to ensure that fair opportunity is provided to all competing users of the spectrum. Techniques such as game theory may be considered. The way in which the CR performance parameters, investigated in this dissertation, are affected by the introduction of multiple SUs should be considered and the concept of SU disruption rate could be investigated as another potential performance parameter. An investigation into how cooperation between multiple SUs could be employed to perform channel allocation should also be considered.

- The incorporation of this work into a complete CR simulator and/or physical CR test bed.

REFERENCES

[1] A. F. Molisch, *Wireless Communications*. Chichester, England: John Wiley & Sons Ltd, Mar. 2009.

[2] Federal Communications Commission, "Mobile broadband: the benefits of additional spectrum," FCC staff technical paper, Oct. 2010.

[3] T. A. Weiss and F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, no. 3, pp. 8–14, Mar. 2004.

[4] Federal Communications Commission, "Spectrum policy task force report," Tech. Rep. ET Docket 02-155, Nov. 2002.

[5] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[6] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, Sept. 2006.

[7] J. Mitola III and G. Q. Maguire Jr., "Cognitive radio: making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[8] J. Mitola III, "Software radio architecture: a mathematical perspective," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 4, pp. 514–538, Apr. 1999.

[9] M. Wellens, A. de Baynast, and P. Mähönen, "Performance of dynamic spectrum access based on spectrum occupancy statistics," *IET Commun.*, vol. 2, no. 6, pp. 772–782, Jul. 2008.

[10] C. Ghosh, C. Cordeiro, D. P. Agrawal, and M. B. Rao, "Markov chain existence and hidden markov models in spectrum sensing," in *Proc. 7th Annu. IEEE Int. Conf. Pervasive Comput. Commun.*, Galveston, TX, 2009, pp. 1–6.

[11] M. A. McHenry, "NSF spectrum occupancy measurements project summary," Shared Spectrum Company, Tech. Rep., Aug. 2005.

[12] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood, "Chicago spectrum occupancy measurements & analysis and a long-term studies proposal," in *Proc. Int. Workshop Technol. Policy Accessing Spectr.*, Boston, MA, 2006.

[13] M. López-Benítez, F. Casadevall, A. Umbert, J. Pérez-Romero, R. Hachemani, J. Palicot, and C. Moy, "Spectral occupation measurements and blind standard recognition sensor for cognitive radio networks," in *Proc. Int. Conf. Cognitive Radio Oriented Wireless Netw. Commun.*, Hannover, Germany, 2009, pp. 1–9.

[14] M. Matinmikko, M. Mustonen, M. Höyhtyä, T. Rauma, H. Sarvanko, and A. Mämmelä, "Distributed and directional spectrum occupancy measurements in the 2.4 GHz ISM band," in *Proc. Int. Symp. Wireless Commun. Syst.*, York, United Kingdom, 2010, pp. 976–980.

[15] M. Wellens, J. Wu, and P. Mähönen, "Lessons learned from an extensive occupancy measurement campaign and stochastic duty cycle," *Mobile Netw. Appl.*, vol. 15, no. 3, pp. 461–474, Jun. 2010.

[16] T. M. Taher, R. B. Bacchus, K. J. Zdunek, and D. A. Roberson, "Long-term spectral occupancy findings in Chicago," in *Proc. IEEE Int. Symp. Dynamic Spectr. Access Netw.*, Aachen, Germany, 2011, pp. 100–107.

[17] J. G. Proakis and M. Salehi, *Digital communications*, 5th ed. Singapore: McGraw Hill, 2008.

[18] T. S. Rappaport, *Wireless communications: principles and practice*, 2nd ed. Prentice Hall, 2002.

[19] V. D. Chakravarthy, Z. Wu, A. Shaw, M. A. Temple, R. Kannan, and F. Garber, "A general overlay/underlay analytic expression representing cognitive radio waveform," in *Proc. Int. Conf. Waveform Diversity Des.*, Pisa, Italy, 2007, pp. 69–73.

[20] *IEEE Standard definitions and concepts for dynamic spectrum access: terminology relating to emerging wireless networks, system functionality, and spectrum management*, IEEE Standard 1900.1, 2008.

[21] ITU-R, "Definitions of Software Defined Radio (SDR) and Cognitive Radio System (CRS)," ITU-R, Tech. Rep. SM.2152, Sept. 2009.

[22] J. Mitola III, "Software radios: survey, critical evaluation and future directions," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 8, no. 4, pp. 25–36, Apr. 1993.

[23] T. Yücek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 1, pp. 116–130, 2009.

