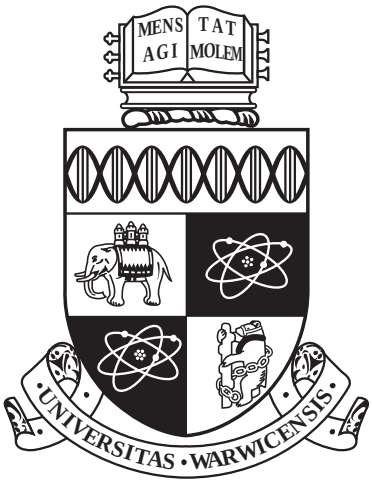


Machine learning and Energy Efficient Cognitive Radio

by

Freeha Azmat

Thesis

Submitted to University of Warwick in partial fulfilment of

the requirements for the degree of

Doctor of Philosophy

School of Engineering

October 2016

THE UNIVERSITY OF
WARWICK

Contents

Acknowledgments

First of all, I thank Almighty ALLAH for giving me strength and ability to complete this thesis. I would like to express my sincere and greatest gratitude to my supervisor, Dr. Yunfei Chen. This work would not have been possible without his generous guidance, suggestions and support. His endless patience and sharp intellect facilitated a highly productive relationship. I have been lucky to have him as my supervisor.

I would also like to sincerely acknowledge my second supervisor, Prof. Nigel Stocks, for providing me constructive advice and feedback related to the research project. It has been a pleasure to work with him and learn from such a kind and experienced individual.

Most importantly, I would like to thank my beloved parents and all family members for their unconditional sacrifices and constant encouragement. Without their boundless love, it would not have been possible for me to complete this journey. I would like to dedicate this thesis to them.

I specially thank Warwick Graduate School for awarding me Chancellors International Scholarship. My dream to accomplish PhD would not come true without this financial support.

I also would like to acknowledge all the fellows at Communication Networks (CoMNet) Lab for providing me inspirational research environment.

I would also like to thank my friends and my house mates who have encouraged me during tough times.

Last but not the least, I would also like to dedicate this thesis to the mem-

ory of my grandfather, Shabbir Hussain (1933-2011), who is a constant source of inspiration. I hope he would have been proud.

Declarations

This thesis is submitted in partial fulfilment for the degree of Doctor of Philosophy under the regulations set out by the Graduate School at the University of Warwick. This thesis is solely composed of research completed by Freeha Azmat, except where stated, under the supervision of Dr. Yunfei Chen and Prof. Nigel Stocks between the dates of January 2013 and October 2016. This thesis has not previously been presented in identical or similar form to any other examination board.

Freeha Azmat

October, 2016

Abstract

With an explosion of wireless mobile devices and services, system designers are facing a challenge of spectrum scarcity and high energy consumption. Cognitive radio (CR) is a promising solution for fulfilling the growing demand of radio spectrum using dynamic spectrum access. It has the ability of sensing, allocating, sharing and adapting to the radio environment. In this thesis, an analytical performance evaluation of the machine learning and energy efficient cognitive radio systems has been investigated while taking some realistic conditions into account. Firstly, bio-inspired techniques, including firefly algorithm (FFA), fish school search (FSS) and particle swarm optimization (PSO), have been utilized in this thesis to evaluate the optimal weighting vectors for cooperative spectrum sensing (CSS) and spectrum allocation in the cognitive radio systems. This evaluation is performed for more realistic signals that suffer from the non-linear distortions, caused by the power amplifiers. The thesis then takes the investigation further by analysing the spectrum occupancy in the cognitive radio systems using different machine learning techniques. Four machine learning algorithms, including naive bayesian classifier (NBC), decision trees (DT), support vector machine (SVM) and hidden markov model (HMM) have been studied to find the best technique with the highest classification accuracy (CA). A detailed comparison of the supervised and unsupervised algorithms in terms of the computational time and classification accuracy has been presented. In addition to this, the thesis investigates the energy efficient cognitive radio systems because energy harvesting enables the perpetual operation of the

wireless networks without the need of battery change. In particular, energy can be harvested from the radio waves in the radio frequency spectrum. For ensuring reliable performance, energy prediction has been proposed as a key component for optimizing the energy harvesting because it equips the harvesting nodes with adaptation to the energy availability. Two machine learning techniques, linear regression (LR) and decision trees (DT) have been utilized to predict the harvested energy using real-time power measurements in the radio spectrum. Furthermore, the conventional energy harvesting cognitive radios do not assume any energy harvesting capability at the primary users (PUs). However, this is not the case when primary users are wirelessly powered. In this thesis, a novel framework has been proposed where PUs possess the energy harvesting capabilities and can get benefit from the presence of the secondary user (SU) without any predetermined agreement. The performances of the wireless powered PUs and the SU has also been analysed.

Numerical results have been presented to show the accuracy of the analysis. First, it has been observed that bio-inspired techniques outperform the conventional algorithms used for collaborative spectrum sensing and allocation. Second, it has been noticed that SVM is the best algorithm among all the supervised and unsupervised classifiers. Based on this, a new SVM algorithm has been proposed by combining SVM with FFA. It has also been observed that SVM+FFA outperform all other machine leaning classifiers. Third, it has been noticed in the energy predictive modelling framework that LR outperforms DT by achieving smaller prediction error. It has also been shown that optimal time and frequency attained using energy predictive model can be used for defining the scheduling policies of the harvesting nodes. Last, it has been shown that wirelessly powered PUs having energy harvesting capabilities can attain energy gain from the transmission of SU and SU can attain the throughput gain from the extra transmission time allocated for energy harvesting PUs.

List of Publications

Published Journals:

1. **Freeha Azmat**, Yunfei Chen, Nigel Stocks, "Bio-inspired collaborative spectrum sensing and allocation for cognitive radios", **IET Communications**, vol. 9, no. 16, May 2015.

2. **Freeha Azmat**, Yunfei Chen, Nigel Stocks, "Analysis of spectrum occupancy using machine learning algorithms", **IEEE Transactions on Vehicular Technology**, vol. pp, no. 99, Oct 2015.

3. **Freeha Azmat**, Yunfei Chen, Nigel Stocks, "Predictive Modelling of RF energy for Wireless Powered Communications", **IEEE Communication letters**, vol. 20, no. 1, Nov 2015.

To be submitted:

1. Freeha Azmat, Yunfei Chen, Nigel Stocks, "Performance Analysis of Wireless Powered energy harvesting cognitive radio networks".

2. Freeha Azmat, Yunfei Chen, Nigel Stocks, "Spectrum prediction using machine learning techniques for cognitive radio networks".

List of Figures

List of Tables

Abbreviations

AP	Access point
APD	Amplitude probability distribution
AI	Artificial intelligence
ABC	Artificial bee colony
AWGN	Additive white Gaussian noise
CA	Classification Accuracy
CR	Cognitive radio
CRN	Cognitive radio network
CAF	Cyclic autocorrelation function
CSD	Cyclic spectral density
CSGC	Color-sensitive graph coloring
CSS	Cooperative spectrum sensing
DSA	Dynamic spectrum access
DT	Decision trees
DL	Downlink
DFS	Dynamic frequency selection
EH	Energy harvesting
EGC	Equal gain combining
EWMA	Exponentially weighted moving average
FFA	Fire Fly algorithm
FCC	Federal communications commission
FSS	Fish School Search

GA Genetic algorithm

GMM Gaussian mixture model

HMM Hidden markov model

HPA High Power Amplifier

ICT Information and communication technology

KNN K-nearest neighbour

KBL Kernel based learning

LEH Low efficiency harvester

LR Linear regression

MEH Medium efficiency harvester

MA Moving average

MAC Media access control

ML Machine learning

MMR Maximum minimum reward

MPF Maximum proportional fair reward

MSR Maximum Sum Reward

NF Noise figure

NBC Naive Bayesian classifier

NRMSE Normalized Root mean square

Ofcom Office of communications

OC Optimal combining

OSA Opportunistic spectrum access

OPEX Operational expenditures

PA Power amplifier

PSO Particle swarm optimization

PU Primary User

PCE Power conversion efficiency

QoS Quality of service

QGA Quantam genetic algorithm

QAM Quadrature amplitude modulation

QPSK Quadrature phase shifting key

RF Radio frequency

SPTF Spectrum policy task force

SNR Signal-to-noise ratio

SUR Secondary user receiver

SUTR Secondary user transmitter

SVM Support vector machine

SWIPT Simultaneous wireless information and transmit power

TPC Transmit power control

UL Uplink

UHF Ultra-high frequency

VHF Very high frequency

WLC Weighted linear combining

WPC Wireless powered communication

WPT Wireless power transfer

WRAN Wireless regional area network

WCMA Weather-conditioned moving average

Important Symbols

Box_{ct} Box constraints in SVM

con^i Consecutive free frequency bins in the ith time slot

$\mathbf{E}[PE]$ Expected harvested energy at the PU

$\mathbf{E}[PE]$ Expected harvested energy at the SUTR

$\mathbf{E}[R_{SUR}]$ Expected throughput of the SUR

γ Signal to noise ratio of the dth radio

$\Gamma(.,.)$ Incomplete Gamma function

$\mathbf{H}_{train}^{i,m-1:m-q}$ Training feature vector for energy harvesting

$H_{train}^{i,m}$ Target label for energy harvesting

λ Decision threshold

L_{oc} Minimum value of occupancy

OC^i Occupancy of the ith time slot

OC^j Occupancy of the jth frequency bin

P^i_{eval} PU occupancy predicted by the classifier

P_{fa} Probability of false alarm

P_d Probability of detection

P_i Power of the primary user

P_0 Power of an AP signal

P_s Power of the SUTR signal

$Q(x)$ Tail probability of the standard normal distribution

$Q_u(a, b)$ The generalized Marcum Q-function

$S^i(j)$ Spectrum status of the PU signal

U_{oc} Maximum value of occupancy

$\mathbf{w}$ Optimal weighting vector

$x^i(j)$ Input PU signal

z_{TR} Noise at SUTR

Chapter 1

Introduction

1.1 Introduction

With the explosive development of wireless products and mobile internet applications, the demand of radio frequency (RF) spectrum has been constantly increasing. The vast majority of the frequency spectrum is restricted to licensed-only access under the current policies defined by the spectrum regulating government bodies, such as the office of communications (Ofcom) and the federal communications commission (FCC) in United kingdom and United States, respectively. The wireless service providers can purchase the license to exclusively utilize a band of certain frequencies within a large geographic region. Under the current policies, the license holders have the right to use the licensed spectrum only and it is forbidden to utilize the other spectrum frequencies regardless of their occupancy status.

Due to the mentioned government regulations, the spectrum scarcity is becoming a critical issue. On the contrary, recent spectrum occupancy measurement studies are illustrating a different story about the spectrum utilization. Interestingly, the spectrum policy task force (SPTF) within the FCC has reported that localised temporal and geographic spectrum utilisation effi-

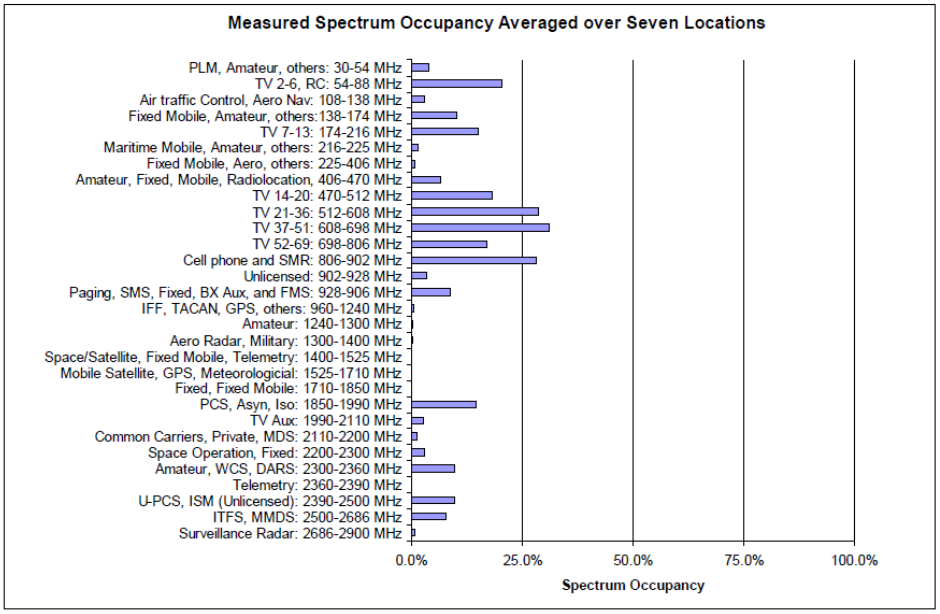


Figure 1.1: A bar chart illustrating the spectrum occupancy measurements evaluated between 30 MHz to 3 GHz for seven different locations in the United States. Adapted from [2].

ciency ranges between 15% to 85% [1]. Following this, the spectrum occupancy measurements were done in different counties in order to attain a clear picture of the actual spectrum utilization in the urban environment. The main objective of all these studies was to find those spectrum bands that have low occupancy and have more chances to be reused for other purposes.

A measurement study was done in [2], where the average spectrum utilization of each frequency band between 30 MHz to 3 GHz is analysed for seven different locations in the United States as shown in Fig. 1.1. It was observed in Fig. 1.2 that majority of the cities have attained average occupancy less than 25%.

In [3], a similar measurement campaign was done in Germany. It was observed that spectrum is nearly 100 % utilized between 20 MHz - 3 GHz and sparsely utilized between 3 GHz -6 GHz for outdoor locations. However,

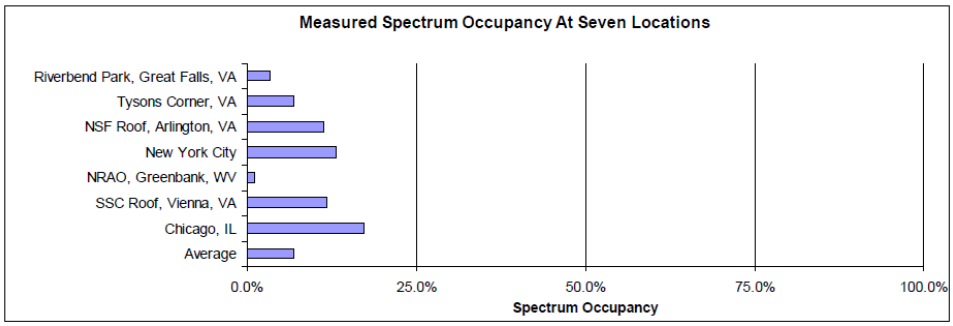


Figure 1.2: Overall spectrum occupancy measured at seven different locations in the United States. Adapted from [2].

the spectrum occupancy for indoor locations is around 32% between 20 MHz - 3 GHz. Another measurement study took place in Singapore in [4] for the frequency bands between 80 MHz - 5.85 GHz, where it was found that GSM 900 is the busiest band. It was also noticed that radar bands and ISM bands have low spectrum occupancy and contain abundant spectrum opportunities.

A survey of the spectral usage was also conducted in Kansas, USA [5]. The measurements were done between 9 kHz - 1 GHz as shown in Fig. 1.3. It was observed that majority of the spectrum is sparsely utilized. The white spaces in Fig. 1.3 can be reused to increase the spectrum utilization. Another study was performed in the outdoor environment at Barcelona between 75 MHz - 3 GHz for two days. The average spectrum occupancy was observed to be 22.57 % between 75 MHz - 3 GHz. It was also observed that cellular bands: GSM 900 and UMTS have very low utilization of about 2.86 % [6].

In all these studies, it was concluded that spectrum scarcity is caused due to two reasons (a) static spectrum allocation policies defined by the regulatory bodies and (b) the inefficient utilization by the license holders. In order to bridge the gap between spectrum underutilization and spectrum scarcity, a novel, adaptable and efficient communication paradigm called cognitive radio,

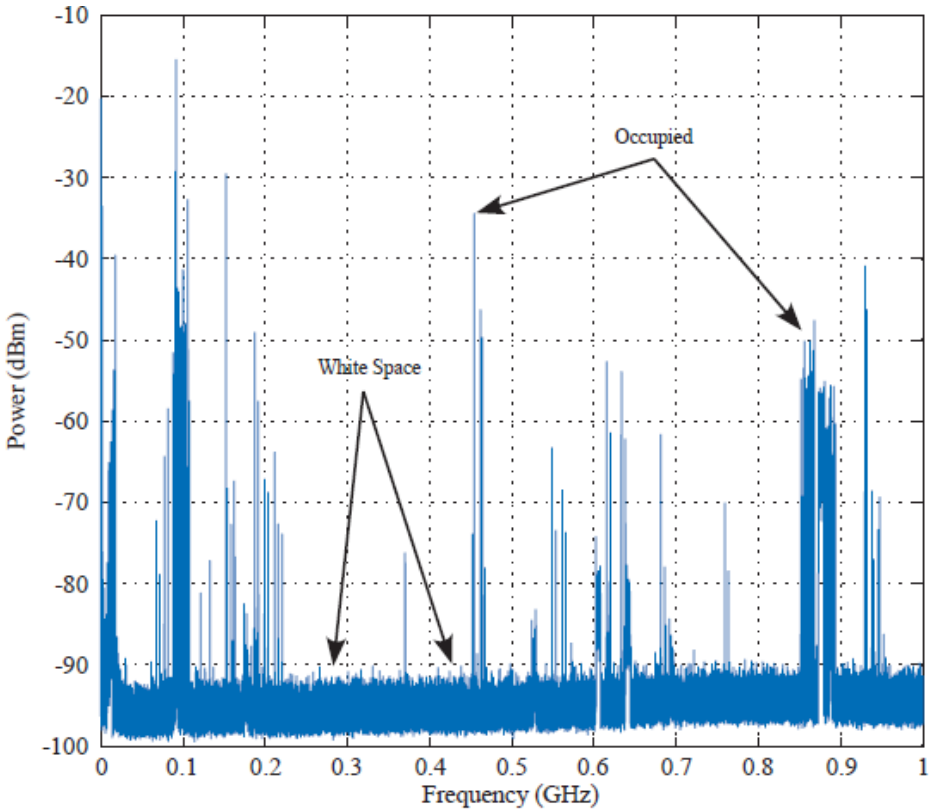


Figure 1.3: The relationship between the average power (dBm) and the frequency range between 9 kHz - 1 GHz using the spectrum occupancy measurements attained at Lawrence, Kansas, USA. Adapted from [5].

was proposed by Mitola [7]. A complete overview of the CR technology is presented in Section 1.2.

1.2 Cognitive Radio

In Mitola's proposed CR network, there is a primary user (PU) who has license (permission) to work in a certain spectrum and a secondary user (SU) who has no spectrum license. CR enables the SU to build transmission links using vacant PU channels such that there is minimum interference between primary and secondary users. It is the vital technology that allows using the spectrum

in a dynamic manner. Cognitive radio (CR) is formally defined in [8] as:

Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), learns from the environment and adapt its internal states according to the statistical variations in the incoming RF stimuli in real-time, with two primary objectives in mind (a) highly reliable communications whenever and wherever needed and (b) efficient utilization of the radio spectrum.

There are two main characteristics of CR that can be derived from this definition:

- Cognitive capability: refers to the ability of radio technology to attain information from the radio environment.
- Reconfigurability: refers to the dynamic programming of the radio according to the environmental conditions [9].

The cognitive capability and reconfigurability together are responsible for attaining self-healing and adaptable cognitive network architecture. It is further explained in Section 1.2.1.

1.2.1 Overview of cognitive radio networks

In this section a brief overview of the cognition cycle, cognitive radio network (CRN) topologies, spectrum sharing modes and the spectrum access techniques is presented.

Cognition Cycle

There are four main functionalities of CR that must be performed continuously in order to ensure the maximum protection to PU activities. The tasks in a *cognition cycle* are shown in Fig. 1.4 and explained as

- Spectrum sensing: refers to the detection of unused spectrum spaces without creating harmful interference to the primary users. Spectrum sensing classifies the spectrum space either as
  - *white space*, one which is completely empty, except for noise. The white spaces are also referred as spectrum holes in literature
  - *gray space*, one which is partially occupied by interfering signals
  - *black space*, one which is fully occupied by the communication signals, interfering signals and the noise.
- Spectrum allocation: refers to the selection of the best available spectrum channels according to the user communication needs.
- Spectrum mobility: refers to fulfilling the user requirements for providing seamless communication while switching the spectrum channels.
- Spectrum sharing: refers to those scheduling methods that ensure fair spectrum allocation among coexisting users.