[24] E. Y. Lee, E. K. Walton, J. Young, S. Gemeny, T.-H. Lee, N. Roberts, E. Bosso, and E. Huang, "The software defined antenna: microstrip antennas with gaps," in *Proc. IEEE Int. Symp. Antennas Propag. Soc.*, Toronto, Canada, 2010, pp. 1–4.

[25] J. Kovitz and Y. Rahmat-Samii, "Micro-actuated pixel patch antenna design using particle swarm optimization," in *Proc. IEEE Int. Symp. Antennas Propag. Soc.*, Spokane, WA, 2011, pp. 2415–2418.

[26] J. Ma and Y. Li, "Soft combination and detection for cooperative spectrum sensing in cognitive radio networks," in *Proc. IEEE Global Telecommun. Conf.*, Washington, DC, 2007, pp. 3139–3143.

[27] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP J. Advances Signal Process.*, vol. 2010, pp. 1–15, 2010.

[28] K. Kim, I. A. Akbar, K. K. Bae, J.-S. Um, C. M. Spooner, and J. H. Reed, "Cyclostationary approaches to signal detection and classification in cognitive radio," in *Proc. 2nd IEEE Int. Symp. New Frontiers Dynamic Spectr. Access Netw.*, Dublin, Ireland, 2007, pp. 212–215.

[29] A. Ghasemi and E. S. Sousa, "Optimization of spectrum sensing for opportunistic spectrum access in cognitive radio networks," in *Proc. 4th IEEE Consum. Commun. Netw. Conf.*, Las Vegas, NV, 2007, pp. 1022–1026.

[30] S. Xu, Z. Zhao, and J. Shang, "Spectrum sensing based on cyclostationarity," in *Proc. 2008 Workshop Power Electron. Intell. Transp. Syst.*, Guangzhou, China, 2008, pp. 171–174.

[31] Y. Lin, C. He, L. Jiang, and D. He, "A cyclostationary-based spectrum sensing method using stochastic resonance in cognitive radio," in *Proc. IEEE Int. Conf. Commun. Workshops*, Cape Town, South Africa, 2010, pp. 1–5.

[32] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor, "Spectrum sensing in cognitive radios based on multiple cyclic frequencies," in *Proc. 2nd Int. Conf. Cognitive Radio Oriented Wireless Netw. Commun.*, Orlando, FL, 2007, pp. 37–43.

[33] R. Tandra and A. Sahai, "Fundamental limits on detection in low SNR under noise uncertainty," in *Proc. Int. Conf. Wireless Netw. Commun. Mobile Comput.*, Maui, HI, 2005, pp. 464–469.

[34] S. Tang and B. L. Mark, "An adaptive spectrum detection mechanism for cognitive radio networks in dynamic traffic environments," in *Proc. IEEE Global Telecommun. Conf.*, New Orleans, LA, 2008, pp. 3009–3013.

[35] G. Atia, S. Aeron, E. Ermis, and V. Saligrama, "On throughput maximization and interference avoidance in cognitive radios," in *Proc. 5th IEEE Consum. Commun. Netw. Conf.*, Las Vegas, NV, 2008, pp. 963–967.

[36] W. Zhang and K. B. Letaief, "Cooperative spectrum sensing with transmit and relay diversity in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4761–4766, Dec. 2008.

[37] W. Zhang, R. K. Mallik, and K. Ben Letaief, "Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5761–5766, Dec. 2009.

[38] G. Ganesan and Y. Li, "Agility improvement through cooperative diversity in cognitive radio," in *Proc. IEEE Global Telecommun. Conf.*, vol. 5, St. Louis, MO, 2005, pp. 2505–2509.

[39] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, pp. 1326–1337, Apr. 2008.

[40] A. Pandharipande and J.-P. M. G. Linnartz, "Performance analysis of primary user detection in a multiple antenna cognitive radio," in *Proc. IEEE Int. Conf. Commun.*, Glasgow, Scotland, 2007, pp. 6482–6486.

[41] M. Höyhtyä, S. Pollin, and A. Mämmelä, "Performance improvement with predictive channel selection for cognitive radios," in *Proc. 1st Int. Workshop Cognitive Radio Adv. Spectr. Manage.*, Aalborg, Denmark, 2008, pp. 1–5.

[42] C. Clancy, J. Hecker, E. Stuntebeck, and T. O'Shea, "Applications of machine learning to cognitive radio networks," *IEEE Wireless Commun.*, vol. 14, no. 4, pp. 47–52, Aug. 2007.

[43] L. Yang, L. Cao, and H. Zheng, "Proactive channel access in dynamic spectrum net-

works," *Physical Commun.*, vol. 1, no. 2, pp. 103–111, Jun. 2008.

[44] G. Yuan, R. C. Grammenos, Y. Yang, and W. Wang, "Performance analysis of selective opportunistic spectrum access with traffic prediction," *IEEE Trans. Veh. Technol.*, vol. 59, no. 4, pp. 1949–1959, May 2010.

[45] S. Yarkan and H. Arslan, "Binary time series approach to spectrum prediction for cognitive radio," in *Proc. IEEE Veh. Technol. Conf.*, Baltimore, MD, 2007, pp. 1563–1567.

[46] T. W. Rondeau, C. J. Rieser, T. M. Gallagher, and C. W. Bostian, "Online modeling of wireless channels with hidden Markov models and channel impulse responses for cognitive radios," in *Proc. IEEE MTT-S Int. Microw. Symp. Dig.*, Fort Worth, TX, 2004, pp. 739–742.

[47] S. Geirhofer, L. Tong, and B. M. Sadler, "A measurement-based model for dynamic spectrum access in WLAN channels," in *Proc. IEEE Military Commun. Conf.*, Washington, DC, 2007, pp. 1–7.

[48] Z.-J. Zhao, S.-L. Zheng, C.-Y. Xu, and X.-Z. Kong, "Discrete channel modelling based on genetic algorithm and simulated annealing for training hidden Markov model," *Chinese Physics*, vol. 16, no. 6, pp. 1619–1623, Jun. 2007.

[49] H. Zhao, T. Luo, and G. Yue, "Multi-channel myopic sensing for opportunistic spectrum access using channel correlation," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Tokyo, Japan, 2009, pp. 696–700.

[50] Y. Yao, Z. Feng, W. Li, and Y. Qian, "Dynamic spectrum access with QoS guarantee for wireless networks: a Markov approach," in *Proc. IEEE Global Telecommun. Conf.*, Miami, FL, 2010, pp. 1–5.

[51] Z. Zhao, Z. Peng, S. Zheng, and J. Shang, "Cognitive radio spectrum allocation using evolutionary algorithms," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4421–4425, Sep. 2009.

[52] I. A. Akbar and W. H. Tranter, "Dynamic spectrum allocation in cognitive radio using hidden markov models: Poisson distributed case," in *Proc. IEEE SoutheastCon*, Richmond, VA, 2007, pp. 196–201.

[53] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed. Wiley & Sons, Inc., 2001.

[54] C.-H. Park, S.-W. Kim, S.-M. Lim, and M.-S. Song, "HMM based channel status predictor for cognitive radio," in *Proc. IEEE Asia-Pacific Microw. Conf.*, Bangkok, Singapore, 2007, pp. 1–4.

[55] C. Ghosh, S. Pagadarai, D. P. Agrawal, and A. M. Wyglinski, "A framework for Statistical Wireless Spectrum Occupancy Modeling," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 38–44, Jan. 2010.

[56] P. Kulkarni, T. Lewis, and Z. Fan, "Simple traffic prediction mechanism and its applications in wireless networks," *Wireless Pers. Commun.*, vol. 59, no. 2, pp. 261–274, Jul. 2011.

[57] W.-K. Kuo and S.-Y. Lien, "Dynamic resource allocation for supporting real-time multimedia applications in IEEE 802.15.3 WPANs," *IET Commun.*, vol. 3, no. 1, pp. 1–9, Jan. 2009.

[58] A. M. Adas, "Using adaptive linear prediction to support real-time VBR video under RCBR network service model," *IEEE/ACM Trans. Netw.*, vol. 6, no. 5, pp. 635–644, Oct. 1998.

[59] D. Thilakawardana, K. Moessner, and R. Tafazolli, "Darwinian approach for dynamic spectrum allocation in next generation systems," *IET Commun.*, vol. 2, no. 6, pp. 827–836, Feb. 2008.

[60] C. Stevenson, G. Chouinard, L. Zhongding, H. Wendong, S. Shellhammer, and W. Caldwell, "IEEE 802.22: The first cognitive radio wireless regional area network

standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, Jan. 2009.

[61] "National radio frequency plan," *Government Gazette, Republic of South Africa*, vol. 541, no. 33409, Jul. 2010.

[62] S. D. Barnes and B. T. Maharaj, "Performance of a hidden Markov channel occupancy model for cognitive radio," in *Proc. IEEE AFRICON Conf.*, Livingstone, Zambia, 2011, pp. 1–6.

[63] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.

[64] C. M. Bishop, *Pattern recognition and machine learning*. New York: Springer, 2006.