Network topologies of cognitive radio

Cognitive radio networks can have centralized or distributed topology. The centralized topology is the one, where a central node coordinates spectrum sensing, allocation and management among all other nodes. On the contrary in distributed topology, the nodes communicate with each other in an adhoc manner. The CRN network topologies can be chosen according to the desired application as each topology has its own advantages and disadvantages. The distributed topology has more computational complexity per node but has less infrastructure costs. However the computational load on each node is lesser

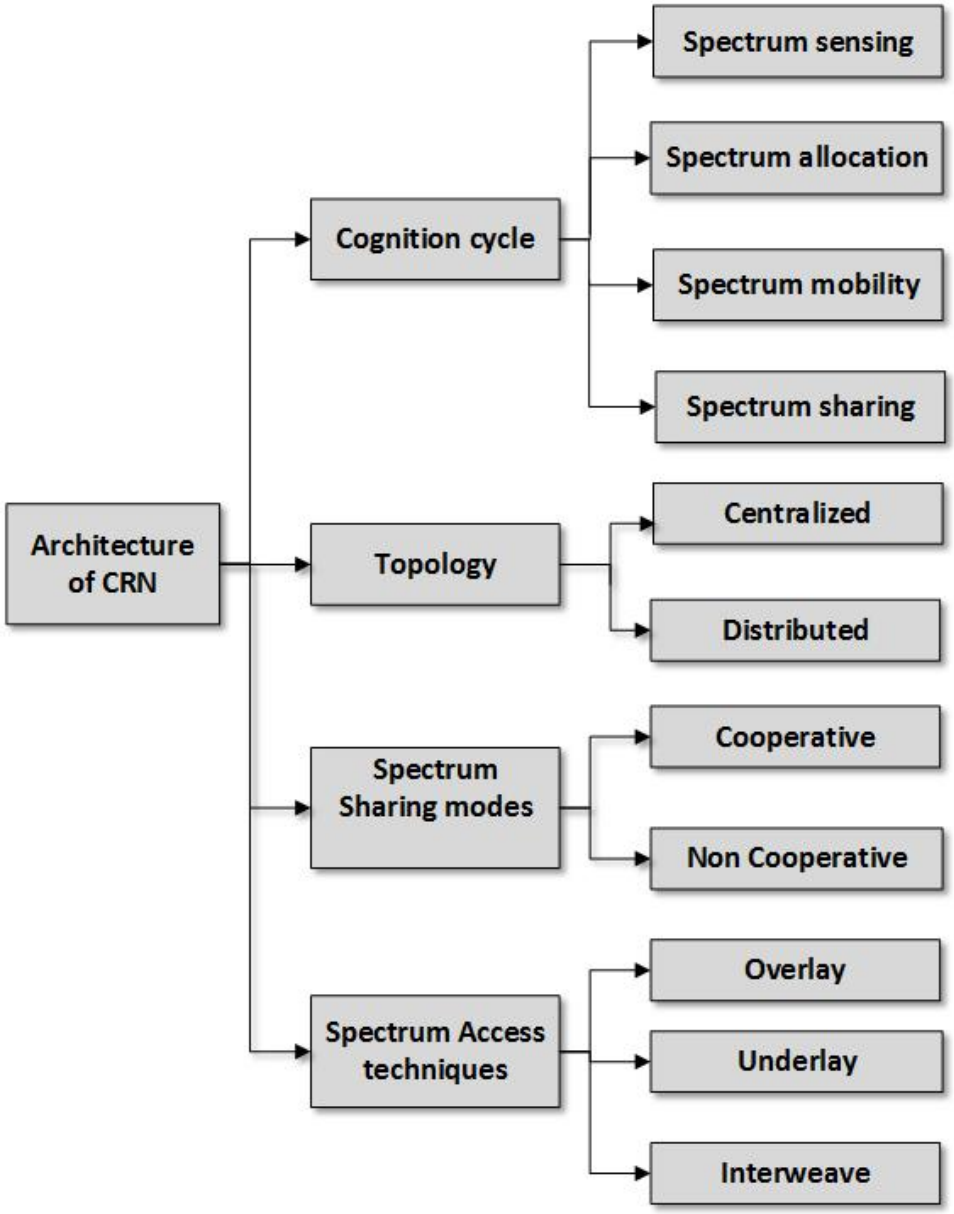


Figure 1.4: Architecture of the cognitive radio networks that comprises of a cognition cycle, network topology, spectrum sharing and spectrum access modes. Adapted from [10]

in the case of centralized topology but it incurs more infrastructure costs as each node shares its sensing results to the central controller [11].

Spectrum Sharing modes

Using either distributed or centralized topology, the results of the spectrum sensing schemes are severely degraded in reality due to multipath fading and shadowing. In order to improve the detection performance, cooperative spectrum sensing (CSS) is encouraged in literature [12], [13]. In cooperative spectrum sensing, the spectrum measurements are evaluated by a group of CRs and the variability of the signal strength rely on various locations. In other words, the probability that all users experience the same fade is very low which eventually decreases the interference to PU. It is expected that the sensing results for a group of CRs would be better than individual sensing. On the other hand, CSS has its own disadvantages as it is more complex than individual sensing due to the increased control traffic between the nodes. Also, the delays incurred due to combining results decreases the time for data transmission in the case of CSS.

Spectrum access techniques

The main objective of cognitive radio is to utilize the spectrum spaces using dynamic spectrum access (DSA). DSA refers to the process of locating those frequency bands and time slots, where CR can send/receive data without causing any degradation to the performance of the PUs [14]. In order to cause minimum interference to the PU, CR adopts one of the following spectrum access techniques given as

- Interweave Access: it is often termed as the interference avoidance mode. The SU finds spectrum holes using sophisticated spectrum sensing tech-

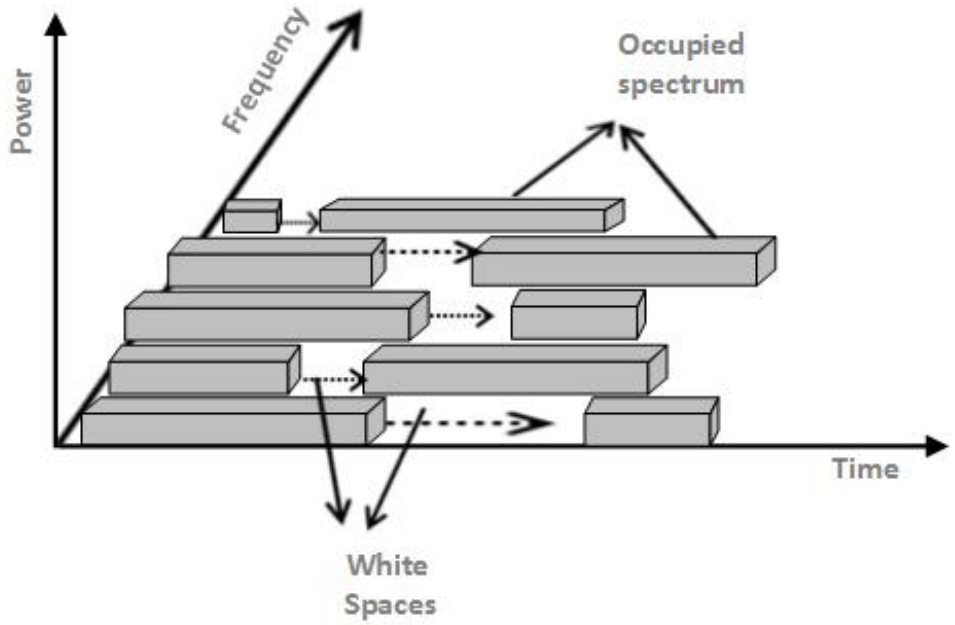


Figure 1.5: Dynamic spectrum access in an interweave mode. Adapted from [1].

niques and uses them for data transmission. The SU should leave the channel, when the PU reappears. The concurrent data transmission from both SU and PU is not allowed. This leads to the forced termination of the SU connection (if there is no other available channel for the SU). This scheme is often named as opportunistic spectrum access (OSA) as shown in Fig.1.5.

- Underlay Access: refers to a technique where SU shares the spectrum with PU, as long as its signal remains below the interference temperature. The acceptable interference temperature refers to the average power that can be tolerated as an interference by the PU receiver [15], [16]. This term was introduced by FCC for 'quantifying and managing interference' [17] and is given as

$$T_L = \frac{P_L}{kW} \tag{1.1}$$

where P_L represents the average interfering power in bandwidth W, k represents the Boltzmann constant and $k = 1.38 * 10^{-23} J/K$. If T_L is set for a specific frequency bandwidth W, then SU should ensure that average interference power it generates remains below kWT_L.

- Overlay Access: refers to a technique where SU has the knowledge of the PU's transmitted data sequence (message) and how the data sequence is encoded (codebook). The knowledge of the PU's codebook can be exploited by SU in various ways to improve the performance of both SU and PU. This access mode is often termed as CR enabled cooperative relaying [18]. The difference between underlay and overlay modes is illustrated in Fig.1.6 and the comparison of three spectrum sharing modes is presented in Table 1.1.

Using the knowledge of spectrum access techniques, it can be concluded that white spaces can be used for interweaving that let PUs to operate in unused regions, black spaces for overlaying where signals are processed in a manner that makes the quality of transmission unimpaired by the SU and grey spaces for underlying where SU ensures to keep the interference towards PUs at a tolerable level.

1.2.2 IEEE Standard: 802.22

Currently, several cognitive radio systems are being finalized by standardization bodies around the world. In the United States, the IEEE standardization committee has worked on IEEE 802.22, where CR system is dealt as a wireless regional area network (WRAN). The WRAN architecture is proposed for very high frequency (VHF) and ultra-high frequency (UHF) ranges between 54-869 MHz. The standard has been finalized in July 2011 and it was developed to

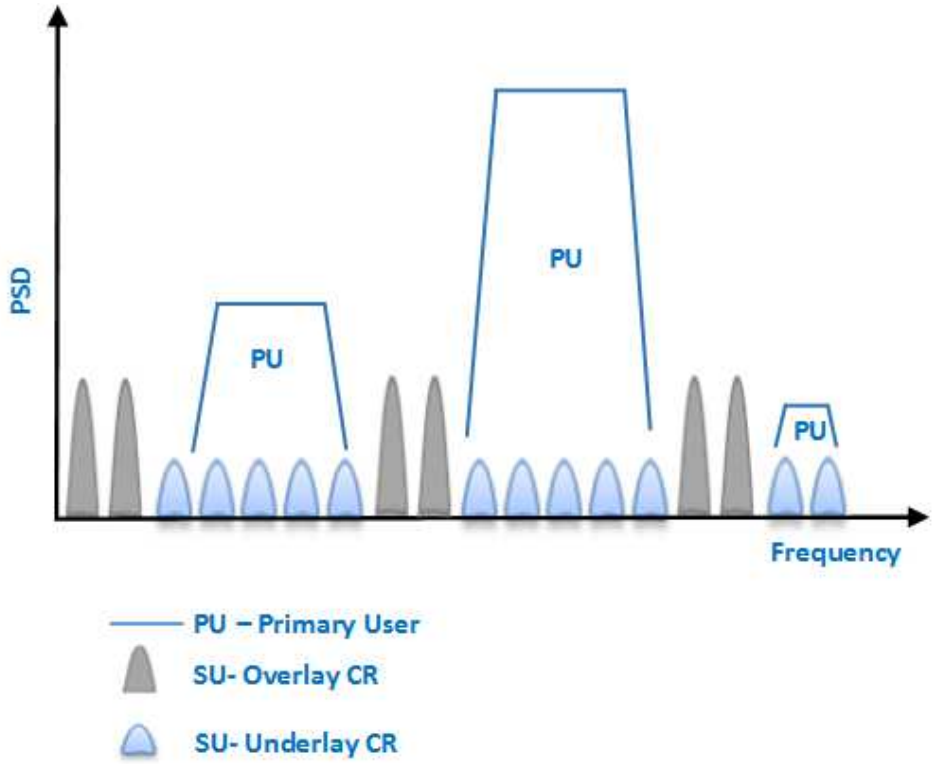


Figure 1.6: Overaly and underlay spectrum sharing modes in cognitive radio. Adapted from [5]

access the unused bands in the TV white spaces, where PUs represent the analogue digital TV channels and low power wireless microphone signals [19].

The operation of the secondary user in IEEE 802.22 WRAN is divided into multiple consecutive super-frames in the time domain, where each super-frame is further subdivided into multiple MAC frames. Each super-frame has 16 media access control (MAC) frames, where each MAC frame is of 10 ms duration which is further subdivided into spectrum estimation and secondary data transmission slot as shown in Fig. 1.7. The super frame structure is proposed for ensuring PU protection and managing secondary data transmission effectively [20].

IEEE 802.22 proposes centralized architecture, where each SU in the WRAN senses the spectrum to analyse the presence of the PU and sends the sensing results to the fusion centre (802.22 base stations). The fusion centre accumulates all the sensing results from various secondary users and makes a final decision regarding the presence/ absence of the PU. The final sensing result will be sent to the SUs by the base station.

Though IEEE 802.22 has been proposed as the first CR based international standard, there are many other IEEE standards such as IEEE 802.11, IEEE 802.15 and IEEE 802.16 that include some degree of CR technology for attaining dynamic spectrum access and coexistence. Many of these standards include dynamic frequency selection (DFS) and transmit power control (TPC) for the purpose of facilitating spectrum sharing. There are many features in IEEE 802.22 that are adopted from IEEE 802.16 and IEEE 802.16e standards such as physical, MAC and quality of service (QoS) features. The frame structure of IEEE 802.22 is also an extension of the IEEE 802.16 and IEEE 802.16e frame structures [21]. A comparison of IEEE 802.22 standard with other popular standards such as IEEE 802.11, IEEE 802.15 and IEEE

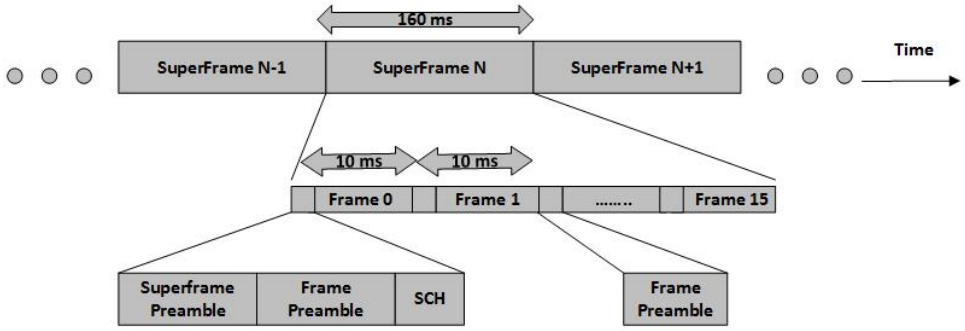


Figure 1.7: A Superframe structure having a duration of 160 ms, where each superframe has 16 frames and every individual frame consumes 10 ms. Adapted from [20].

802.16 is presented in Table 1.2.

1.2.3 Potential applications of cognitive radio

Although the most common application of CR is in the TV white spaces; where CR enabled SUs opportunistically utilize the unused spectrum without interfering with PUs of spectrum, namely TV transmitters. Apart from this, they can be used for the following:

- Cognitive radios can possibly be used to provide 'data boost' by opportunistically offloading the traffic of the existing PUs to the white spaces. In this manner, CRs can help to alleviate high loads on the cellular network and meet the quality of service (QoS) requirements of the delay sensitive traffic like voice, streaming and video etc [23].

- The usage of wireless technologies like Bluetooth, Infrared and WiFi has been increasing in every home and it has eliminated the need of wired communication. The problem associated with majority of these technologies is that they are not designed to operate in coherence with each other and causes interference to the receivers, which are not part of their network. A central controller is required for the coexistence

of several wireless technologies. A cognitive digital home can operate on white spaces in the unlicensed bands to carry control information between devices [24], [25]. A cognitive home can be the solution for all the interference issues due to multiple competing technologies.

- CR can be used for vehicle to vehicle communication networks, where each vehicle can have CR and form a peer to peer network with other vehicles around. This configuration can be used for sending traffic alerts and geographic applications services.

Feature	Underlay	Overlay	Interweave
Network side Information	SUs know the interference caused to primary receivers	SUs know channel gains, encoding techniques and the transmitted data sequence of the PUs.	SUs identify spectrum holes where PUs are absent.
Simultaneous transmission	SUs transmit simultaneously with PU as long as the interference caused is below an acceptable limit	SUs cooperate with PUs and relay PUs data in order to coexist with PUs.	SUs transmit simultaneously with a PU only when there is missed detection of the PU activity.
Transmit power limits	SUs transmit power is limited by a constraint on the interference caused to the PUs	SUs can transmit at any power and the interference to PUs can be offset by relaying the PUs data.	SUs transmit power is limited by the range of PU activity it can detect (alone or via cooperative sensing)
Hardware	SUs must measure the interference they cause to PUs by either exploiting channel reciprocity or via cooperative sensing	SUs must also listen to PUs. Encoding and decoding complexity is also higher than other paradigms.	Receiver must be frequency agile or have a wideband front end for spectrum hole detection.

Table 1.1: Comparison of underlay, overlay and interweave spectrum access techniques in cognitive radio techniques. Adapted from [23].

Network type	Industry standards	Frequency	Max data rate	Range
PAN	IEEE 802.15	2.4 GHz	10 Mb/s	20-50 m
LAN	IEEE 802.11 a	5 GHz	54 Mb/s	20 m
LAN	IEEE 802.11 b	2.4 GHz	11 Mb/s	33m
MAN	IEEE 802.16	¡ 2.4 GHz	54 Mb/s	1-2 km
RAN	IEEE 802.22	54 - 862 MHz	6 - 31 Mb/s	30 km

Table 1.2: The comparison of IEEE 802.22 wireless radio access network (WRAN) to the other popular wireless standards. Adapted from [22].

1.3 Thesis Outline

1.3.1 Research Motivation

Although cognitive radio networks have been extensively researched as a potential candidate for mitigating the spectrum scarcity and increasing spectrum utilization, there are still a lot of problems that need to be addressed.

First of all, in order to utilize the spectrum effectively with minimum interference to PUs, the SU needs to sense the spectrum efficiently with less chances of error. The reliable spectrum sensing is only guaranteed using the conventional spectrum sensing techniques when signal-to-noise ratio (SNR) is high, however the detection performance degrades, otherwise. For combating this problem, the cooperative spectrum sensing is encouraged. In CSS, the local sensing is performed at each SU and the individual results are sent to the data fusion centre via common control channel [26]. A data fusion centre combines energy measurements from all cooperating cognitive radios to make a final detection decision. In order for CSS to be operative, the local measurements sensed by each radio should be weighted according to their reliability during data combining. This is because; the received SNR value at each radio can vary intensely due to path loss and shadowing in realis-

tic scenarios. It was observed in [27] that CSS outperforms the standalone energy detector. Similarly [28] derived optimal and sub-optimal weights for a linear combination of measurements in the data fusion centre. A popular weighted energy combining method using CSS is proposed in literature known as 'weighted linear combining (WLC)'. This scheme determines the optimal weighting vector using a heuristic technique proposed in [29] which minimizes the probability of detection error. However there is still a need for investigating self-managing, self-configuring energy based combining techniques that can adapt according to the time varying nature of the wireless channels. For that, the bio-inspired approaches, with appealing features like self-adaptation, autonomy and collaborative decision making abilities, can be investigated to address the complexity of the CSS systems. Furthermore, the optimal weights attained for CSS in literature deals with the linear PU signals, however in reality, the PU signal may suffer from non-linear distortions. Therefore, there is a need to consider both linear and non-linear PU signals for CSS analysis. Apart from spectrum sensing, there is a need to investigate spectrum allocation method based on spectrum sensing results that can ensure conflict free spectrum allocation. A common method used for spectrum allocation in literature is color-sensitive graph coloring (CSGC) [30]. Three evolutionary algorithms are presented in [31] that outperform CSGC by attaining higher value of the spectrum allocation rewards compared to CSGC. However there is a still a need to analyse the self-adaptable algorithms that can converge quickly and can attain higher value of the spectrum allocation rewards at the same time.

Secondly, many studies have been performed to understand the spectrum occupancy statistics. For instance, the statistical and spectral occupation analysis of the spectrum measurements was presented in [32] in order to study

the traffic density in all frequency bands. In [33], auto-regressive model was used to predict the radio resource availability using occupancy measurements in order to achieve uninterrupted transmission of the secondary users. Similarly, in [34] - [36], the occupancy statistics were utilized to select the best channels for control and data transmission purposes so that less time is required for switching transmission from one channel to the other for the case, when the PU appears. All of the aforementioned works have evaluated the spectrum occupancy models by using conventional probabilistic or statistical tools. These tools are often limited to the assumptions required to derive their theories. For example, one has to determine whether the value is a random variable or a random process in order to use the probabilistic and statistical tools. Therefore, there is a need to investigate spectrum occupancy using those techniques that do not have prerequisites on data. The correct modelling of the spectrum occupancy can yield to better spectrum sensing.

Furthermore, the energy consumption due to radio frequency (RF) devices is increasing exponentially with an increase in the usage of wireless applications. It is reported that the energy consumption of the information and communications technology (ICT) infrastructure is increasing 16 - 20 % approximately per annum and generates about 2 % of the worldwide CO_2 emissions [37], [38]. Therefore, it is important to optimize the energy efficiency of the wireless networks as it will not only cut the overall cost of the network but will also decrease the adverse effects on the environment. Energy harvesting devices could be a potential source of energy. In particular, radio frequency (RF) energy harvesting is an upcoming technology that allows ambient RF signals to be collected by an antenna and converted into DC power using a rectifier [39]. For making CRN efficient, it was proposed in [40] to allow secondary user to harvest energy using renewable energy sources. Powering

mobile devices using harvested energy from ambient sources such as solar, wind, and kinetic activities makes wireless networks not only environmentally friendly but also self-sustaining. On the other hand, the amount of RF energy that could be harvested changes with time and frequency. For example, there are more mobile signals during the day than during the night time in commercial areas. Thus, it is very important for RF energy harvesters to choose the right operating time and frequency for harvesting maximum energy.

Last, but not least, energy harvesting in CRN is a hot research topic and the majority of the literature investigates those energy harvesting CRN, where the SU harvests energy from the nearby PU or transmits information, if the PU is far away [41], [42]. There are also few studies that encourages cooperation between PU and SU [43]- [45]. All these works assumed that the PU does not have any energy harvesting capability. Therefore, there is a need to investigate the framework where PU can also have energy harvesting capability and can get benefit from the presence of SU.

1.3.2 Chapter Outlines

Motivated by the above observations, the performance of the cognitive radio network is analyticity evaluated in this thesis, while taking several realistic conditions into account. Numerical results and discussion are presented to evaluate the performance of the algorithms. The introduction and conclusion sections are presented in each chapter to provide readers with overall summaries of the chapters. The thesis is organized as follows.

In Chapter 2, a comprehensive overview of the spectrum sensing algorithms is presented. Then, a detailed description of the machine learning (ML) framework and bio-inspired techniques is presented. Also, the previous works that deploy machine learning and bio-inspired algorithms in the area of

the cognitive radio network is illustrated. Finally, an overview of the energy harvesting techniques and wireless energy transfer methods are discussed.

In Chapter 3, a collaborative spectrum sensing and allocation framework using bio-inspired techniques is proposed which provides an optimal weighting vector for the data fusion centre. Three bio-inspired algorithms: firefly algorithms (FFA), fish school search (FSS) and particle swarm optimization (PSO) are used in this chapter, where FFA and FSS have not been used for both collaborative spectrum sensing and allocation before. In order to analyse the practical CSS systems, both linear non- linear primary user signals with interference and fading losses are considered. Furthermore, a spectrum allocation approach based on the optimal weighting vector evaluated by the data fusion centre is proposed that outperforms the conventional method in literature by attaining higher value of the spectrum allocation awards.

In Chapter 4, the machine learning algorithms are utilized to investigate and classify the spectrum measurements taken at the University of Warwick in order to gain an insight of the spectrum occupancy. Machine learning plays a vital role in the artificial intelligence (AI) field, where algorithms are implemented to analyse data using past experiences. Machine learning refers to a paradigm, where an algorithm learns from the inputs in a manner that its expected future performance improves. The machine learning algorithms are often heuristic, as they do not have any prerequisites or assumptions on data. As a result, in many cases, they provide higher accuracy than conventional probabilistic and statistical tools. Three supervised algorithms, naive Bayesian classifier (NBC), decision trees (DT), support vector machine (SVM), and one unsupervised algorithm, hidden markov model (HMM) are used in Chapter 4 to classify the occupancy status of the time slots. The classified occupancy status is further utilized for evaluating the blocking probability. Furthermore,

a new technique that combines SVM with fire fly algorithm (FFA) is also proposed that outperforms all supervised and unsupervised algorithms.

In Chapter 5, the energy prediction framework is proposed for choosing the optimal frequency and time for harvesting maximum energy. Most current works assume a theoretical framework for analysing energy profile, where majority of the models can be utilized for the management of energy efficient networks using the given set of conditions. However none of the previous works considered the real time measurements from practice to estimate the amount of energy, that can be harvested using the prior knowledge. Two machine learning techniques: linear regression (LR) and decision trees (DT), are employed in Chapter 5 to predict the optimal frequency and time for harvesting the maximum energy using low and medium-efficiency harvesters.

In Chapter 6, the performance of an energy harvesting CRN is analysed, where it is assumed that PU is a wireless powered system with energy harvesting capability. In wireless powered communication (WPC), the power is sent first to the PU from an access point (AP) and further PU uses the harvested power for sending the information. It is proposed in Chapter 6 that SU sends data while PU is harvesting power from the AP. In this manner, PU gets chance to harvest some energy from the SU data transmission as well. Using this notion, PU gets benefit from the SU without allocating extra resources. Also, SU coexists with the PU and changes its strategies in a manner that bring advantages to the PU. The proposed strategy is compared with the conventional energy harvesting CRN, where PU does not have any energy harvesting capabilities. It is shown in Chapter 6 that proposed methodology outperforms the conventional energy harvesting CRN.

Finally in Chapter 7, the research results and analysis are summarised. Also, the future work directions and suggestions are presented.

Chapter 2

Background and Literature Review

2.1 Introduction

With an increasing usage of dynamic mobile applications, it is becoming important for wireless devices to learn from environment. Cognitive radio (CR) is one of the pioneer technologies that propose to learn from surroundings. Apart from the environment awareness, CR should have capability to remember, judge and analyse the given situation. This will help CR to forecast the network resources by analysing network load, location and user's mobility. Learning from previous experiences can make the decision process and adjustment of transmission/ reception parameters faster in the CR systems. An intelligent CR architecture has three important stages [46]:

- *Perception*: refers to the ability of sensing environment and acquiring data. The fundamental block to attain environmental awareness in CR systems is spectrum sensing. The aim of the spectrum sensing is to

detect spectrum holes. An extensive overview of the spectrum sensing techniques is presented in Section 2.1

- *Learning*: refer to those steps that help to extract knowledge using the information acquired from the RF environment For example, the PU status (presence/ absence) is learnt using the data attained from the environment.

- *Reasoning*: that deals with the ability to use the knowledge for taking actions according to the given policies and conditions. Spectrum allocation is an example of reasoning, where the frequency channel would be allocated to SU after evaluating PU status. The machine learning and bio-inspired intelligent algorithms are utilized in this thesis for learning and reasoning. They are further explained in Section 2.2.

In order to acquire perception, learning and reasoning, a lot of computations are required to be done by CR that consumes energy. For making CRs energy efficient, a novel energy harvesting approach is also presented in this thesis. The overview of the energy harvesting cognitive radio networks and the wireless energy transfer methods is presented in Section 2.3.

2.2 Perception

The process of getting awareness about the PU signals in a given geographical area at a given time is referred as perception. This is a challenging task, because the detection should be done at SUs independently without causing any external interference to the PUs [23]. In the simplest form, the received signal at SU ($y^i(j)$) can be written as

$$y^i(j) = x^i(j) + w^i(j) \tag{2.1}$$

where $x^i(j)$ represents the PU signal (x(t)=0, when PU is not transmitting) and $w^i(j)$ represents the additive noise respectively. Specifically, it is assumed that

$$w^i(j) = g^i(j) + f^i(j) \tag{2.2}$$

where $g^i(j)$ and $f^i(j)$ represents the white Gaussian noise and additive interference, respectively. A single detector must decide between the two hypotheses

$$y^i(j) = \begin{cases} w^i(j), & H_0, \\ x^i(j) + w^i(j), & H_1, \end{cases} \tag{2.3}$$

The spectrum sensing performance is measured using two parameters

- Probability of false alarm: $P_{fa} = P_r(H_1|H_0)$, that refers to the probability when SU mistakenly detects the presence of PU, however it is absent.
- Probability of detection: $P_d = P_r(H_1|H_1)$, represents the probability when SU detects the presence of PU correctly.

These two probabilities are not independent, however they are linked by a function, which illustrates the detector performance, called as receiver operating characteristic (ROC). The ROC curve expresses the dependency between P_{fa} and P_d for a given detector [47], [48]. Some authors also use, the probability of missed opportunity, $P_{mo} = P_r(H_1|H_0) = P_{fa}$ and the probability of mis-detection $P_{md} = P_r(H_0|H_1) = 1 - P_d$, as the performance metric in the literature. Using P_{fa} and P_d, three classes of spectrum sensing systems can be defined given as:

- **conservative**, if $P_{fa} > 0.5$ and $P_{md} < 0.5$
- **aggressive**, if $P_{fa} < 0.5$ and $P_{md} < 0.5$
- **hostile**, if $P_{fa} < 0.5$ and $P_{md} > 0.5$.

A conservative system has small probability of interference with PU as it has high value of P_{fa}, hence small spectrum utilization rate. An aggressive system has high spectrum utilization rate and small probability of interference. A hostile system has low value of P_{fa}, with a large probability of interference [23]. Depending on the system model and the sensing quality, the sensing algorithms can be chosen. There are some sophisticated spectrum sensing methods used in the literature discussed as follows:

2.2.1 Energy Detection

In this method, the energy of the received signal is compared with a decision threshold (λ), determined by the noise level [49]. Let $y^i(j)$ represents the received signal at the SU, the decision metric Y can be computed using N received samples, given as

$$Y = \frac{1}{N} \sum_{i=0}^{N-1} |y^i(j)|^2 \tag{2.4}$$

where y_n represents the received signal and it can be in time-domain or frequency domain. The binary testing hypothesis using energy detection can be formulated as

$$status = \begin{cases} Y < \lambda, \;\; H_0, \\ Y > \lambda, \;\; H_1, \end{cases} \tag{2.5}$$

There is no fading assumed in this model except the additive noise that is independent, circularly symmetric zero-mean complex Gaussian signal with

variance $\mathbf{E}[|w^i(j)|^2]$. The performance of the energy detector is dependent on the estimation of noise variance in the model. This is because, the test statistic (Y) is compared with decision threshold, where decision threshold is dependent on the observed signal model, hence on the noise variance.

Energy detector has low detection performance when the noise variance is unknown to the receiving node. An error on the estimate of the noise variance affects the performance of the detector, which becomes vulnerable to noise power inaccuracies. When the signal-to-noise ration is very low, it would be hard to distinguish between the radio signal and noise signal, therefore the prior knowledge of the noise power can be used to improve the detection performance of the energy detector. If the noise power level is perfectly known at the receiver, the energy detector can work with arbitrary values of probability of detection and probability of false alarm, even in low SNR regimes, by using a sufficiently long observation time.

In summary, it can be concluded that energy detection has low computational complexity and works well if two prerequisites are already been achieved (a) the noise must be statistically stationary and (b) noise variance is known to the detector [50].

2.2.2 Coherent Detection

This method assumes that PU signal is fully known at the receiver. The prior knowledge about the features of the primary signals (such as modulation type, pulse shape, data rate or statistical properties) can be used for increasing the performance of the spectrum sensing detection. Provided the received signal is $y^i(j) = x^i(j) + w^i(j)$, the demodulation of the PU signal ($x^i(j)$) can be coherently done once timing and carrier synchronization sequences are attained. The optimal coherent detection consists of a filter matched to $x^i(j)$,

for the case having no interference and having stationary Gaussian noise only. However the filter consists of $hx^i(j)$, for the case when fading is present in the model, where h represents the fading coefficient known at the receiver. The maximum -likelihood of the received signal obtained after sampling is given as

$$Y = \frac{1}{N}\sum_{i=0}^{N-1} y^i(j)x^i(j)^* \tag{2.6}$$

where $x^i(j)^*$ represents the conjugate of $x^i(j)$ and N represents the total number of samples, where $n = 0, 1, 2, .., N-1$. The SNR of the matched filter output, is $\gamma = \frac{P}{\sigma_w^2}$, where σ_w^2 represents the average noise power and P represents the average signal power under the assumption that primary signal is stationary and the sensing time N is long enough given as

$$P = \frac{1}{N}\sum_{i=0}^{N-1} |x^i(j)|^2 \tag{2.7}$$

This technique has better detection performance than energy detection and consumes less time [51], [52]. However, it needs complete knowledge of the PU signal, which is impractical in reality [53]. Also, the synchronisation between the PU signal and the detector is a prerequisite for good performance in this technique because synchronisation errors can severely degrade the performance. In order to detect different kinds of the PU signal, the correlation of the received signal needs to be computed with different signal patterns, which eventually will increase the implementation complexity and power consumption.

2.2.3 Feature detection

In practical wireless communication systems, the transmitted signals have some distinct features that can assist receiver to detect. These features could

be evaluated by exploiting the second order statistics of the received signals, where in special cases the signals show cyclostationary properties. Two feature detection schemes are explained as follows:

Cyclostationary-based detection

In reality, coherent detection is difficult to implement because it is cumbersome to attain the complete knowledge of the PU signal. However, some features of the PU signal can be used for defining the test static. For example, modulated signal can be distinguished from the noise because most of the modulated signals are not stationary, however they are cyclostationary. The cyclostatioarity refers to the property, where the mean and the autocorrelation function are the periodic functions of time [54]. In this scenario, the cyclic autocorrelation function (CAF) of the received signal can be analysed. Generally, the CAF is periodic for the data signals but aperiodic for the noise signals [55], [56]. Consider a linearly modulated signal $x^i(j) = \sum_n a_n g(t - nT_0)$, where T_0 represents the symbol period $g(t)$ is the modulation pulse and a_n is assumed to be independent and identically distributed wide sense stationary data sequence with zero mean and autocorrelation, $r_{n-m} = E[a_n a_m^*]$. The autocorrelation function for $x(t)$ is given as

$$A_x(t,\tau) = \sum_n \sum_m r_{n-m} g(t + \tau - nT_0) g^*(t - nT_0) \tag{2.8}$$

As $A_x(t,\tau)$ in (2.8) is a periodic function of t with period T_0, so it can be represented using Fourier series as

$$A_x(t,\tau) = \sum_n A_x^\alpha(\tau) e^{+j2\pi\alpha t}, \qquad \alpha = \frac{n}{T_0} \tag{2.9}$$

where Fourier co-efficient ($A_x^\alpha(\tau)$) is defined as

$$A_x^\alpha(\tau) = \frac{1}{T_0} \int_{\frac{-T_0}{2}}^{\frac{T_0}{2}} A_x(t,\tau) e^{-j2\pi\alpha t} dt, \qquad \alpha = \frac{n}{T_0} \tag{2.10}$$

The Fourier coefficient given in (2.10) represents the CAF of $x^i(j)$, where α represents the cyclic frequency and $\alpha = 0$ represents the autocorrelation of $x^i(j)$. The cyclic spectral density (CSD) of $x(t)$ is defined as the Fourier transform of $A_x^\alpha(\tau)$:

$$S_x^\alpha(f) = \int_{-\infty}^{\infty} A_x^\alpha(\tau) e^{-j2\pi f\tau} d\tau \tag{2.11}$$

The CAF can be used for detecting the presence of PU using $A_x^\alpha(\tau)$, provided that SU knows the T_0 of the PU signal. In contrast to the energy detector, the CSD detector can differentiate the PU's signal from the noise and the interference signals. The computational complexity of this technique is higher than energy detector because managing the cyclostationary features of all possible PUs is a complex task.

Autocorrelation-based detection

Signal autocorrelation is computed in this method in order to differentiate the signal from the noise. Let the received signal be represented as $y^i(j) = x^i(j) + w^i(j)$, then the sample autocorrelation matrix of $y^i(j)$ is given as

$$\mathbf{A}_y = \frac{1}{N} \sum_{n=L-1}^{N-1} \mathbf{y}^i(j)[\mathbf{y}^i(j)]^* \tag{2.12}$$

When $N \to \infty$, provided that signal is uncorrelated with noise and the noise sequence is white Gaussian, A_y converges to autocorrelation of $\mathbf{y}_n$ given as

$$\mathbf{A}_y = \mathbf{A}_x + \sigma_w^2 \mathbf{I}_L \tag{2.13}$$

where σ_w^2 represents the average noise power, $\mathbf{A}_x$ is the autocorrelation matrix of the column vector $x^i(j)$ and $\mathbf{I}_L$ represents the identity matrix. When there is no PU signal then $\mathbf{A}_y = \sigma_w^2 \mathbf{I}_L$. Different test statistics can be calculated from the autocorrelation matrix. This technique has several algorithm options depending on the test statistics choices that include: eigen value-ratio test, maximum-eigenvalue test and minimum-eigenvalue ratio test. Each of them is briefly discussed as follows:

- **Eigen value-ratio test** Assume ρ_{max} and ρ_{min} as the maximum and minimum eigenvalues value of $\mathbf{A}_x$, and μ_{max} and μ_{min} represent the maximum and minimum eigenvalues of $\mathbf{A}_y$, respectively, given as

$$\mu_{max} = \rho_{max} + \sigma_w^2 \tag{2.14}$$

$$\mu_{min} = \rho_{min} + \sigma_w^2 \tag{2.15}$$

  where $\frac{\mu_{max}}{\mu_{min}} = 1$, represents the condition when there is no PU signal and $\frac{\mu_{max}}{\mu_{min}} > 1$, otherwise [57].

- **Maximum-eigenvalue test** Considering (2.14), if there is no PU signal then $\rho_{max} = 0$, otherwise PU is present. In other words, $\mu_{max} > \sigma_w^2$, when PU is present while $\mu_{max} = \sigma_w^2$, otherwise [58].

- **Minimum-eigenvalue ratio test** This test statistic compares the energy of the signal (y_n) with the minimum eigen value of A_y. If their ratio exceeds a given threshold, then the primary signal is assumed to be present [59].

This method is robust to noise uncertainty and needs no synchronization [60]. The advantage of autocorrelation based detectors is that they do not require a priori knowledge of the PU's signal, channel and noise power. However the

maximum eigen value detector can be used when noise variance is known [61]. The performance of the autocorrelation based detector is better than energy detector when signal is correlated; however it matches energy detection when the primary user is independent and identically distributed signal. Due to autocorrelation computation, the complexity of this algorithm is higher than energy detector.

The comparison of the detection schemes is presented in Table 2.1. Following the discussion, it can be concluded that the spectrum sensing technique can be chosen according to the given scenario. Due to low computational complexity, energy detection is chosen for the spectrum sensing in this thesis.

The complexity of the spectrum sensing receiver increases when wideband sensing is performed. For that case, the radio frequency components like high resolution analogue to digital converters and high speed digital signal processors with large operating bandwidth are required. Sensing frequency that represent the number of times a cognitive radio should perform spectrum sensing should be carefully adjusted. The optimum value depends on the capabilities of cognitive radio itself and the temporal characteristics of PUs in the environment. The three spectrum sensing detectors differ each other in terms of complexity and accuracy as shown in Figure 2.1 where it is clearly observed that coherent detector is the most accurate detector with maximum complexity. This is because coherent detector requires complete knowledge of PU signal before deciding spectrum status [62].

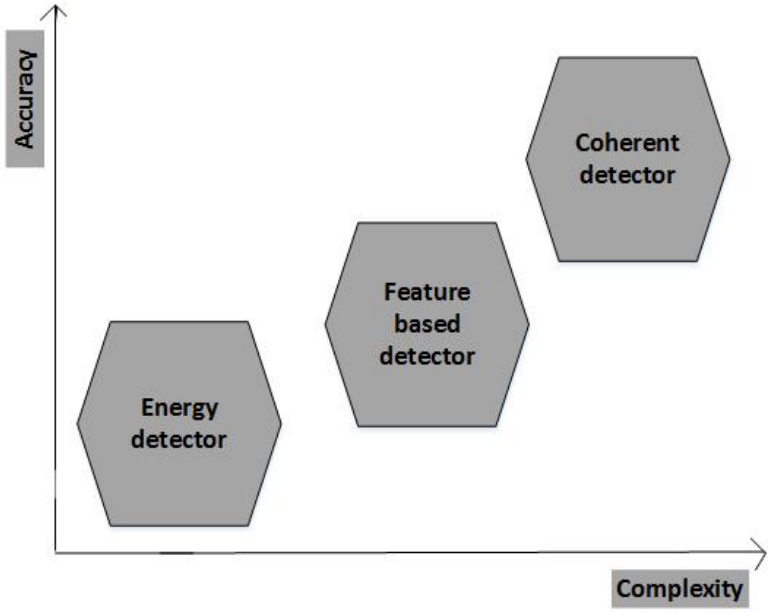


Figure 2.1: Comparison of spectrum sensing detectors in terms of complexity and accuracy [62].

Comparison between Spectrum Sensing Techniques			
Technique	**Required Information** σ^2 | x	**Distinguish PU from** Noise | other signals	**Complexity**
Energy Detector	Yes | No	SNR dependant | NO	Relatively simple
Coherent Detector	No | Yes	Yes | Yes	Maximise SNR but require synchronization
Cyclostationary Detector	No | Yes	Yes | Yes	Complex processing and high sensing time
Autocorrelation based Detector	No | No	Yes | Yes	Some test statistics require knowledge of the PU signal.

Table 2.1: Comparison between Spectrum Sensing Techniques. Adapted from [62]

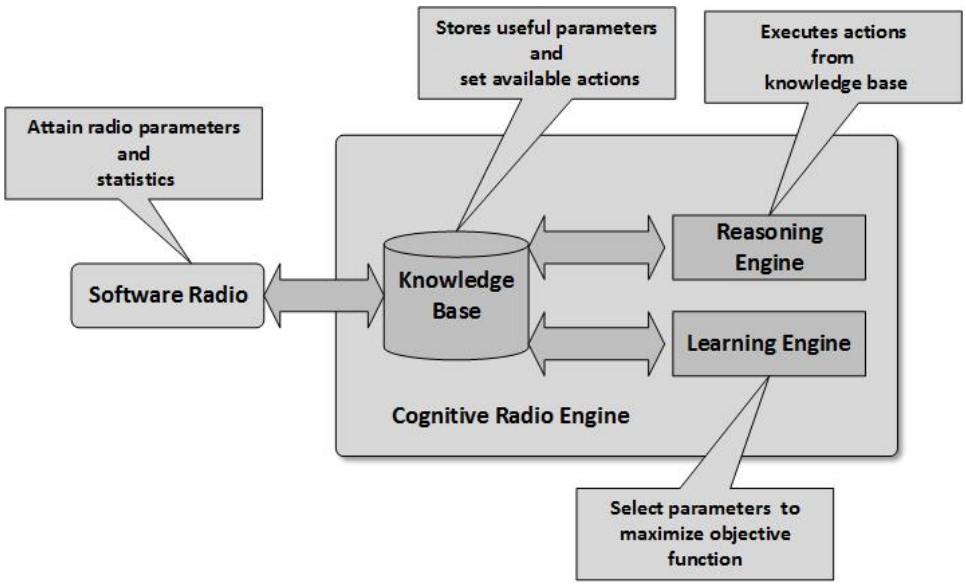


Figure 2.2: Cognitive radio engine illustrating the relationship between the software radio, the knowledge base, the learning and the reasoning engine. Adapted from [46].

2.3 Learning and Reasoning

An intelligent CR is composed of a software radio, knowledge base, learning and reasoning engine as shown Fig. 2.2. The knowledge base of the cognitive radio system stores the basic information like SNR, bit error rate, transmit power, coding rate and symbol constellation. The main objective of the learning engine is to evaluate the input state that will optimize the objective function, where the objective function depicts the goal of the application that could be the maximization of the throughput or the minimisation of the interference. The learning is mandatory in those situations, where the effect of changing inputs on outputs is not known. The channel statistics can be estimated using a learning engine which can be used for making optimal decisions using reasoning engine [63]- [64]. A comprehensive overview of the learning and reasoning engine using machine learning and bio-inspired algorithms is explained as follows.

2.3.1 Bio-Inspired Intelligence

Traditionally, AI has been focussed to imitate the cognitive abilities of the human brain. However in last few decades, there has been emergence of the new AI methods inspired from the biological processes such as the immune system, colonies of ants, bees and swarm of birds, to mention just a few [65]. The bio-inspired approaches, with appealing features such as self-adaptation, self-organization and collaborative intelligence, have been extensively applied to complex problems nowadays. The two most dominant classes of bio-inspired algorithms are evolutionary algorithms and swarm based algorithms, explained as follows:

- Evolutionary algorithms: These methods take inspiration from the fundamental biological principle: survival of the fittest and deletion of the poor solutions from the population. Each iteration evaluates the objective function for all individuals in a population and compares them. The fittest individuals are selected for the next iteration. The popular example is genetic algorithm and its variants [66]

- Swarm Based Algorithms: These methods exploit the coordination principles especially related to the distributed communication between individuals in nature. Each individual in this domain try to adjust its position according to the location of the fittest individual in the population. Popular examples include bird flocking, ant colonies and fish schooling [66].

An overview of the applications using evolutionary based algorithms and swarm based algorithms in cognitive radio networks is presented as follows

2.3.2 Bio-Inspired Cognitive radio

Inspired from the biological systems, a bio-inspired radio access scheme is proposed in [67], where the behaviour of the flock of birds swarming in search for food is simulated. Another technique, artificial bee colony (ABC) optimization is utilized in [68], for evaluating the optimal spectrum allocation matrix. It was observed in [68] that proposed method has outcomes near to the optimal values, which are calculated using exhaustive search. Similarly in [69], ant colony optimization is utilized for spectrum detection and allocation in multi-radio environment. In [70], a co-operative spectrum sensing method using particle swarm optimization (PSO) is presented that results in higher probability of detection compared to the previous spectrum sensing techniques. Further in [31], genetic algorithm (GA), quantum genetic algorithm (QGA) and PSO are utilized for the spectrum allocation. It was observed in [31] that all bio-inspired algorithms perform better than the traditional technique used for spectrum allocation in the literature (color sensitive graph algorithm) and specifically PSO outperforms both GA and QGA. A comparison between ABC and GA is presented in [71] for the optimal spectrum allocation and it was observed that ABC performs better than GA.