[65] E. Elbeltagi, T. Hegazy, and D. Grierson, "Comparison among five evolutionary-based optimization algorithms," *Adv. Eng. Informat.*, vol. 19, no. 1, pp. 43–53, Jan. 2005.

[66] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. IEEE Conf. Evol. Comput.*, Anchorage, AK, 1998, pp. 69–73.

[67] B. Allaoua, A. Laoufi, B. Gasbaoui, and A. Abderrahmani, "Neuro-Fuzzy DC motor speed control using particle swarm optimization," *Leonardo Electron. J. Pract. Technol.*, vol. 8, no. 15, pp. 1–18, Mar. 2009.

[68] C.-S. Yang, L.-Y. Chuang, Y.-J. Chen, and C.-H. Yang, "Feature selection using memetic algorithms," in *Proc. Int. Conf. Convergence Hybrid Inf. Technol.*, Busan, South Korea, 2008, pp. 416–423.

[69] P. Mertz and B. Freisleben, "A genetic local search approach to the quadratic assignment problem," in *Proc. 7th Int. Conf. Genetic Algorithms*, San Diego, CA, 1997, pp. 465–472.

[70] S. J. Russell and P. Norvig, *Informed search and artificial intelligence: a modern approach*. Upper Saddle River, NJ: Prentice-Hall, 2003.

[71] E. Ifeachor and B. Jervis, *Digital signal processing: a practical approach*, 2nd ed. Prentice-Hall International, 2002.

[72] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, United Kingdom: Cambridge University Press, 2004.

[73] P. Paysarvi-Hoseini and N. C. Beaulieu, "Optimal wideband spectrum sensing framework for cognitive radio systems," *IEEE Trans. Signal Process.*, vol. 59, no. 3, pp. 1170–1182, Mar. 2011.

[74] R. I. C. Chiang, G. B. Rowe, and K. W. Sowerby, "A quantitative analysis of spectral occupancy measurements for cognitive radio," in *Proc. IEEE Veh. Technol. Conf.*, Dublin, Ireland, 2007, pp. 3016–3020.

[75] T. Erpek, K. Steadman, and D. Jones, "Spectrum occupancy measurements: Dublin, Ireland, April 16-18, 2007," Shared Spectrum Company, Tech. Rep., Apr. 2007.

[76] T. Erpek, M. Lofquist, and K. Patton, "Spectrum occupancy measurements Loring Commerce Centre Limestone, Maine," Shared Spectrum Company, Tech. Rep., Sept. 2007.

[77] K. A. Qaraqe, H. Celebi, A. Gorcin, A. El-Saigh, H. Arslan, and M.-S. Alouini, "Empirical results for wideband multidimensional spectrum usage," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Tokyo, Japan, 2009, pp. 1262–1266.

[78] *Super-M Ultra Base Data Sheet*, MP Antenna, Ltd., Elyria, OH, no. 08-ANT-0861.

[79] *Coaxial Low Noise Amplifier*, Mini-Circuits, Brooklyn, NY, model No. ZX60-33LN+.

[80] "Aerial photograph of the University of Pretoria," Dec. 2011. [Online]. Available: http://g.co/maps/tzdhw

[81] T. Szepesi and K. Shum, "Cell phone power management requires small regulators with fast response," Analog Devices Inc., San Jose, CA, Tech. Rep., Feb. 2002.

APPENDIX A

FURTHER RESULTS

Table A.1: Required number of HMM training algorithm iterations.

	L	Theoretical	Freq. band A			Freq. band B		
VA								
Iterations	200	4.0	5.49	1.49	62.82	4.80	0.80	80.06
	300	4.0	15.14	11.14	-178.50	4.98	0.98	75.50
	400	4.0	5.65	1.65	58.83	5.16	1.16	71.00
BWA								
Iterations	200	106.8	77.35	-29.45	72.43	101.2	-5.60	94.76
	300	116.6	79.37	-37.23	68.07	91.84	-24.76	78.77
	400	126.5	90.43	-36.07	71.49	90.48	-36.02	71.53
PSO								
Iterations	200	263.90	263.18	-0.72	99.73	263.54	-0.36	99.86
	300	268.10	267.22	-0.88	99.67	267.39	-0.71	99.74
	400	269.60	267.84	-1.76	99.35	270.55	0.95	99.65
MA								
Iterations	200	329.76	341.40	11.64	96.47	336.80	7.04	97.87
	300	334.05	355.60	21.55	93.55	343.00	8.95	97.32
	400	333.74	362.80	29.06	91.29	351.40	17.66	94.71
			Meas.	**Diff.**	**Similar.**	**Meas.**	**Diff.**	**Similar.**