Although both evolutionary based algorithms and swarm based algorithms have been extensively applied in the cognitive radio network but there are still some complex problems that could be solved and optimized by exploiting the bio inspired algorithms. In this thesis, three bio-inspired techniques: particle swarm optimization, fire fly algorithm and fish school search are utilized for optimizing the collaborative spectrum sensing and allocation in Chapter 3. Both particle swarm optimization and fire fly algorithm are swarm based algorithms. They both exploit and explore the problem space to get the best optimal configuration for attaining spectrum sensing results with

highest probability of detection and for achieving fair spectrum allocation. However fish school search is a hybrid algorithm, that inherits some characteristic from both swarm based and evolutionary based algorithms. Both fire fly algorithm and fish school search are not utlized for addressing CSS before. It is observed in Chapter 3 that proposed algorithms performs better than PSO (conventional swarm based scheme known in literature for achieving the best spectrum sensing results in [70] and the best spectrum allocation results in [31].)

2.3.3 Machine Learning

Machine learning is a pioneer field of the artificial intelligence, where algorithms are implemented to analyse the data (input examples) based on the previous experiences. There are three ways to learn from the input examples: supervised learning, unsupervised learning and reinforcement learning [72] as shown in Fig. 2.2 and explained as follows:

Supervised learning

In this method, a part of the data samples is used for training the objective function. During training phase, a target value corresponding to each input sample is directly mapped as an output. The main objective of the supervised learning is to predict the future input-output observations with minimum error by finding a deterministic function which maps any input to an output. The recognition of the handwritten letters or digits is a popular example of the supervised learning [72].

Supervised learning is of two types: classification and regression. In classification, the output space can have discrete values. For example, the presence/absence of the cancerous cells can be detected using classification.

On the other hand, the learning problem is categorized as regression if the output space is formed by the values of the continuous variables [73]. For example, the prediction of shares in the stock exchange market is a regression problem.

Unsupervised learning

In this method, the input samples are not associated with the target values. The dataset for unsupervised learning could be a set of clusters or a probability density stating that how likely it is to observe a certain object in the future. For example, image and text segmentation can be solved using unsupervised learning. These learning techniques can also be used for the dimensionality reduction (transforming high dimensional data into low dimension space ensuring that the original information of the data is preserved) [74].

Reinforcement learning

In reinforcement learning, there is not an optimal output for a given input however the algorithm identifies an action that maximises the reward. In this scenario, the algorithm learns itself what needs to be done. Learning to play chess is an example reinforcement learning [76].

An overview of the machine learning based cognitive radio networks is presented as follows.

2.3.4 Machine Learning Cognitive Radio

Conventional spectrum sensing techniques need ample time to attain satisfactory performance and are computationally complex as well. On the other hand, machine learning is an effective tool for attaining promising results for dynamic spectrum access in CRN because ML techniques are adaptive to the

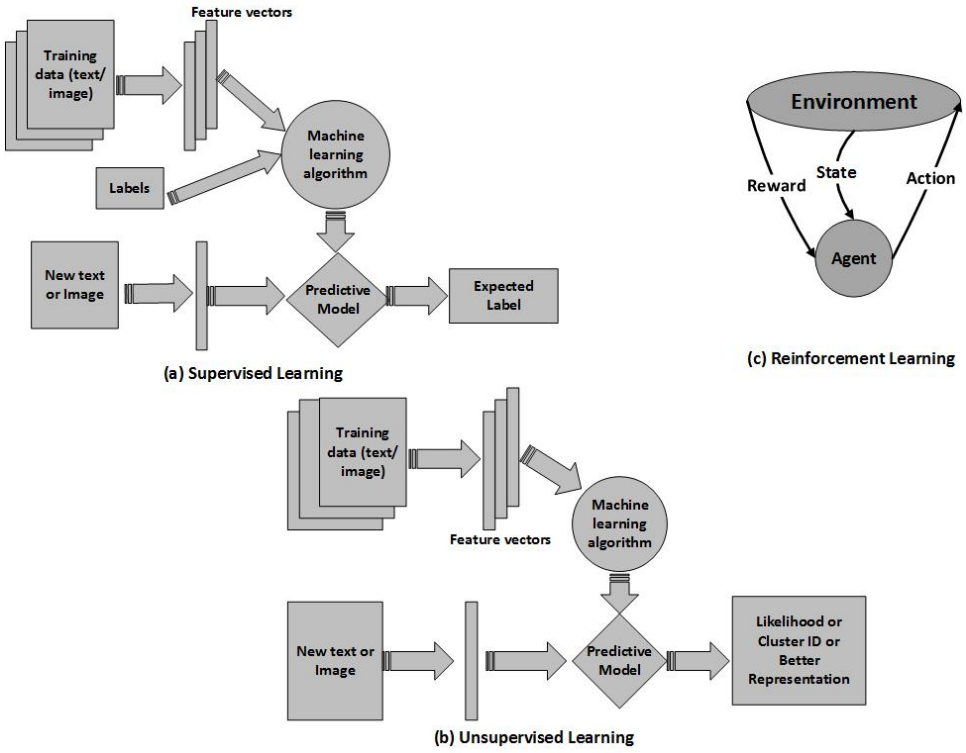


Figure 2.3: Three methods of learning from data in machine learning (a) supervised learning (b) unsupervised learning and (c) reinforcement learning. Adapted from [75].

changing radio environment. The same ML technique can be used for analysing the data of different days and locations [77].

Recently, ML has been extensively used in CRN for analysing the historic data to predict future outcomes. In [77], k-means clustering, Gaussian mixture model (GMM), support vector machine (SVM) and weighted k-nearest-neighbour (KNN) are utilized for the pattern classification in the collaborative spectrum sensing framework. The performance of each classification technique is quantified in terms of the average training time, the sample classification delay and the ROC curve. In [78], a ML based collaborative multi-band spectrum sensing policy for CRs is proposed that provides high throughput for the SUs, reduces miss detections and improves the energy efficiency in the system. Similarly in [79], the ML algorithms are utilized for the classification of the signal from different MAC protocols. In [80], an extensive discussion is presented on the application of kernel based learning (KBL) methods to the statistical signal processing in CRNs.

Though machine learning has been applied for improving the performance of cognitive radio networks but there exists some problems that can be addressed using ML. In this thesis, ML is used for analysing spectrum occupancy using the real time data acquired at the University for Warwick. The spectrum occupancy is not analysed using ML algorithms before. The machine learning framework for spectrum occupancy classification is presented in Chapter 4, where ML algorithms are also compared with a statistical approach in the literature used for analysing occupancy. Also, a novel energy predictive modelling framework using machine learning algorithms is presented in Chapter 5, which has not been proposed before. Two supervised algorithms: linear regression and decision trees are utilized for predicting those time slots and frequency bins where more energy can be harvested. The proposed approach

is also compared with the conventional prediction scheme in the literature.

2.4 Energy harvesting cognitive radio networks

Nowdays, energy harvesting has become a hot research topic for the next generation wireless communications systems. This is because energy cost paid by the operators for running their access networks is becoming a significant factor of their operational expenditures (OPEX) [81].

An extensive research has been done on energy saving hardware, energy friendly software applications and energy aware network architecture etc. As CR is a promising approach for increasing the spectrum efficiency using the concept of opportunistic spectrum access, there is a need to address the energy harvesting capacilities in CRN as well [82]. The deployment of energy harvesting capabilities in CR can provide a way to utilize the limited transmission power efficiently.

An energy harvesting CRN is proposed in [83], where SU harvests energy using renewable energy sources (solar, wind, and kinetic activities). Harvesting energy from the renewable sources not only makes wireless networks environmentally friendly but also self-sustaining. Apart from the renewable sources, radio frequency (RF) energy harvesting is also proposed for CRN in [84]- [86] briefly explained as follows.

2.4.1 RF Energy harvesting

The key component for converting radio waves into electrical energy in a RF energy harvester is rectenna, comprised of an antenna and a RF to DC converter. In addition to transmitting and receiving data, the wireless interface of the cognitive radio device can be reused for energy harvesting which will save cost and reduce implementation complexity [87]. A general overview of

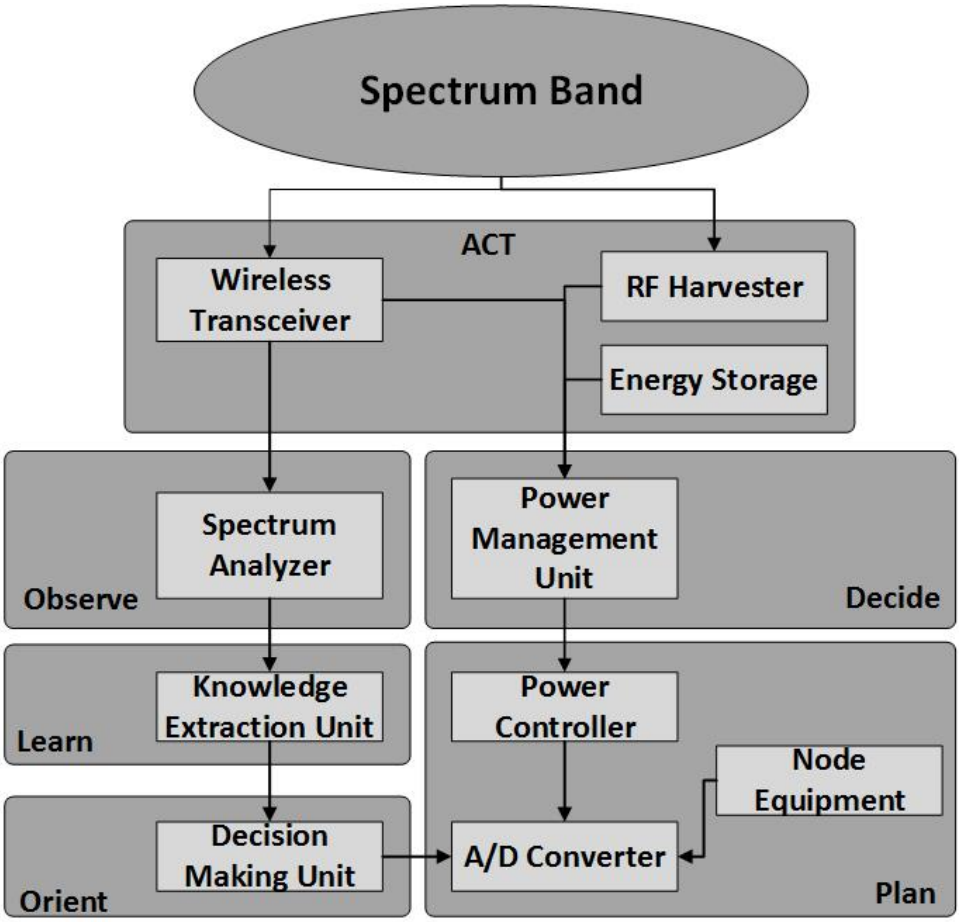


Figure 2.4: The relationship of cognition cycle with the components of RF powered cognitve radio network device. Adapted from [87]

the cognitive radio with RF energy harvesting features is presented in Fig 2.4. This includes wireless transceiver, spectrum analyser, knowledge extraction unit, decision making unit, analogue to digital converter, power controller, power management unit and a RF energy harvester. It is clearly depicted in Fig 2.4 that cognition cycle (as mentioned in Section 1.2) is achieved using observe, learn, orient, decide and plan modules. The cognition cycle along with a RF harvester and energy storage unit composes RF energy harvesting CR.

The output power of a RF harvester is typically in the micro watt to milli watt range which depends on number of factors such as distance and output power of the source. Table 2.2 summarizes the experimental measurements of the RF energy harvested from various RF energy sources [87]. Though the amount of energy that can be harvested using RF is very small, it can be used to provide the energy to the low power devices such as temperature sensors, hearing aids, and wristwatches. The advantage of RF energy harvesting is that it does not depend on nature, and hence it provides relatively predictable energy supply.

In conventional RF energy harvesting CRN approaches, SUs harvest RF energy from the transmission of nearby PUs opportunistically, store the harvested energy in batteries and uses it for transmitting the data, when primary user is passive. Therefore, SUs must not only identify spectrum spaces for the data transmission but it should also search the occupied spectrum to harvest the maximum energy [88].

Source	source power	Frequency	Distance	Harvested Energy
Isotropic RF transmitter	4W	902-928 MHz	15m	5.5 μW
Isotropic RF transmitter	1.78W	868 MHz	25m	2.3 μW
Isotropic RF transmitter	1.78W	868 MHz	27m	2 μW
TX91501 Powercaster Transmitter	3W	915 MHz	5m	189 μW
KING-TV tower	960kW	674-680 MHz	4.1km	60 μW

Table 2.2: Experimental data of RF energy harvesting. Adapted from [87].

In this thesis, an energy harvesting approach is proposed that considers

the idea of using wireless power transfer (WPT). There has been an increasing research interest in wireless power transfer methods as it can increase the lifetime of the power limited network. A brief overview about WPT is presented as follows.

2.4.2 Wireless power transfer methods

The concept of WPT was coined by Nikola Tesla in 1890, where it was suggested to transmit energy from the power source to the destination using a wireless medium. By using WPT, the cost for planning and installing power cables can be saved but one of the main challenges to implement WPT is its low energy transfer efficiency. Recently the research in WPT is motivated due to the widespread use of low-power devices that can be charged wirelessly [89].

The WPT using RF signals has recently attracted considerable attention due to two reasons: (a) it is a deterministic power transfer method (b) the information and power can be transferred simultaneously as RF signal. The WPT methods are categorized into two main types according to the transfer model and protocol (a) simultaneous wireless information and power transmit (SWIPT), where data and power signals are sent at the same time (b) wireless powered communication (WPC). In WPC, power signal is transferred first to the receiver and then harvested power is utlized by the receiver to transmit the information, thus it is also named energy harvesting communication as shown in Fig. 2.5. The information and power transfer is related more closely in WPC compared to SWIPT because the harvested power can affect the information rate [90], [91] where energy is sent first and then harvested power is used for sending the data signals.

In this thesis, a novel RF energy harvesting CRN is presented where PUs are wirelessly powered and have energy harvesting capabilities. In proposed

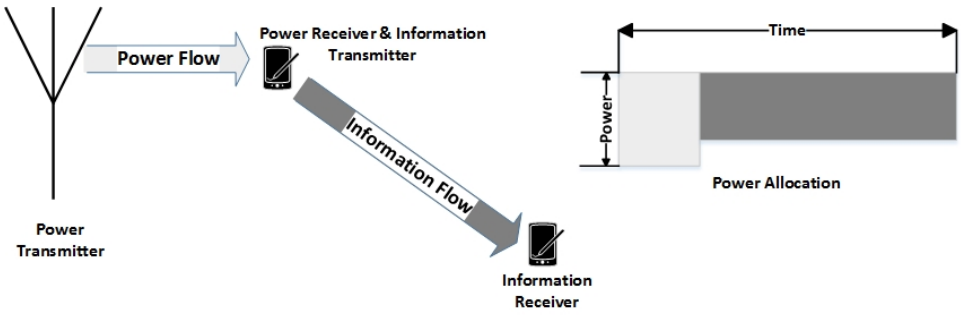


Figure 2.5: Wireless powered communication. Adapted from [91].

framework PUs are employing harvest-then-transmit protocol, where power is transmitted first and then the harvested power is used for sending information. The detailed framework is explained in Chapter 6.

Chapter 3

Bio-Inspired Cognitive Radios

3.1 Introduction

In collaborative spectrum sensing with energy detection, a data fusion centre combines energy measurements from all cooperating cognitive radios to make a final detection decision. The energy measurements from each cognitive radio are weighted for making final decision. As mentioned in Chapter 1, it was observed in [27] that cooperative spectrum sensing (CSS) outperforms the standalone energy detector. Reference [28] derived optimal and sub-optimal weights for a linear combination of measurements in the data fusion centre. More works on spectrum sensing can be found in [29] and the references therein.

After spectrum sensing, spectrum allocation is often performed to allocate the detected channels. A common method for spectrum allocation in literature is CSGC [30]. However, a bio-inspired technique, particle swarm optimization has outperformed CSGC in [31] by attaining higher value of spectrum allocation rewards. There is still a need to analyse those algorithms that can converge faster than PSO and can ensure higher spectrum allocation rewards. In this chapter, a cooperative spectrum sensing and allocation

framework is proposed using bio-inspired techniques, which not only provides an optimal weighting vector for the data fusion centre but also ensures conflict free spectrum allocation. The contributions of this chapter are summarized as follows:

- Three bio-inspired algorithms: firefly algorithm (FFA), fish school search (FSS) and particle swarm optimization (PSO) are used in this chapter, where FFA and FSS have not been used for collaborative spectrum sensing and allocation before. It will be shown in Section 3.4 that firefly and fish school search algorithm outperforms conventional schemes used earlier for collaborative spectrum sensing and allocation scenarios.

- In previous works, linear primary user signals were considered for CSS. However in reality, primary user signals may suffer from non-linear distortions, that is, if the power amplifier does not have enough gain, the input of the power amplifier at the primary user will be non-linearly distorted. In this case, the linear weights of the measurements as used in [28] may not be optimal any more. In order to address this, both linear and non-linear primary input signals with interference and fading losses are considered in this chapter.

- For CSS, the optimal weighting vector in the data fusion centre is computed using the popular algorithm 'weighted linear combining (WLC)' [28], and numerical results in Section 3.4 will show that the bio-inspired algorithms outperforms the WLC method.

- A cooperative spectrum allocation approach has also been proposed, which is dependent on the optimal 'weighting vector' evaluated by the data fusion centre using spectrum sensing results. The relationship between the optimal 'weighting vector' computed during cooperative spec-

trum sensing and the spectrum allocation module is not presented before.

The rest of the chapter is organized as follows. System model is presented in Section 3.2. The description of the spectrum sensing and spectrum allocation models using bio-inspired techniques is presented in Section 3.3. Numerical results and discussions are presented in Section 3.4. Finally, conclusions are given in Section 3.5.

3.2 System Model

Following the binary hypothesis test in Section 2.1, the received signal $r_d(m)$ is formulated as

$$H_0 : r_d(m) = z_d(m) \tag{3.1}$$

$$H_1 : r_d(m) = G_d s(m) + z_d(m). \tag{3.2}$$

In (3.1) and (3.2), $d = 1, 2, ...D$ and $m = 1, 2, ...M$, where D is the total number of radios and M is the total number of samples at each cognitive radio. Also, $s(m)$, G_d and $z_d(m)$ represent the primary signal, the channel gain and the zero-mean additive white Gaussian noise with variance $\sigma^2_{z_d}$, respectively. In order to model the real time scenario, the fading is introduced in the system where fading represents the fluctuation in the magnitude of the trasnmited signals due to variaiton in the propogation media. Both flat fading and frequency selective fading channels are considered. The frequency selectivity assumes the tapped delay line model given in [92]. Furthermore, RF power amplifiers (PA) are used in transmitters for converting the low power signal into higher power signal. The main objective of using PA are power efficiency, heat dissipation and input-output impedance matching. In this system, it is assumed that the primary signal $s(m)$ is a non-linear signal where the non-linearity is induced by

passing the signal through the PA at the primary user transmitter. Typically, the non-linear behaviour can be characterised due to amplitude compression (amplitude to amplitude (AM/AM)) or phase distortion (amplitude to phase (AM/PM)), and in some types of PAs both of these effect the system [93]. Two models are considered in this system that model PA non-linearity given as:

- *Memory less Polynomial Model*: As explained in [93], a strictly memory-less PA can be described in pass band as a non-linear function that maps the real valued input to the real valued output. This memory-less non linearity can be approximated using a power series given as

$$s(m) = \sum_{p=1}^{P} b_p [I(m)]^p \tag{3.3}$$

  where b_p are the real-valued coefficients, $I(m)$ is the pass-band PA input and $s(m)$ is the pass band PA output signal in (3.3) that will be sampled by cognitive radio.

- *High Power Amplifier (HPA) Model*: The nonlinear HPA model in the transmitter represents the nonlinear distortion imposed on the signal. A useful nonlinear HPA model is the Saleh model [94, 95]. Utilizing the Saleh model, the output of the HPA, $s(m)$, is given by

$$s(m) = \frac{I(m)A(r_s)e^{j\phi(r_s)}}{r_s}. \tag{3.4}$$

  In (3.4), $A(r_s)$ is an odd function of r_s, with a leading term representing AM/AM conversion and $\phi(r_s)$ is an even function of r_s, with a quadratic leading term representing AM/PM conversion. As AM/AM conversion

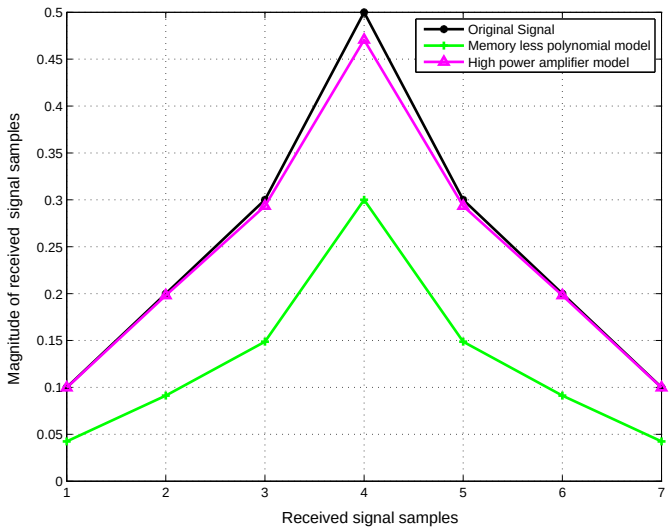


Figure 3.1: Comparison of high power amplifier model and memory less polynomial model.

is only considered in this thesis, $\phi(r_s)$ is not taken into account. In (3.4), $r_s = |I(m)|$ and $A(r_s) = \frac{e_A r_s}{1+f_A r_s^2}$, where e_A and f_A are constants [95].

It is shown in Fig. 3.1 that both high power amplifier model and memory less polynomial model affects the magnitude of original signal by different amounts. This is because the functions responsible for bringing non-linearities in both models are different. For spectrum sensing, the dth cognitive radio calculates its received signal energy using $E_d = \sum_{m=0}^{M-1} |r_d(m)|^2$ and sends it to the fusion centre. The fusion centre evaluates an output for decision as $y_{dd} = \sum_{d=0}^{D-1} w_d E_d = w^T E$, where $\mathbf{E} = [E_1, E_2, ...E_D]$ and $\mathbf{w} = [w_1, w_2, ...w_D]^T$ represents the weighting vector. Following [96], the probability of detection (P_d) is derived as

$$P_d = Q\left(\frac{Q^{-1}(P_f)\sqrt{\mathbf{w}^T X_1 \mathbf{w}} - M E_s G_d^T \mathbf{w}}{\sqrt{\mathbf{w}^T X_2 \mathbf{w}}}\right) \tag{3.5}$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$, P_f is the probability of false alarm, $X_1 =$

$2Mdiag^2(\sigma_{z_d})$, $X_2 = 2Mdiag^2(\sigma_{z_d}) + 4E_s diag(G_d)$, $E_s = \sum_{m=0}^{M-1}|s(m)|^2$, $\sigma_{z_d} = [\sigma^2_{z_1}, \sigma^2_{z_2}, ..\sigma^2_{z_D}]$, $G_d = [|G_1|^2, |G_2|^2...|G_D|^2]$ and $diag()$ represents the diagonal matrices.

It is evident in (3.5) that P_d resulting from CSS can be optimized by optimizing the weighting vector $\mathbf{w}$. Therefore P_d is a function of $\mathbf{w}$, or $f(\mathbf{w}) = P_d(\mathbf{w})$.

Three bio-inspired techniques are used in this chapter for optimizing $\mathbf{w}$, where $\mathbf{w}$ represents the position of bio-creature in the proposed system. The motivation of using bio-inspired algorithms comes from 'foraging', where every bio-creature tries to detect the best location of food with the help of their mates. In this work, the concept of the 'high food concentration' in bio-creatures is assumed as the 'high value of P_d' in CRN and the 'positions of bio-creatures' are assumed as the 'weighting vector'. The optimal weighting vector refers to the specific location where bio-creatures finds the maximum food concentration. The bio-creatures (fire flies, fishes and birds) are referred as particles in further explanation.

A slow changing spectrum access of primary users is assumed in this model, where particle m represents a sample sensed by cognitive radios. A particle m has a specific position in D dimensions, where dimensions are assumed equal to the number of CRs. The analogies between CRN and bio-inspired techniques are illustrated in Table 3.1. Each optimal weighting vector $\mathbf{w} = [w_1, w_2, ...w_D]^T$ has D dimensions, where the value of each dimension, say w_D, represents the weight assigned to dth CR as shown in Fig. 3.2. The weight for each CR determines its priority over the other, so it is named as 'priority weight'. It is evident that the change in particle's position will consequently affects the priority weights of all CRs in the system. During spectrum sensing, the fusion centre evaluates the optimal position of the particle that

yields highest P_d. Once optimal $\mathbf{w}$ is known by CR, the spectrum allocation is performed by the fusion centre as shown in Fig. 3.2. During spectrum allocation, it is proposed that the frequency channels will be allocated according to the value of 'priority weights' in optimal w. Let w_1 and w_2 be two values of priority weights belonging to CR1 and CR2 respectively in optimal $\mathbf{w} = [w_1, w_2, ...w_D]^T$, where $w_1 > w_2$, which shows that priority will be given to CR1 compared to CR2. Spectrum allocation is explained in detail in Section 3.3.2.

The probability of detection (P_d) in (3.5) will be used as an objective function for the evaluation of both bio-inspired spectrum sensing and allocation metrics. In the literature, the optimal value of $\mathbf{w}$ was determined in three ways: equal gain combining (EGC), weighted linear combining (WLC) and optimal combining (OC) [28]. In these techniques, it was observed that WLC and OC outperform the EGC method because their received energy measurements are 'weighted' as shown in [[28], Eq.5] and [[28], Eq.7]. To determine the optimal weighting vector in the WLC method, a heuristic technique was proposed in [96] that minimizes the probability of detection error as

$$\mathbf{w}_{wlc} = \frac{\gamma_d}{1 + 2\gamma_d} \tag{3.6}$$

where $\gamma_d = \frac{E_s|G_d|^2}{\sigma^2_{z_d}}$ is signal to noise ratio and $\mathbf{w}_{wlc}$ represents the weight for each dth radio respectively. Though WLC outperforms EGC, the proposed bio-inspired algorithms will perform better for the case of both linear and non-linear signals as will be explained in Section 3.4.

CRN	Bio-Inspired Techniques
Number of samples sensed by CR	Number of particles
Number of CRs	Number of dimensions of the position of mth particle
Fitness Function: P_d	Fitness function: food concentration
Optimal weighting vector	Best position of particle with maximum value of fitness function

Table 3.1: Analogies between CRN framework and proposed bio-inspired techniques.

3.3 Proposed Algorithms and Methodology

Three bio-inspired meta-heuristic algorithms are used for the collaborative spectrum sensing and allocation in this approach, where meta-heuristic algorithms are iterative search processes that efficiently perform the exploration and exploitation in the solution space for efficiently finding the near optimal solutions. In this context, three types of bio-inspired meta-heuristics: PSO, FSS and FFA were devised to find the optimal solutions of noisy non-linear continuous mathematical models. FFA is potentially more powerful in solving noisy non-linear optimization problems. The FFA not only includes the self-improving process with the current space, but it also includes the improvement in its own space from the previous stages. In [97], it was evaluated using benchmark functions that FFA outperforms the PSO in noisy situations. PSO is a powerful optimization tool but sometimes it cannot tackle dynamic optimization problems. It occurs because the entire swarm often increases the exploitation around a good region of the search space, reducing the overall diversity.

3.3.1 Spectrum Sensing

Let $\mathbf{w}_m^k = [w_{m1}^k, w_{m2}^k, ..w_{mD}^k]$, where $\mathbf{w}_m^k$ represents the mth particle during iteration k in D dimensions and $m = 1, 2, ..M$, where M represents the total

number of particles. The optimized $\mathbf{w}_m^k$ is evaluated using three bio-inspired techniques in this approach where first two steps are common in each algorithms given as

Step 1: Set k = 0, and generate initial positions of particles as $\mathbf{w}_m^k \epsilon a$, where a is a uniform random variable between [0,1].

Step 2: Evaluate the fitness of each particle ($P_d(\mathbf{w}_m^k)$) using objective function in (3.5).

Step 3: This step is specific for each technique and explained below.

Firefly Algorithm

(a) If ($P_d(\mathbf{w}_2^k) > P_d(\mathbf{w}_1^k)$), then an update in $\mathbf{w}_1^k$ occur as follows [99]:

$$\mathbf{w}_1^{k+1} = \mathbf{w}_1^k + \beta e^{-\gamma r_f^2}(\mathbf{w}_2^k - \mathbf{w}_1^k) + \alpha(rand - 0.5) \tag{3.7}$$

where r_f represents the Euclidean distance between $\mathbf{w}_1^k$ and $\mathbf{w}_2^k$, α represents the attractiveness between particles at initial stage, β is a positive constant, γ is the absorption co-efficient of the medium and $rand$ is the uniform random number generator. The third term $\alpha(rand - 0.5)$ is added for randomization with α being the randomization parameter [99].

(b) After comparison of the fitness of all M particles, the particle with the highest fitness is selected; which represents the optimal weighting vector of the kth iteration.

Fish School Search

(a) If ($P_d[\mathbf{w}_m^k] > P_d[\mathbf{w}_m^{k-1}]$, then $\mathbf{w}_m^k$ is updated as [100]

$$\mathbf{w}_m^k = \frac{P_d[\mathbf{w}_m^k] - P_d[\mathbf{w}_m^{k-1}]}{max|P_d[\mathbf{w}_m^k] - P_d[\mathbf{w}_m^{k-1}]|} \tag{3.8a}$$

(b) After all particles have moved individually, a weighted average of individual movements based on the instantaneous success of all particles is computed and added to current particle position given as [100]

$$\mathbf{w}_m^k = \mathbf{w}_m^{k-1} + \frac{\sum_{m=1}^{M} \Delta\mathbf{w}_m P_d[\mathbf{w}_m^k] - P_d[\mathbf{w}_m^{k-1}]}{\sum_{m=1}^{M} P_d[\mathbf{w}_m^k] - P_d[\mathbf{w}_m^{k-1}]} \tag{3.8b}$$

where $\Delta\mathbf{w}_m = \mathbf{w}_m^k - \mathbf{w}_m^{k-1}$ shows the displacement of particle due to individual movement in step (a). This step ensures that those particles who had successful individual movements influence the search direction more than other ones. (c) The successful particles produce their successor, where successful particle is determined by computing the maximum ratio of weight over distance for all particles. This process is called as breeding.

Particle Swarm Optimization

(a) Initialize the particle's velocity $\mathbf{v}_m^k \ \epsilon \ [-v_{max}, +v_{max}]$, where v_{max} and v_{min} represents the maximum and minimum value of velocity.
(b) The local best particle ($\mathbf{w}_l^k$) is evaluated in each iteration, who has the highest fitness compared to others. Similarly, the global best particle ($\mathbf{w}_g^k$) is selected in each iteration, who possess the maximum fitness value among all local best particles. The update in $\mathbf{w}_m^k$ and $\mathbf{v}_m^k$ given as [101]

$$\mathbf{v}_m^k = c_1\mathbf{v}_m^{k-1} + c_2\zeta(\mathbf{w}_l^{k-1} - \mathbf{w}_m^{k-1}) + c_3\eta(\mathbf{w}_g^{k-1} - \mathbf{w}_m^{k-1}) \tag{3.9}$$

$$\mathbf{w}_m^k = \mathbf{w}_m^{k-1} + \mathbf{v}_m^k \tag{3.10}$$

where ζ and η are uniform random variables between 0 and 1, c_1, c_2 and c_3 are positive constants which are selected by the practitioner to control the behaviour and efficiency of the PSO . The selection of the parameters is explained in detail in Section 3.4.1.

Step 4: If it reaches maximum generation, terminate the spectrum sensing algorithm and assign the optimal **w** to the spectrum allocation module; else, go to step 2.

In summary, the three technique discussed so far are partial search algorithms that may provide a good solution to the optimization problem. There is a distinct property in each of them that makes them different from each other. For example, PSO has exploration capability and emphasises on global search, where exploration depicts a phenomenon where large space is searched for finding promising solutions. This step brings diversification by avoiding local optimum traps. However FFA has more of exploitation capabilities along with exploration, where exploitation depicts a search method that is performed to improve the solutions that are already at hand. By having exploitation capability, FFA outperforms PSO in noisy situations even with faster speed and performs better when objective function has multiple peaks. The third algorithm, FSS, has breeding operator that let successful particles to go through the evolution process. The breeding operation ensures the survival of the fittest particles in the population. All three approaches have been utilized for several applications in literature, for example PSO for vehicle routing problem and anomaly detection. However FFA has been used for feature selection and antenna design while FSS for parameter estimation and function optimization respectively.

3.3.2 Spectrum Allocation

As discussed in Section 1.2.1, spectrum allocation refers to the selection of the best available spectrum channel according to the user communication needs. The main objective of spectrum allocation is to avoid conflict between users and treat them on fair basis. The conflict between users can occur when two

users try to utilize one channel at the same time. In this scenario, there is a need for a criteria that can give precedence to one user over another. The optimal value of weighting vector ($\mathbf{w} = [w_1, w_2, ..w_D]$), evaluated using FFA, FSS and PSO in Section 3.3.1 plays a vital role for the conflict free spectrum allocation that gives precedence to one user over another. For example, when two users try to access the same channel, then the value of $\mathbf{w}$ of one user is compared with another and the user having larger value of $\mathbf{w}$ will be allowed to transmit. This is because, P_d increases proportionally with an increase in the value of $\mathbf{w}$. This means maximum PU protection can be guaranteed by a user having larger value of $\mathbf{w}$ because larger value of $\mathbf{w}$ ensures higher value of P_d. The proposed spectrum allocation criteria not only resolves the conflict between two users by using optimal $\mathbf{w}$ but also ensure maximum protection to PU.

The spectrum allocation of cognitive radio can be explained with channel availability matrix, channel reward matrix and conflict free assignment matrix. It is assumed that D cognitive radios needs to communicate and U idle channels can be used, where $d = [1, 2, ...D]$ and $u = [1, 2, ..U]$. As CRs sense M channels, so it is assumed that U out of M channels are vacant. The concerned matrices are defined as:

- **Channel Availability matrix (L):**

$L = l_{d,u} \in [0, 1]$, that is D by U matrix, where $l_{d,u} = 1$, if channel u can be utilized by cognitive user d otherwise $l_{d,u} = 0$.

- **Conflict Free Channel Assignment matrix (A):**

The conflict free channel assignment matrix $A = a_{d,u} \in [0, 1]_{D*U}$, is a D by U matrix where $a_{d,u} = 1$, if channel u is assigned to user d or $a_{d,u} = 0$ otherwise.

For conflict free assignment, *assignment precedence matrix* (C) is proposed, that is D by U matrix. Each row in C is given as

$$C_d = w_d * L_{d,u} \tag{3.11}$$

where each row d in L is multiplied with corresponding component of optimal **w** as shown in Fig. 3.2. During sensing, for example; if $w_d > w_{d+1}$, then more precedence is given by the fusion centre to the dth radio. Similarly in spectrum allocation, if two radio users try to access the same channel, then the conflict is resolved by assigning the channel to the specific radio user, who has higher value of w_d.

• **Channel Reward Matrix (T):**

The channel reward matrix $T = t_{d,u}$ is a D by U matrix where $t_{d,u}$ represents the reward attained by user d for utilizing channel u. The reward attained by user d for utilizing channel u is r_d that is given by

$$r_d = \sum_{u=1}^{U} a_{d,u} * t_{d,u} \tag{3.12}$$

By following the above model, the spectrum allocation problem can be defined as the optimization problem that is dependent on the optimization of r_d. Following [31], the objective functions considered in this model for optimized spectrum allocation are (a) Maximum proportional fair reward (MPF): $MPF = (\prod_{d=1}^{D} r_d + 10^{-6})^{1/D}$, (b) Maximum Sum Reward (MSR) $MSR = \sum_{d=1}^{D} r_d$ and (c) Max-Min-reward (MMR): $MMR = min_{1<=d<=D} r_d$. The performance analysis is given in Section 3.4.

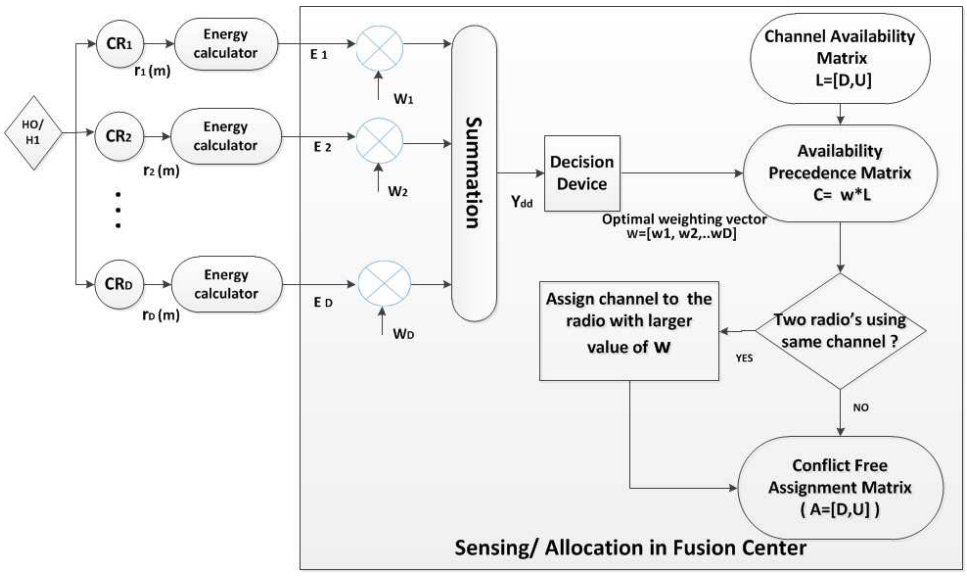


Figure 3.2: A cooperative spectrum sensing and allocation framework.

3.4 Numerical Results and Discussion

3.4.1 Parameter selection for Bio-Inspired Algorithms

The convergence speed and optimization accuracy of the bio-inspired algorithms is affected by the choice of parameters [102]. For FFA, β represents the attractiveness and for most cases $\beta = 1$ and $\alpha\epsilon[0, 1]$. The parameter γ characterizes the variation of the attractiveness, and its value is crucially important in determining the convergence speed and behaviour of the FFA. Thus, in most applications, it typically varies from 0.01 to 100 [99]. Following the constraints in [99], $\alpha = 1, \beta = 1$ and $\gamma = 1.3$ are chosen. The parameters c_2 and c_3 are the learning parameters or acceleration constants in PSO. A traditional way of improving the PSO method is by manually changing its behavioural parameters. The standard version of PSO with the learning parameters $c_2 \approx c_3 \approx 2$ are used [103]. Various studies have been reported in the literature [104], [105] regarding the choice of the inertia weight (c_1) and the velocity boundaries in PSO, which is believed to influence the degree of exploration versus exploita-

tion. After conducting extensive survey and experiments, it is recommended in [106] that $c_1 \epsilon [-2, 2]$ and $v_m^k \epsilon [-4, 4]$ for optimization experiments. Following [106], $c_1 = 1$ and $v_m^k = [-2, 2]$ are chosen. Following [94], $e_A = 1$ and $f_A = 0.25$ are chosen, for high power amplifier model (model1). The sum of coefficients of memory less polynomial model (model2) are assumed to be unity, so b_p coefficients are assumed as 0.4, 0.2 and 0.4, where $P = 3$. For the initial analysis, the number of particles are assumed as $M = 15$ and the number of radios as $D = 7$, where the received SNR of each radio is in range of $[0, -5]$ dB given as : 0, - 0.75, - 1.5, - 2.25, - 3, - 3.75 and - 4.5 dB. The effect of using different values of SNR and D are analysed in Section 3.4.4 and 3.4.5 respectively.

3.4.2 Comparison of PSO, FFA and FSS

Non-linear signals

In order to compare the convergence performances of the bio-inspired algorithms, the relationship between the 'number of iterations' and the maximum value of 'P_d' is demonstrated in Fig.3.3, by setting $P_f = 0.1$. A rectangular pulse is considered in Fig.3.3(a) as the input pulse and undergoes non-linear distortion using model1 and model2, while Fig.3.3(b) deals with cosine pulse. The results have shown that FFA outperformed PSO and FSS for both non-linear models. The maximum value of P_d attained by PSO, FFA and FSS for rectangular pulse using model1 is 0.9709, 0.9749, 0.9623 and using model2 is 0.8528, 0.8632, 0.8565 respectively. However the maximum value of P_d attained by PSO, FFA and FSS for cosine pulse using model 1 is 0.9424, 0.9760, 0.9615 and using model 2 is: 0.8491, 0.8602, 0.8555 respectively. It is observed that model 1 outperformed model2 for both input pulses. The difference in performance of model1 and model2 is due to the nature of non-linear functions

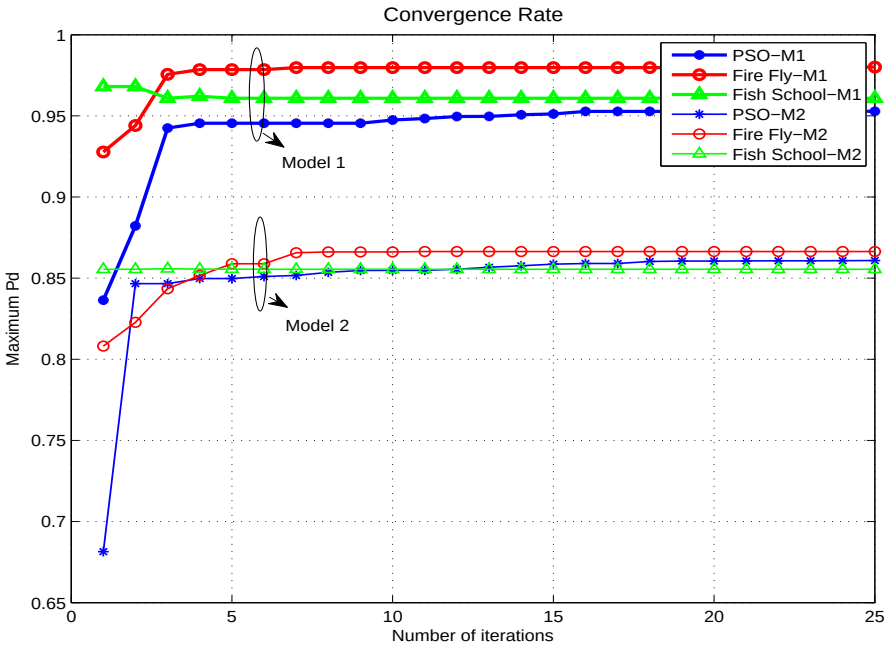
involved in the system.

Linear Signals

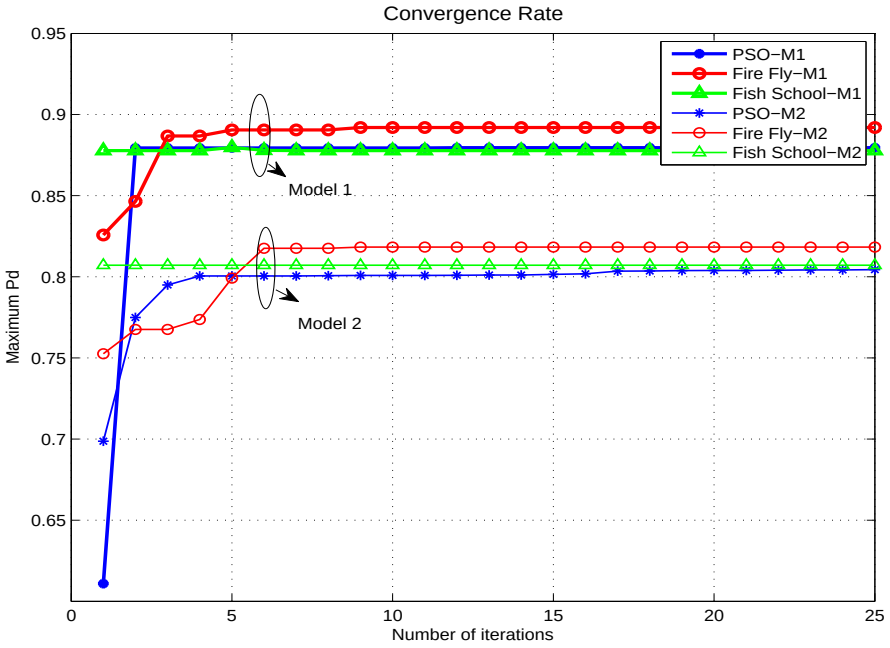
Both rectangular and cosine pulses without non-linearities are considered in Fig.3.4. It was observed that FFA algorithm outperforms PSO and FSS again. The maximum value of P_d attained by PSO, FFA and FSS for the rectangular pulse is: 0.9722, 0.9754 and 0.9630, while the maximum P_d value attained by PSO, FFA and FSS for cosine pulse is: 0.9259, 0.9432 and 0.9234 respectively. It was observed that the value of P_d for linear signals is higher than non-linear signals. The non-linear distortions induced by model1 and model2 are responsible for the P_d degradation in non-linear signals.

3.4.3 Comparison of Bio-Inspired Algorithms with WLC

The convergence of bio-inspired algorithms is examined for fixed values of P_f in Fig.3.3 and Fig.3.4. However, in real scenarios, P_f can change anytime so the effect of a changing value of the P_f on P_d is plotted as the ROC curve in Fig. 3.5 and Fig. 3.6 respectively. The probability of detection is affected by the probability of false alarm. This is illustrated using two state hypothesis in Section 2.2. In Fig. 3.5 and Fig. 3.6, the value of P_f is changed in each iteration by setting its $P_f = 0.01$ in the first iteration and incremented with a value of $P_f = 0.01$ in each iteration. The performance of WLC is compared with bio-inspired techniques and the results have shown that all three bio-inspired algorithms outperform the WLC method for both model1 and model2.



(a)



(b)

Figure 3.3: Using $P_f = 0.1$, the convergence rate of (a) rectangular pulse using model1 (M1) and model2 (M2) (b) cosine pulse using model1 (M1) and model2 (M2).

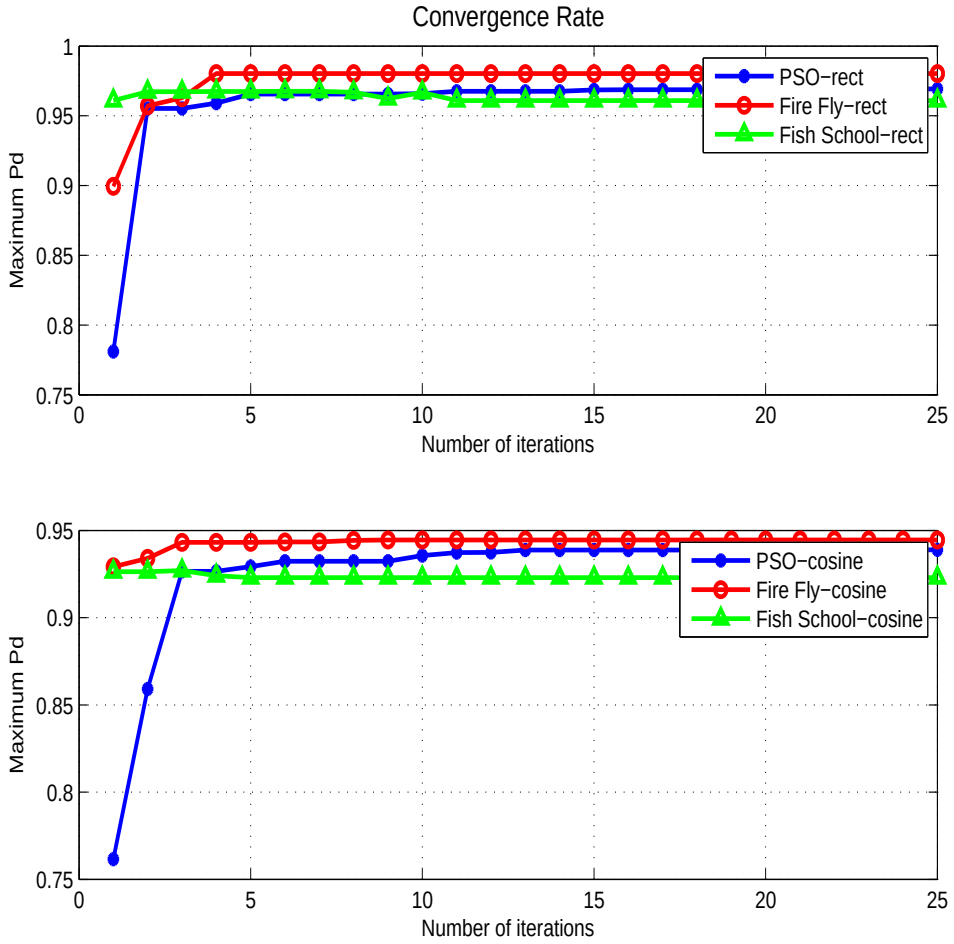


Figure 3.4: Using $P_f = 0.1$, the convergence rate of linear rectangular and cosine pulses.

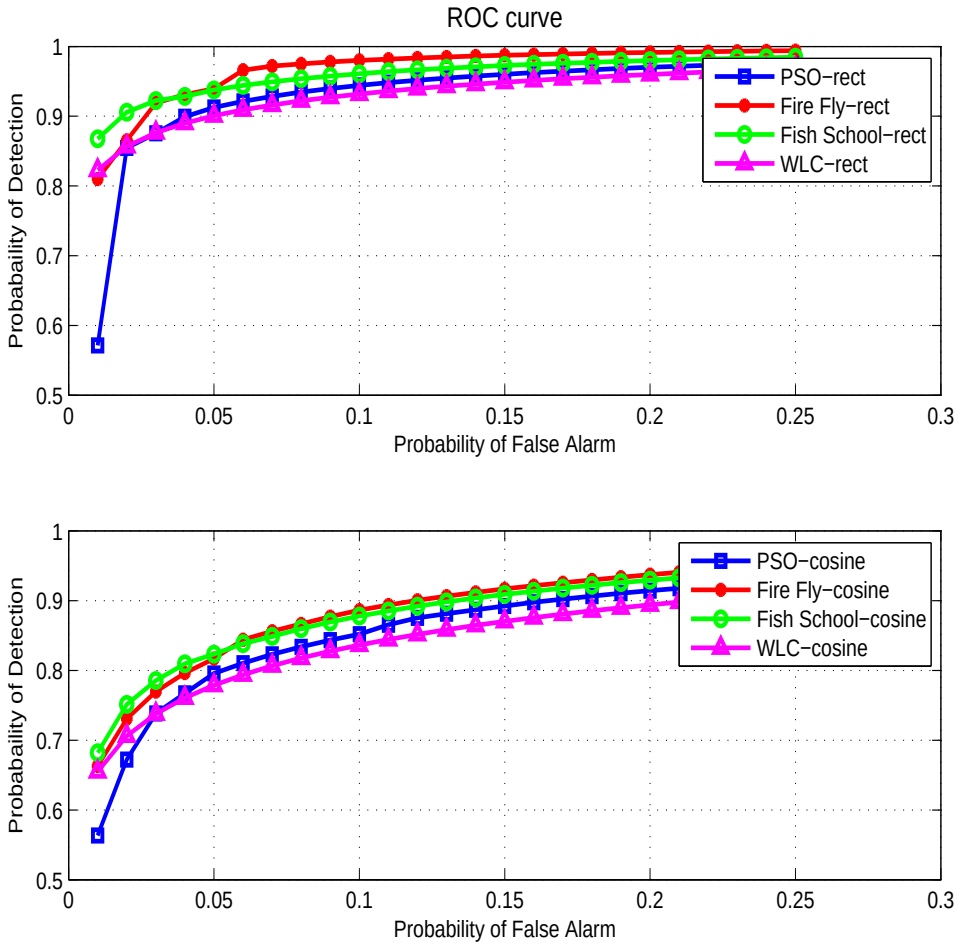


Figure 3.5: Comparison of bio-inspired algorithms with WLC using model1 by considering rectangular and cosine pulses for different values of P_f.

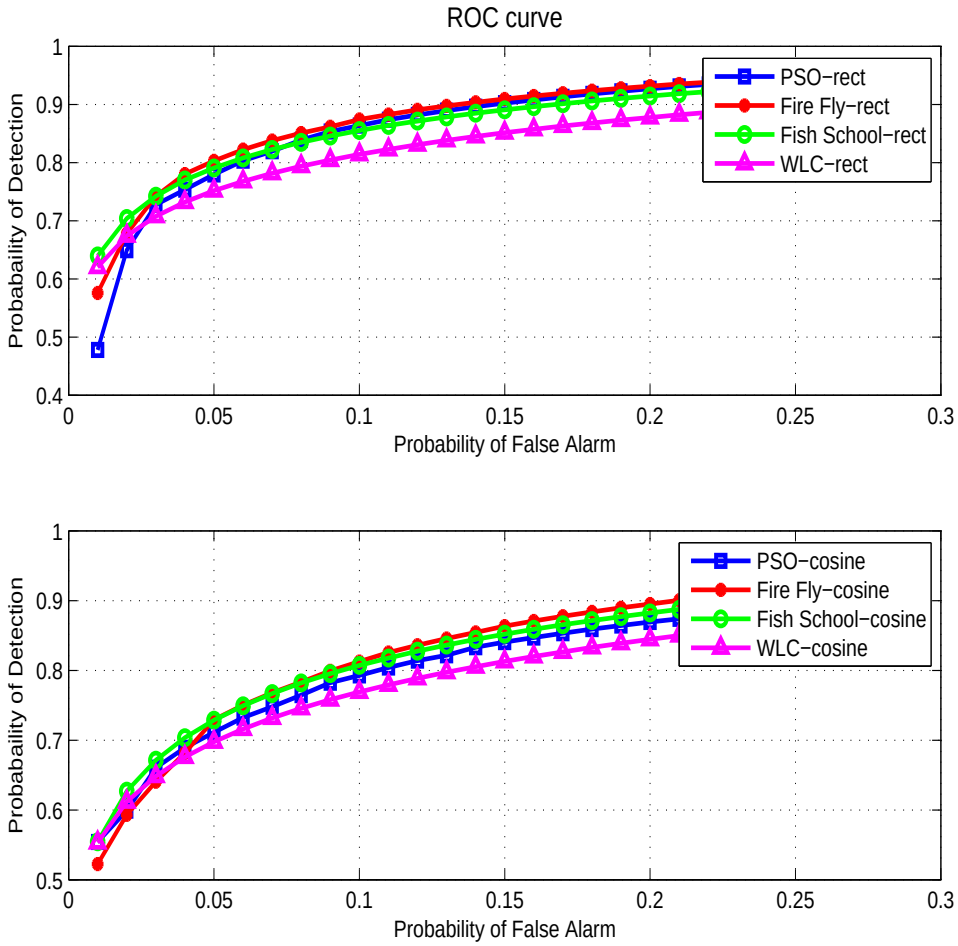


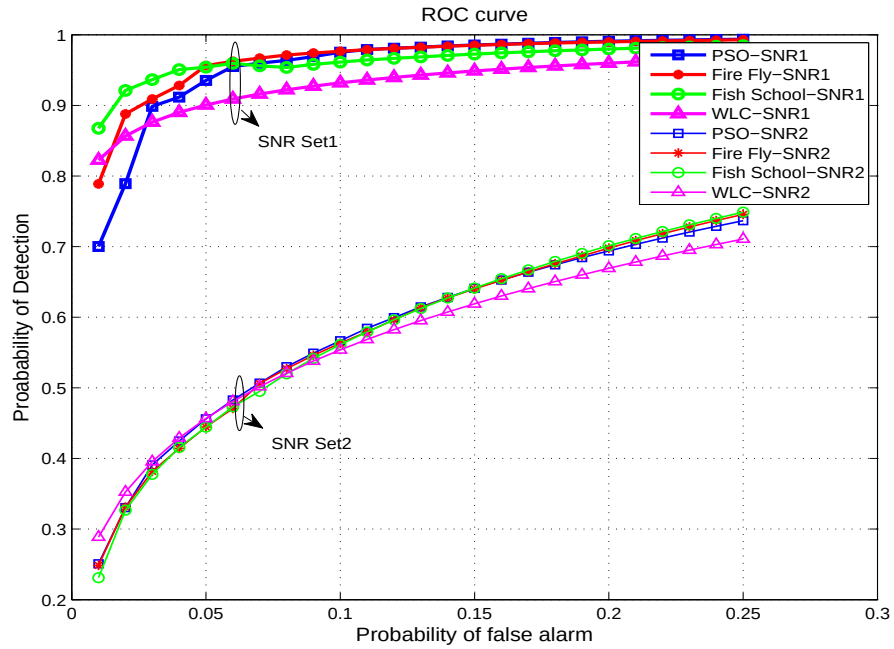
Figure 3.6: Comparison of bio-inspired algorithms with WLC using model2 by considering rectangular and cosine pulses for different values of P_f.

3.4.4 Effect of SNR

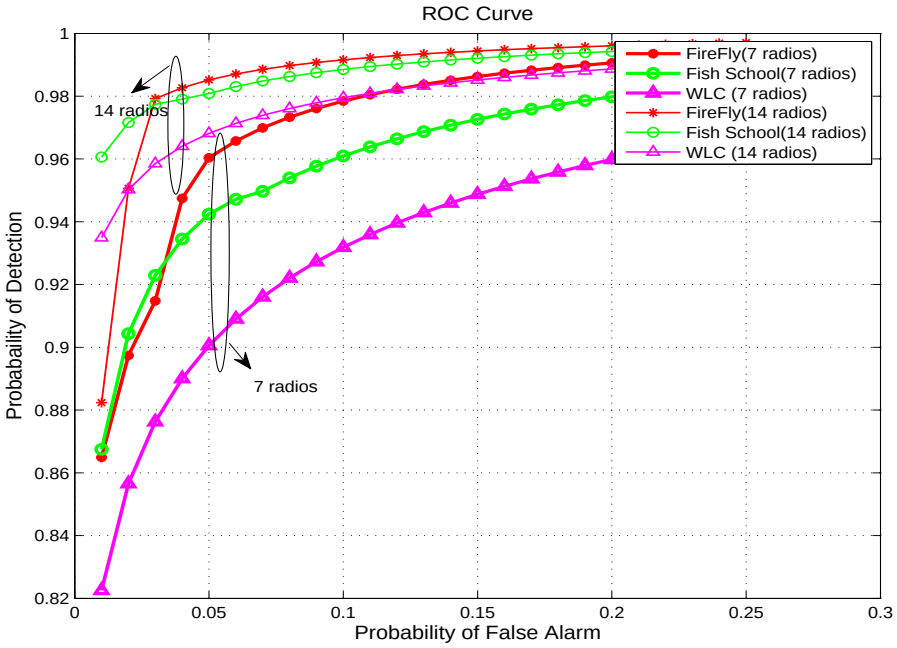
In the above subsections, different values of SNR for each radio are used while keeping the SNR for an individual radio constant for all simulation runs. However in real time, SNR changes according to the radio environment. There are possibilities to receive SNR as low as -10 dB that can directly affects the probability of detection. Therefore, the effect of different SNR sets is demonstrated in Fig. 3.7(a), where SNR set1 is assumed to be the same as mentioned in Section 3.4.1. SNR set2 is assumed as : - 5.25, - 6, - 6.25, - 7, - 7.75, - 8.5 and - 9.25 dB. It was observed that performance of SNR set2 is low compared to SNR set1. This is because SNR set2 ranges between [- 5, - 10] dB, while SNR set1 ranges between [0, - 5] dB. As P_d is directly proportional to SNR, so an increase in the value of P_d increases the value of SNR for each radio.

3.4.5 Effect of the number of radios

The effect of increasing the number of CRs from $D = 7$ to $D = 14$ is analysed in Fig.3.7(b). By increasing D, the dimensions of the particle's position actually increases, which eventually results in better performance. It was observed in Fig.3.7(b), that FFA and FSS outperforms WLC for both cases, $D = 7$ and $D = 14$. Though increase in number of radios, increases the mean P_d value for all algorithms; however it results an increase in the computational time and memory as well. Using MATLAB on a system with Core i7 processor and 8GB RAM, it was observed that the occupied memory and execution time for $D = 7$ is 964 MB and 15.9 sec respectively, however the occupied memory and execution time for $D = 14$ is 975 MB and 20.6 sec, respectively.



(a)



(b)

Figure 3.7: (a) Effect of different values of SNR on the value of P_d (b) Effect of changing the number of radios on the value of P_d.

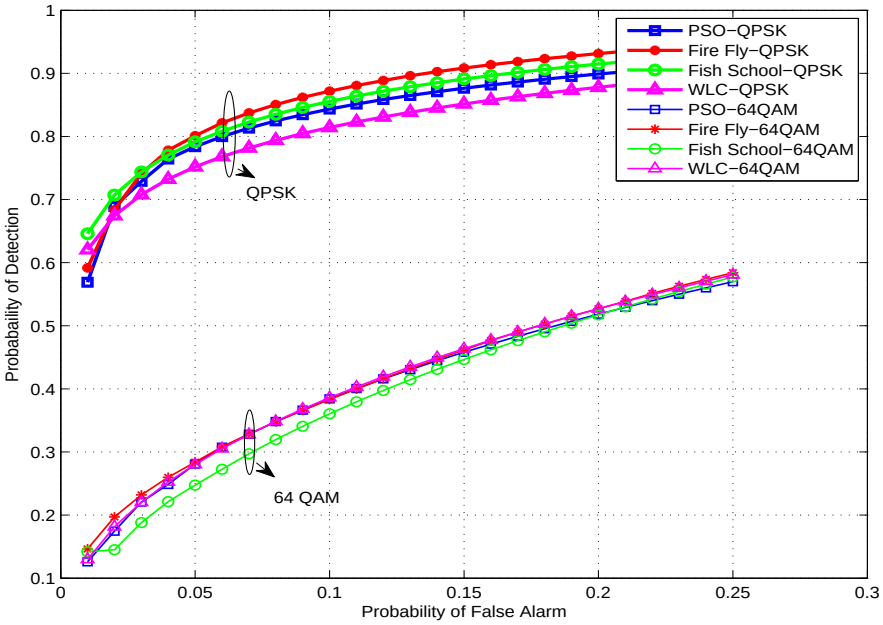
3.4.6 Effect of Modulation, Interference and Fading

The modulation is an important module of transmitter that takes place before sending data to the receiver. In the proposed approach, it is assumed that input data is modulated first and then amplified using PA that induces further non-linearities. The modulated non-linearly amplified output is sent to the channel, where the magnitude of the transmitted is degraded due to fading and interference. It is assumed that Gaussian noise is present always. The effect of different modulation schemes, channel fading and interference is discussed as follows.

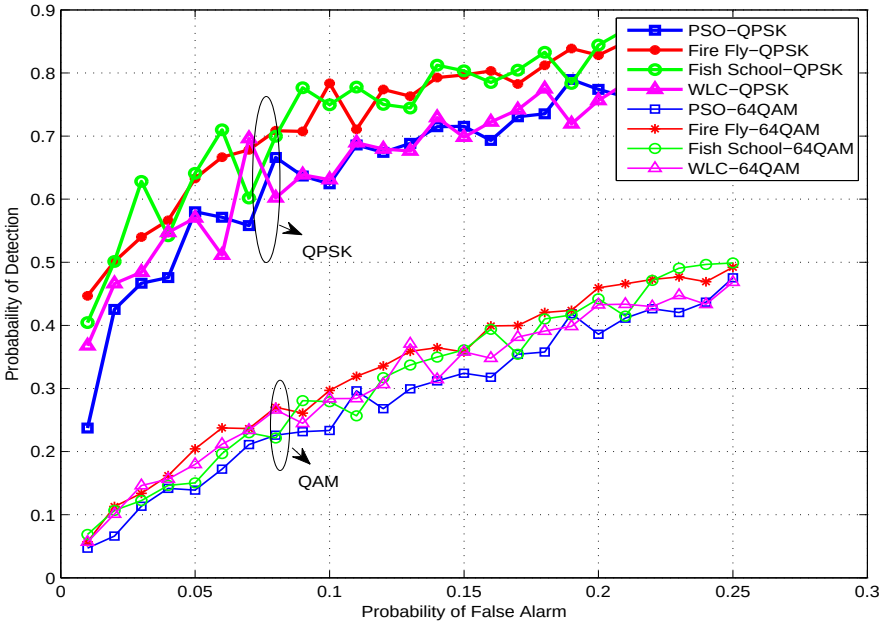
Two modulation schemes: 64 QAM and QPSK are considered for analysing the case, when $I(m)$ is fed into non-linear amplifier as modulated input. The result is shown in Fig. 3.8(a) using rectangular input pulse. It was observed in Fig. 3.8(a), that modulated rectangular pulse using QPSK attains higher P_d than 64 QAM. The FFA has outperformed other algorithms but there is 11% decrease in the value of its P_d compared to the non-linear model2 without modulation in Section 3.4.2.

Furthermore, external interference is introduced to the non-linear model modulated using QPSK and 64 QAM following the approach in [92]. The interference represents the noise faced by each CR caused due to PUs. It is observed in Fig. 3.8(b), that the value of P_d decreases more with the introduction of interference. The degradation in the value of P_d occurs because $r_d(m)$ faces both white Gaussian noise and interference in the channel. It is again observed that FFA has outperformed other schemes but interference degrades the performance of FFA 10% more compared to non-linear modulated model mentioned above.

Fig. 3.9 analyses the effect of frequency selective fading channels using the non-linear modulated model with interference, mentioned above. The



(a)



(b)

Figure 3.8: Effect on the value of P_d using (a) different modulation schemes (QPSK and 64 QAM) (b) modulation schemes (QPSK and 64 QAM) plus interference

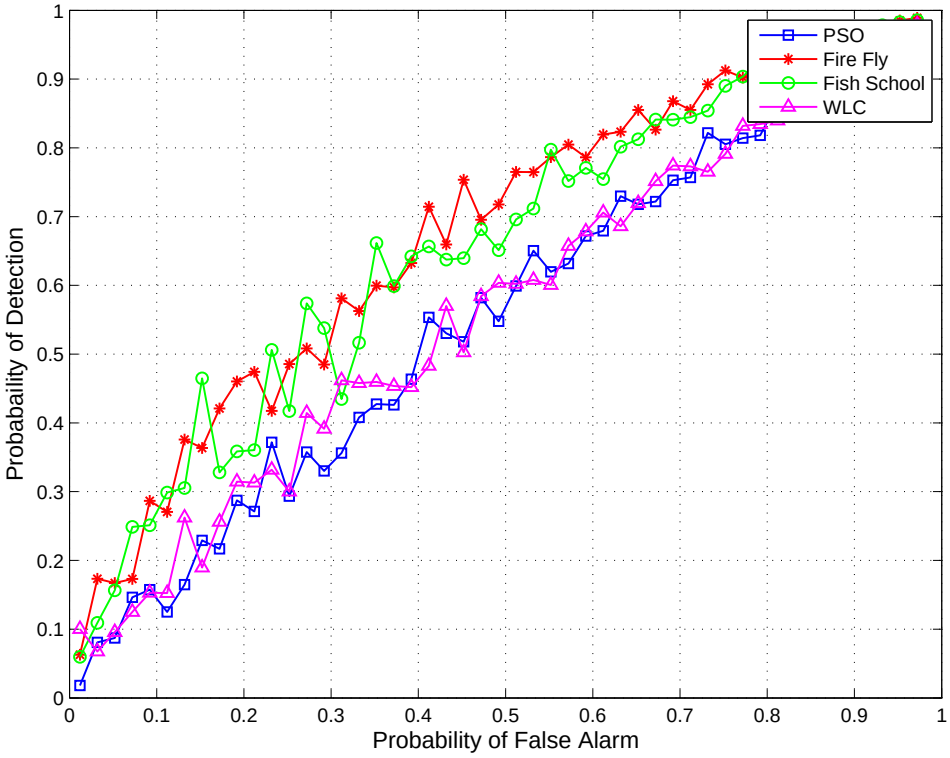


Figure 3.9: Average values of non-linear frequency selective faded model modulated using QPSK with $k = 250$

channel impulse response (G_d) is modelled as T time delayed taps with independent Rayleigh fading gains following [107], for 250 iterations. The QPSK is considered as the modulation scheme. The mean P_d values attained by PSO, FFA FSS, WLC is 0.5519, 0.6704, 0.6506, 0.5615 respectively. With the introduction of fading, the performance of FFA is decreased 7% more compared to the non-linear modulated model mentioned above. Also the execution time increases, when fading is considered. The time required for 250 iterations is 1 hour (3545.35 sec).

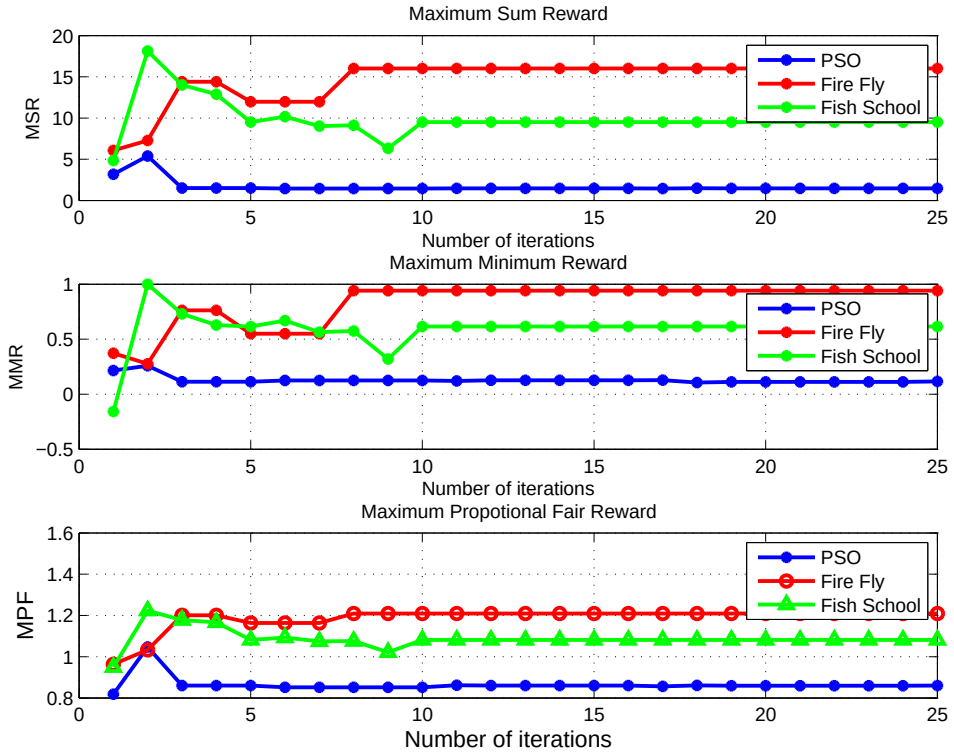


Figure 3.10: Spectrum allocation rewards using rectangular pulse and model2.

3.4.7 Spectrum allocation functions

As discussed in Section 3.3.2, the spectrum allocation is performed to optimize three objective functions: MSR, MMR and MPF, where MSR maximizes the utilization without considering fairness while MMR maximizes the reward for user with least allocated spectrum and MPF assigns resources to the user with fairness. The three objective functions are plotted in Fig.3.10 using model2, where rectangular pulse is used as input primary signal. The mean MSR attained by PSO, FFA and FSS is 1.6641, 20.2950 and 9.8550 respectively, while MMR is 0.1104, 0.8632 and 0.5628 respectively, and MPF is 0.8678, 1.1987 and 1.0831 respectively. It was observed that FFA outperformed FSS and PSO for all three objective functions.

The detailed comparison of linear model, non-linear model, modulated

non-linear model and modulated non-linear faded model is presented in Table 3.2 where it was observed that FFA outperforms PSO and fish school by attaining higher value of P_d for all models. It was also observed that FFA attain maximum MPF reward compared to other two algorithms. Furthermore, the computational time required by each model increases as one moves from left to to right in the table and its value is maximum for the non-linear faded model.

Performance Comparison													
	Algorithm	Linear Model			Model 2			A=Model2 +QPSK			A+Fading		
		MPF	P_d	T'	MPF	P_d	T'	MPF	P_d	T'	MPF	P_d	T'
K=25 D=7	PSO	0.87	0.91	0.62	0.79	0.83	0.63	0.77	0.71	2.09	0.71	0.48	30.9
	Fire Fly	1.19	0.97	0.71	1.11	0.86	0.72	1.06	0.79	2.91	0.91	0.62	35.1
	Fish School	1.07	0.95	0.64	1.01	0.85	0.69	0.99	0.75	2.11	0.91	0.52	32.5
K=45 D=7	PSO	0.88	0.94	0.91	0.81	0.81	1.15	0.76	0.71	3.75	0.74	0.46	54.5
	Fire fly	1.27	0.98	0.99	1.25	0.88	2.11	1.09	0.73	4.01	0.93	0.53	60.5
	Fish School	1.11	0.96	0.94	1.09	0.86	1.71	1.06	0.72	3.91	0.42	0.66	55.1
K=25 D=10	PSO	0.93	0.97	0.67	0.85	0.84	0.68	0.80	0.77	3.41	0.75	0.55	132.5
	Fire fly	1.20	0.98	0.87	1.17	0.87	0.95	1.09	0.80	8.62	0.97	0.69	160.2
	Fish School	1.15	0.97	0.69	1.07	0.86	0.64	1.01	0.70	3.56	0.94	0.60	130.5
K=45 D=10	PSO	0.97	0.98	0.97	0.87	0.87	0.92	0.80	0.60	4.35	0.77	0.54	159.7
	Fire fly	1.38	0.98	0.97	1.29	0.88	1.74	1.18	0.68	15.2	0.98	0.58	288.1
	Fish School	1.20	0.97	0.98	1.16	0.87	1.32	1.12	0.76	3.84	0.99	0.57	155.1

Table 3.2: Performance comparison of linear, non-linear, non-linear modulated and non-linear modulated faded models

3.5 Conclusion

A framework for collaborative spectrum sensing and allocation using used bio-inspired techniques in cognitive radios is presented in this chapter. Both linear and non-linear signals have been considered. It was observed that the non-linearities induced using high power amplifier model degrades the performance of spectrum sensing more compared to the memory less polynomial model.

It was observed that all bio-inspired techniques performed equally well in the presence of Gaussian noise and outperformed the conventional spectrum sensing weighting method: WLC. Bio-inspired techniques performed better because they are the iterative search processes which efficiently find near optimal solutions using exploration and exploitation principles. However, it was observed that bio-based solutions and WLC are affected by the change in SNR values. The increase in noise, interference and fading degrades the performance of all algorithms. PSO is affected more compared to FFA, because FFA is potentially more powerful in solving noisy non-linear optimization problems compared to PSO. Similarly FSS auto-regulates its exploitation and exploration capabilities compared to PSO, so it was concluded that FSS performs better in noisy conditions compared to PSO but its performance is worse than FFA. A precedence based spectrum allocation framework based on spectrum sensing weighting vector is also presented. It was noticed that bio-inspired techniques not only helps to attain higher value of probability of detection but also ensure conflict free spectrum allocation.

Chapter 4

Machine Learning Cognitive Radios

4.1 Introduction

Various spectrum measurement campaigns covering a wide range of frequencies have been performed all over the world in order to evaluate the spectrum utilization [108]. These spectrum measurement studies have found significant amount of unused frequency bands in the case of normal usage due to the static spectrum regulations. This has led researchers to understand the spectrum occupancy characteristics in depth for exploiting the free spectrum.

Many studies have been performed to understand the occupancy statistics. As mentioned in Chapter 1, the statistical and spectral occupation analysis of the measurements was presented in [32] in order to study the traffic density in all frequency bands. In [33], auto-regressive model was used to predict the radio resource availability using occupancy measurements in order to achieve uninterrupted transmission of secondary users. In [34], the occupancy statistics were utilized to select the best channels for control and data transmis-

sion purposes, so that less time is required for switching transmission from one channel to the other in the case when the PU appears. Further, in [109], [110], the bandwidth efficiency was maximized by controlling the transmission power of cognitive radio using spectrum occupancy measurements.

In [111], different time series models were used to categorize the specific occupancy patterns in the spectrum measurements. In [112], a novel time-varying statistical model for spectrum occupancy is proposed that uses real time wireless frequency measurements for predicting the arrival rate of PUs in each frequency bin by assuming the Poisson distribution for the arrival rates of PUs and the exponential distribution for the idle durations. All of the aforementioned works have evaluated the spectrum occupancy models by using conventional probabilistic or statistical tools. These tools are often limited due to the assumptions required to derive their theories. For example, one has to determine whether the value is a random variable or a random process in order to use the probabilistic and statistical tools. On the other hand, machine learning (ML) is a very powerful tool that has received increasing attention recently [113]. The machine learning algorithms are often heuristic, as they don't have any prerequisites or assumptions on data. As a result, in many cases, they provide higher accuracy than conventional probabilistic and statistical tools. There are very few works on the use of ML in spectrum occupancy. The ML works related to cognitive radio in [114]- [118] discussed cooperative spectrum sensing and spectrum occupancy variation. However, in this chapter, a comprehensive investigation on the use of ML for analysing spectrum occupancy is presented. The contributions of this chapter can be summarized as follows:

- The use of ML algorithms in spectrum occupancy study is proposed. Both supervised and unsupervised algorithms are used. In [114], [115],

ML was used for cooperative spectrum sensing. However in proposed work, ML is used for spectrum occupancy modelling that may be used in all CR operations, including spectrum management, spectrum decision and spectrum sensing.

- Three supervised ML algorithms: naive Bayesian classifier (NBC), decision trees (DT), support vector machine (SVM) and one unsupervised algorithm, hidden markov model (HMM) has been used to classify the occupancy status of time slots. The classified occupancy status is further utilized for evaluating the blocking probability.

- A new technique that combines SVM with fire fly algorithm (FFA) is proposed in this work that outperforms all supervised and unsupervised algorithms.

The rest of the chapter is organized as follows: Section 4.2 explains the system model that includes the details of the SU model, the PU model and the machine learning framework. Section 4.3 discusses the supervised and unsupervised algorithms. Section 4.4 presents the numerical results and discussion. Finally conclusions are presented in Section 4.5.

4.2 System Model

4.2.1 Measurement setup and data

The data was measured using RFeye evaluation system at University of Warwick as shown in Figure. 4.1. The system is portable carry case having omnidirectional antenna whose frequency ranges from 800 MHz to 6 GHz. This system has wideband reception with a universal power adapter. The receiver noise figure (NF) varies with the frequency, for example NF is 8 dB for 10 MHz

- 3 GHz while it is 11 dB for the frequency range from 3 GHz-6 GHz. More details about the equipment can be found in [120]. The RFeye evaluation system was placed in a lab ,that is located on fourth floor of School of Engineering at University of Warwick. Due to plug and play capability of RFeye, a USB was connected to the equipment where power measurements were stored. The data was measured indoor for approximately four months (6th Feb-18th June 2013) from 880 MHz to 2500 MHz.

The range 880 MHz to 2500 MHz has eight main radio frequency bands that are: 880-915 MHz, 925-960 MHz, 1900-1920 MHz, 1920-1980 MHz, 1710-1785 MHz, 1805-1880 MHz, 2110-2170 MHz and 2400-2500 MHz. The number of the frequency bins in each band varies. A frequency bin is a frequency channel having a specific bandwidth. For example, the band 925-960 MHz contains 192 frequency bins, each occupying a bandwidth of 0.18 MHz, while the band 1710-1785 MHz contains 448 frequency bins, each occupying a bandwidth of 0.167 MHz. The dataset for each band is arranged in a two dimensional matrix $(y^i(j))$, where $i = 1, 2, .., n$ and $j = 1, 2, .., k$. Each row i of matrix $y^i(j)$ represents a time slot that contains the data of k frequency bins; while each column j of matrix $y^i(j)$ represents a frequency bin that contains the power of n time slots. As the data for four months is measured which constitute 131 days ($N = 188917$ minutes), therefore the number of rows in each band are 188917 while the number of columns varies according to the number of the frequency bins in a particular band. The total number of frequency bins (k) for each band and their corresponding bandwidth is given in Table 4.1.

4.2.2 SU Model

In a network of licensed users, SU is allowed to access the licensed band without causing any harmful interference to the PU. Let $y^i(j)$ be the sample sensed

(a)

(b)

Figure 4.1: (a) RFeye evaluation system (b) RFeye battery and monitor.

Band (MHz)	k	**Bandwidth of each frequency bin (MHz)**
880 -915	192	0.18
925 -960	192	0.18
1710 -1785	448	0.16
1805 -1880	448	0.16
1900 -1920	128	0.15
1920 -1980	384	0.15
2110 -2170	384	0.15
2400-2500	640	0.15

Table 4.1: UK cellular frequency bands measured at University of Warwick from Feb - June 2013.

at the i^{th} time slot in the j^{th} frequency bin. Following binary hypothesis in Section 2.2, one has

$$y^i(j) = x^i(j) + w^i(j) \tag{4.1a}$$

$$or \qquad y^i(j) = w^i(j) \tag{4.1b}$$

where $x^i(j)$ represents the received PU signal and $w^i(j)$ represents the additive white Gaussian noise (AWGN) with zero mean and variance σ_w^2. Using energy detector, the received signal energy (Y) is computed as: $Y = \frac{1}{n}\sum_{i=0}^{n-1}|y^i(j)|^2$. The received signal energy is compared with a decision threshold (λ) to attain the spectrum status. The spectrum status ($S^i(j)$) is given as

$$S^i(j) = \begin{cases} 1, & Y > \lambda \\ 0, & Y < \lambda. \end{cases} \tag{4.2}$$

The selection of λ is very important because small values of λ will cause false alarms while large values will miss spectrum opportunities. The computation of λ is explained in [119]. In the proposed approach, the threshold is dynamic and its selection is discussed in Section 4.4.1. Using ($S^i(j)$), the occupancy for the ith time slot (OC^i) is defined as

$$OC^i = \frac{\sum_{j=1}^{k} S^i(j)}{k}. \tag{4.3}$$

In order to clearly illustrate the concept of OC^i, a example of a three-minute interval for the band 880 - 890 MHz having 9 frequency bins is shown in Fig.4.2. Each bin in Fig.4.2 occupies 1 MHz. For each frequency bin, $S^i(j)$ is decided. Once $S^i(j)$ is evaluated, the occupancy OC^i is calculated using (4.3). It is observed that more frequency bins are occupied for the first minute than for the second and third minutes, that imply SU has less chance to transmit

in the first minute. Following the discussion above, it is required to set the criteria for quantifying the transmission chance based on the occupancies.

4.2.3 PU Model

In proposed approach, the status of PU (P^i) for the i^{th} time slot can be decided using the following rules:

$$P^i = \begin{cases} (Condition\ 1) : 1, & OC^i > U_{oc} \\ (Condition\ 2) : 1, & L_{oc} <= OC^i <= U_{oc} \\ \qquad AND \quad con^i <= B & \\ (Condition\ 3) : 0, & L_{oc} <= OC^i <= U_{oc} \\ \qquad AND \quad con^i > B & \\ (Condition\ 4) : 0, & OC^i < L_{oc} \end{cases} \tag{4.4}$$

where U_{oc} and L_{oc} represent the maximum and minimum values of the occupancy for all n time slots, con^i represents the number of consecutive free frequency bins in the ith time slot and B represents the maximum value of con^i, when PU is considered present. Each condition is explained as follows:

1. Condition 1 and Condition 4: The values of U_{oc} and L_{oc} vary with the frequency band, the day and the threshold. For fixed frequency band and day, U_{oc} and L_{oc} are evaluated using different thresholds in Section 4.4.2. In order to guarantee the PU protection and ensure SU transmission, when the values of OC^i lie in the range between L_{oc} and U_{oc}, further criterion is applied.

2. Condition 2 and Condition 3: It is difficult to apply Condition 1 and Condition 4, when $L_{oc} <= OC^i <= U_{oc}$, so con^i is evaluated for each time slot. If $con^i > B$ for $L_{oc} <= OC^i <= U_{oc}$, there exists at least B consecutive free frequency bins in the ith time slot; thus SU can transmit and vice versa

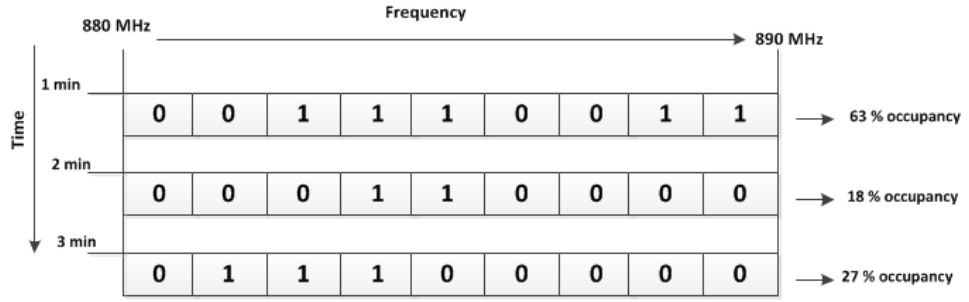


Figure 4.2: Occupancy for different time slots between 880 -890 MHz using $k = 9$ and $n = 3$.

when $con^i <= B$. The value of B is selected to provide PU protection. This will be explained in Section 4.4.2.

4.2.4 Machine learning framework for SU and PU Model

ML algorithms learn the target function that maps the input variables to output variables in the best manner given as $output = f(input)$. The function f needs to be learnt from data using different ML schemes. This is difficult problem to learn function from the data and that's the main motivation behind existence of various ML methods. When a function (f) is learnt, it means function (f) has been estimated using the data. The estimation can have errors as it represents the hypothetical mapping from output given input. In the proposed approach, a classifier is constructed to estimate P^i using $\mathbf{S}^i$, where $\mathbf{S}^i = [S^i(1), S^i(2), ..S^i(k)]$ represents the input (feature vector) and P^i represents the output labels (the corresponding response to the feature vector). The proposed system can be given as $P^i = f(\mathbf{S}^i)$, where function (f) is learnt with the help of supervised and unsupervised ML algorithms. There are two steps for constructing a classifier using supervised algorithms:

Training

During training phase, the input feature vector is mapped against the output labels so that function (f) can be learnt. In the proposed approach, $\mathbf{S}^i_{train} = [S^i(1)_{train}, S^i(2)_{train}, ..., S^i(k)_{train}]^T$ denote the training spectrum status (input feature vector) and P^i_{train} represent the output label for the ith time slot, respectively, where $i = 1, 2, ..n_1$ and n_1 represents the number of training time slots fed into the classifier. Once the classifier is successfully trained, which means it has learnt function (f) using the data then it is ready to receive the test vector for making estimation.

Testing

The testing phase is performed to quantize that how accurate the function (f) has been estimated. For that, the new feature vectors are fed into the machine learning system without any output labels. It should be noted that testing is performed for estimating the output labels of the new feature vectors. In the proposed system, $\mathbf{S}^i_{test} = [S^i(1)_{test}, S^i(2)_{test}..., S^i(k)_{test}]^T$ denote the testing spectrum status and P^i_{test} represent the actual testing PU status for the ith time slot, respectively, where $i = n_1{+}1, n_1{+}2, ..n_2$ and n_2 represents the length of the testing sequence. It is assumed that $n = n_1 + n_2$. The value P^i_{test} is not used during the testing but as a reference for computing the classification error.

Classification Accuracy (CA)

Let P^i_{eval} denote the PU status determined by the classifier for the ith time slot. The classifier categorizes the testing vector $\mathbf{S}^i_{test}$ as the 'occupied class' (i.e., $P^i_{eval} = 1$) or 'unoccupied class' (i.e., $P^i_{eval} = 0$). Therefore, the PU status is correctly determined, when P^i_{eval}= P^i_{test}, giving $CA^i = 1$. The mis-detection

occurs when $P^i_{eval} = 0$ and $P^i_{test} = 1$, while false alarm occurs when $P^i_{eval} = 1$ and $P^i_{test} = 0$, both giving $CA^i = 0$.

4.2.5 Blocking Probability

Let $\mathbf{P}^i_{eval}$ be a vector evaluated by each classifier and define out_{su} as the minimum value of consecutive free time slots required by SU for transmission. The blocking probability measures the chance when SU cannot find out_{su} consecutive free time slots for transmission given as [121]:

$$P_r(SU_{blocking}) = 1 - P_r(SU_{transmit}) \tag{4.5a}$$

where

$$P_r(SU_{transmit}) = \sum_{c=1}^{C} P_r(FB_c) \tag{4.5b}$$

where FB_c is the block of free consecutive time slots of length out_{su}, $c = \{1, 2, ..C\}$ and C represents the total number of free blocks in $\mathbf{P}^i_{eval}$. The probability for a free block starting at index, say r, in $\mathbf{P}^i_{eval}$ is evaluated using the following equation

$$P(FB_c) = \prod_{i=r}^{r+out_{su}} OC^i \tag{4.5c}$$

4.3 Proposed Algorithms

In the proposed approach, four machine learning algorithms are utilized to predict the future PU status using the occupancies, which are a function of time, frequency and threshold. Among four ML algorithms, three are supervised learning algorithms: NBC, DT and SVM, while one is an unsupervised algorithm, HMM. The motivation to use four different algorithms is to find the best machine learning algorithm for predicting the future status as each

one has different characteristics.

4.3.1 Naive Bayesian Classifier

It is also called 'independent feature model' because it does not take dependency of the features into account. The feature vector for the *ith* time slot in this model contains all the samples which are independent of each other, since every feature represents a specific frequency bin. For example, the status vector of the *ith* time slot is given as $\mathbf{S}^i = [S^i(1), S^i(2), S^i(3), .., S^i(k)]$, where $S^i(1)$ is independent of $S^i(2)$. However, the response variable in this approach i.e. PU status (P^i) is a dependent variable which is affected by each frequency bin. The probability of $\mathbf{S}^i$ belonging to the class P^i evaluated using the Bayes theorem is formally defined as [122]

$$P_r(P^i, \mathbf{S}^i) = P_r(P^i) * P_r(\mathbf{S}^i | P^i). \tag{4.6}$$

When $P^i = 0$, $\mathbf{S}^i$ will be classified as 'idle' class, while when $P^i = 1$, $\mathbf{S}^i$ will be classified as 'occupied' class. The goal is to find the class with the largest posterior probability in the classification phase. The classification rule is given as

$$classify(\hat{\mathbf{S}^i}) = argmax_{\mathbf{S}^i}\{P_r(P^i, (\hat{\mathbf{S}^i})\} \tag{4.7}$$

where $\hat{\mathbf{S}^i} = [\hat{S^i(1)}, \hat{S^i(2)}...\hat{S^i(k)}]$. NBC is sensitive to the choice of kernel and the kernel selection will be further explained in Section 4.4.3.

4.3.2 Decision Trees

Decision trees are one of the most popular approaches used for classification in machine learning, pattern recognition and data mining areas. The decision tree consist of parent node/root (the node which has no incoming edges),

decision nodes and leaves. Each decision node splits the data on the base of defined criteria into two subspaces. Each leaf node is assigned to one of the classes which represent the appropriate target value [123]. The decision trees used in this approach are classification trees whose leaf represents the class labels. Unlike NBC, it can handle feature interactions and dependencies. In DT, the decision is made on each internal node, which is used as a basis for dividing the data into further two subsets while leaf nodes represent the class labels (in the case of classification trees) or the real numbers (in the case of regression trees). Data comes in the form:

$$(\mathbf{S}^i, P^i) = (S^i(1), S^i(2), S^i(3).., S^i(k), P^i). \tag{4.8}$$

where P^i is the dependent variable representing the class label of the ith time slot. The class labels P^i are assigned by using one of the splitting criteria [124] explained as follows

- Entropy: It is given by

$$Entropy(t) = -\sum_{id=0}^{Z} P_r(id|t)\, log_2 P_r(id|t). \tag{4.9}$$

  where $P_r(id|t)$ denote the fraction of records belonging to class id at a given node t and Z represents the total number of classes. In our approach, $Z = 1$ i.e. $id = [0, 1]$, where '0' represents the available class and '1' represents the occupied class. The smaller entropy implies that all records belong to the same class.

- Gini Diversity index (gdi): It a measure of node's impurity and is given by

$$gdi = 1 - \sum_{id} [P_r(id|t)]^2 \tag{4.10}$$

For a pure node (one class), gdi will be zero. It will be discussed in Section 4.4.3 that how splitting criteria affects the classification accuracy of DT.

4.3.3 Support Vector Machines

SVM is a discriminative classifier with high accuracy. The phenomenon of over fitting [1] often happens in DT, however SVM tends to be resistant to over-fitting and can be used for online learning [2]. There are two types of classifiers in SVM: linear SVM for separable data [3] and non-linear SVM for non-separable data.

The training feature and response vectors in our system is represented as $D = (P^i, \mathbf{S}^i)$ where $P^i \in \{0, 1\}$. Following the SVM [127], the two classes are separated by defining a hyper plane H, represented as $\mathbf{w}.\mathbf{S}^i = \rho$, where $\mathbf{w}$ represent the normal vector and ρ represents the constant separating occupied and idle classes given as:

$$P^i = +1 \quad when \quad \mathbf{w}.\mathbf{S}^i > \rho \qquad (Occupied\, class) \tag{4.11a}$$

$$P^i = 0 \quad when \quad \mathbf{w}.\mathbf{S}^i < \rho \qquad (Idle\, class) \tag{4.11b}$$

Two margins are defined on both sides of H to maximize the gap between two classes. The length of the margins is controlled by a parameter called box constraint Box_{ct}. The parameter Box_{ct} controls the relative weighting between the goal of making the margin small and ensuring that each dataset

[1] It is a condition when the ML model fits the training set very well but fails to generalize to the unseen examples [125].

[2] It is learning scenario in which training data is provided one example at a time, as opposed to the batch mode in which all examples are available at once. [126].

[3] Two sets of points A and B are linearly separable if there exists n real numbers $w_1, w_2, w_3, .., w_n$, such that every point $a_i \in A$ satisfies $\sum_{i=1}^{n} w_i.a_i > \rho$ and every point $b_i \in B$ satisfies $\sum_{i=1}^{n} w_i.b_i < \rho$, where ρ represents the constant separating two sets A and B [125].

have functional margins of at least 1. The optimal value of Box_{ct} is evaluated using a bio-inspired technique i.e. FFA in this approach explained as follows

4.3.4 SVM with Fire Fly Algorithm

The performance of SVM model is dependant on the selection of parameters. The problems like over-fitting can be avoided in SVM by optimally tuning the parameters that will increase accuracy consequently. For selecting the optimal parameters of SVM, a meta heuristic algorithm FFA is utilized. As discussed in in Section 3.3, meta heuristic methods can reach a solution by iteratively updating the candidate solution. Although the combination of SVM with FFA increases the computational complexity on one hand, but on the other hand it results an increase in classification accuracy. Due to its high accuracy, this algorithm has been used in [128] for estimating wind speed distribution, in [129] for global solar radiation prediction and in [130] for forecasting malaria incidences. The SVM+FFA algorithm is explained as follows.

In FFA, let X be a group of fire flies, $X = [l_1, l_2, ..l_X]$, initially located at specific positions $a_X = [a_{l_1}, a_{l_2}, ..a_{l_X}]$. Each fire fly moves and tries find a brighter fire fly, which has more light intensity than its own. The objective function $f(x)$, used for evaluating the brightness of the fire fly in this approach is the classification accuracy i. e. $f(x) = CA(a_X)$. When a fire fly, say l_1 finds another brighter fire fly l_2 at another location having more intensity compared to its own, it tends to move towards fire fly l_2. The change in position is determined as [131]

$$a_{l_1}^{v+1} = a_{l_1}^{v} + \beta_0 e^{-\psi_{l_1 l_2} r_{l_1 l_2}^2}(a_{l_2}^{v} - a_{l_1}^{v}) + \alpha(rand - 0.5) \tag{4.12}$$

where v represents the number of iterations, a_{l_1} and a_{l_2} represents the position of fire fly l_1 and l_2 respectively, α, β_0 and $\psi_{l_1 l_2}$ are constants, $rand$ is a uniformly

distributed random number and $r_{l_1 l_2}$ represents the Euclidean distance between l_1 and l_2. For the proposed approach, the starting positions of the X fire flies are initialized, where the position of each fire fly represents the value of box constraints Box_{ct} and the best position of the firefly will be selected for evaluating the optimal value of Box_{ct}, to be further used for classification. The FFA runs for specific number of iterations to achieve convergence as defined by the user. The pseudo code for FFA is given below.

Algorithm 1 Fire Fly Algorithm

```
while v < Maximum Iterations do
  for l1 = 1 to X do
    for l2 = 1 to l1 do
      if CA(a_{l2}^v) > CA(a_2^v)
      Fire Fly l1 will move towards firefly l2 using (4.12)
      Calculate new solutions by computing CA for each fire fly
    end for
  end for
  Compare the fireflies and evaluate the best one
end while
```

4.3.5 Hidden Markov Models

It is an unsupervised algorithm for modelling the time series data. The motivation to use the unsupervised algorithm is that it does not need the training phase. In HMM, the sequence of states can be recovered by an analysis of the sequence of observations. The state alphabet set U and observations alphabet set G are given as

$$U = (u_1, u_2, ...u_N) \tag{4.13}$$

$$G = (g_1, g_2, ...g_M) \tag{4.14}$$

where u_1 and u_2 represent the states when $P^i = 0$ and $P^i = 1$, respectively in the proposed methodology. The observations g_1 and g_2 represents the value of OC^i corresponding to each P^i state. The fixed state sequence (Q) of length T

and corresponding observations O is represented as:

$$Q = q_1, q_2, .., q_T \tag{4.15}$$

$$O = o_1, o_2, .., o_T \tag{4.16}$$

HMM is formally defined as

$$\lambda_{hmm} = (C_h, D_h, \pi) \tag{4.17}$$

where C_h is the transition array, D_h is the observation array and π is the initial probability array. HMM has two main steps of execution in this model, where in the first step, the sequence of observations O, transition probability matrix C_h and the emission probability matrix D_h are utilized to find the probability of observations O given hmm model λ_{hmm} shown as [133]

$$P_r(O|\lambda_{hmm}) = \sum_{Q} P_r(O|Q, \lambda_{hmm}) P_r(Q|\lambda_{hmm}) \tag{4.18}$$

where the probability of observations O for a specific state sequence Q is defined as

$$P_r(O|Q, \lambda_{hmm}) = \prod_{t=1}^{T} P_r(o_t|q_t, \lambda_{hmm}) = g_{q_1}(o_1) * g_{q_2}(o_2)....g_{q_T}(o_T) \tag{4.19}$$

In the second step, the hidden state sequence that is most likely to have produced an observation is decoded using Viterbi algorithm. The most likely sequence of states Q_L generated using Viterbi algorithm is matched with expected fixed state sequence Q to compute the classification accuracy. HMM can be also be supervised by adding two extra steps of execution in the above mentioned stepwise algorithm as follows:

Step(a): Use initial guesses of C_h and D_h to compute Q and O, that are used for computing $P_r(O|\lambda_{hmm})$ in forward algorithm

Step(b): Use O, D_h and C_h evaluated in step(a) to estimate the transition probability matrix $C_{h'}$ and emission probability matrix $D_{h'}$ using maximum likelihood estimation [134].

The $C_{h'}$ and $D_{h'}$ collectively form estimated HMM model (λ_e) that can be further used for evaluating $P_r(O|\lambda_e)$ and Q_L using forward algorithm and Viterbi algorithm respectively. Though, the classification accuracy increases by training HMM ahead however the complexity increases. For example, the computational time required for classifying the data of one day using HMM is 0.011s on Core i7, 8GB machine however it takes 0.1717s using trained HMM. The computational time increases with an increase in the amount of training data, therefore the complexity of supervised HMM will increase with an increase in the amount of data to be classified. The main motivation for using supervised HMM is to compare its performance with the other supervised algorithms and find the best supervised algorithm for analysing the spectrum occupancy.

4.4 Numerical Results and Discussion

In order to analyse the occupancy of the eight bands, the statistics of data for all bands from 880 to 2500 MHz are presented in Section 4.4.1. The classification criteria is explained in Section 4.4.2. The selection of the optimal parameters for each model is discussed in Section 4.4.3. In Section 4.4.4, the classification models with the optimal parameters are compared to find the best classifier using CA as the selection criterion. CA is defined as: $CA = \frac{\text{No. of correct classifications}}{\text{Total number of test samples}}$.

4.4.1 Statistics of Data

The CDF plot is shown in Fig.4.3 which gives the summarized view of all power ranges for the eight bands. It can be observed from Fig.4.3 that the eight bands can be categorized into two main groups. Group A includes 925 - 960 MHz, 1805 - 1800 MHz and 2110 - 2170 MHz while Group B has five bands: 880 - 915 MHz, 1710 - 1785 MHz, 1900 - 1920 MHz, 1920 - 1980 MHz and 2400 - 2500 MHz. The differences of Group A and Group B are explained as follows:

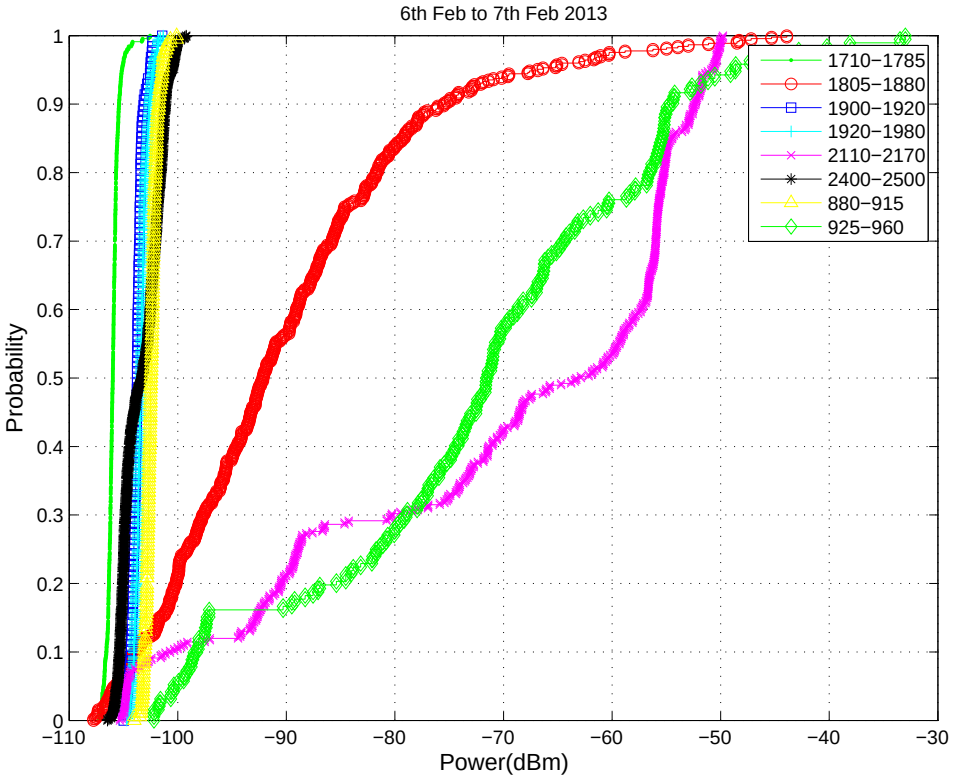


Figure 4.3: The CDFs for the eight bands between 880-2500 MHz.

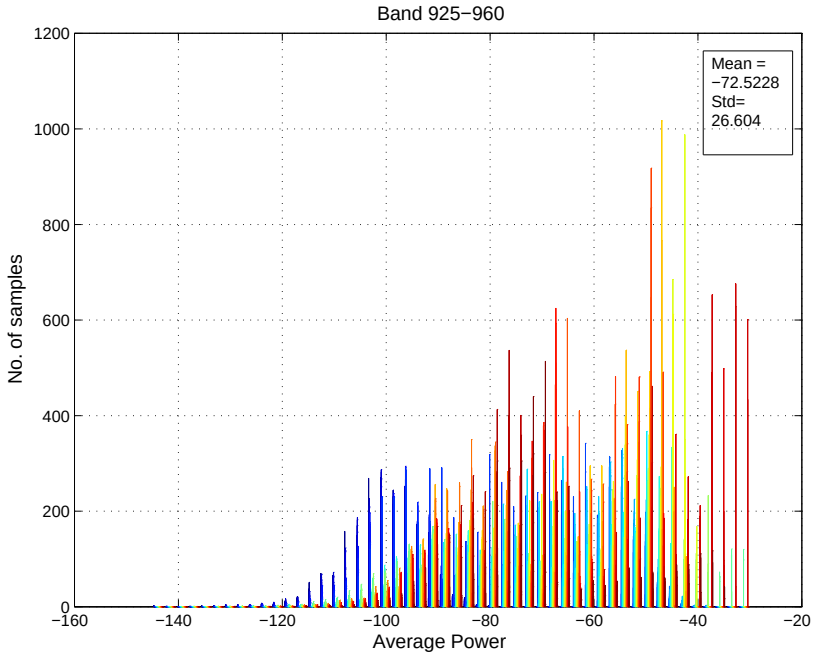
Average Power

The average power in Group A and Group B bands is evaluated using data of four months (Feb - June, 2013) and illustrated using histograms. The histograms for two Group A bands: 925-960 MHz and 1805-1880 MHz are shown in Fig 4.4(a) and Fig. 4.5(a) respectively, while the histograms for two Group B bands: 880-915 MHz and 1710-1785 MHz are shown in Fig 4.6(a) and Fig. 4.7(a) respectively. The data samples for 120 days are considered, that is the total number of samples utilized for each histogram are 172800 min (n= 120* 1440=172800, where 120 represents the number of days and 1440 represents the number of minutes in each day). It was observed that Group

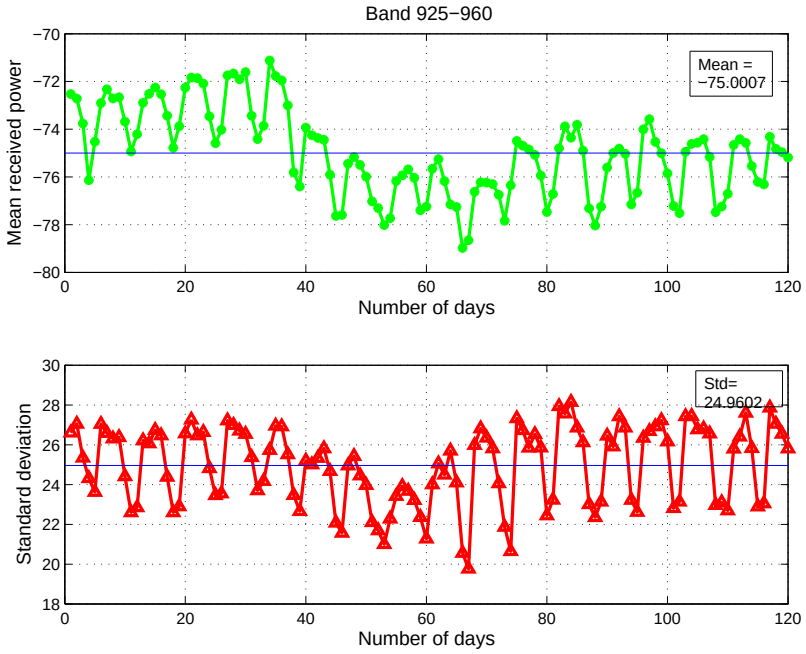
A bands have higher value of the average power than Group B bands. For example, the average power for 925-960 MHz in Fig. 4.4(a) is higher than the value of average power for 880-915 MHz in Fig.4.6(a). The similar behaviour is observed for 1805-1880 MHz and 1710-1785 MHz in Fig.4.5 (a) and Fig. 4.7(a) respectively. The average power for each day is also evaluated individually over the period of 120 days in Fig 4.4(b)-Fig. 4.7(b) respectively, where n=1440 min (since the number of minutes in 1 day=24 * 60=1440) are utilized. It was observed that the value of average power for each day is approximately same as the total value of average power evaluated using $n = 172800$ min in the histograms. The average power for all bands belonging to both Group A and Group B is illustrated in Table 4.2.

Standard Deviation (σ)

The maximum standard deviation of the power measurements in each time slot (σ^i) is also evaluated. It was observed in Fig 4.4(a) and 4.5(a) that Group A bands have higher values of σ^i compared to Group B bands in Fig. 4.6(a) and Fig.4.7(a) respectively. The σ^i is also evaluated for each day separately in Fig. 4.4(b)-Fig. 4.7(b) using n=1440 min. It was observed in Fig. 4.4(b)-Fig. 4.7(b) that σ^i for each day is approximately the same as over the period of four months, however Group A bands have higher value of average σ^i than Group B bands.

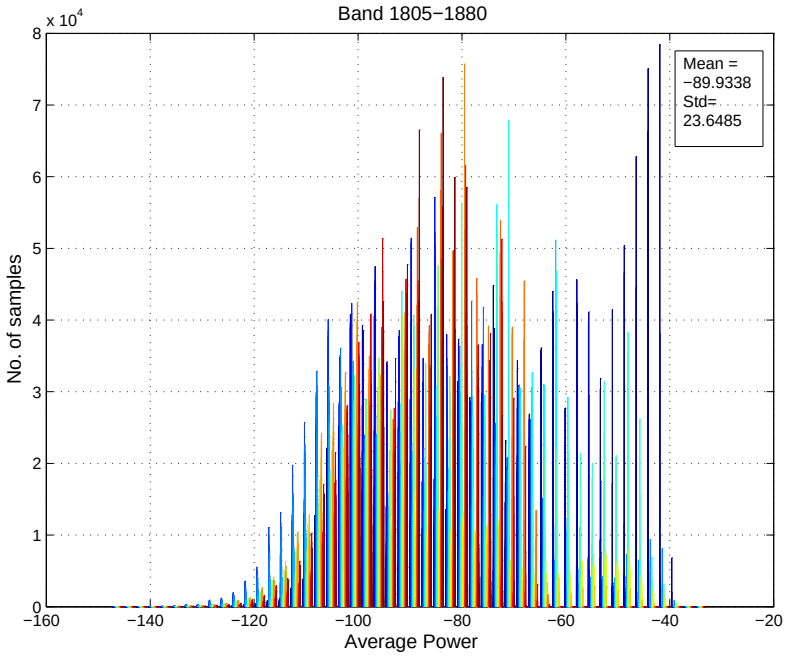


(a)

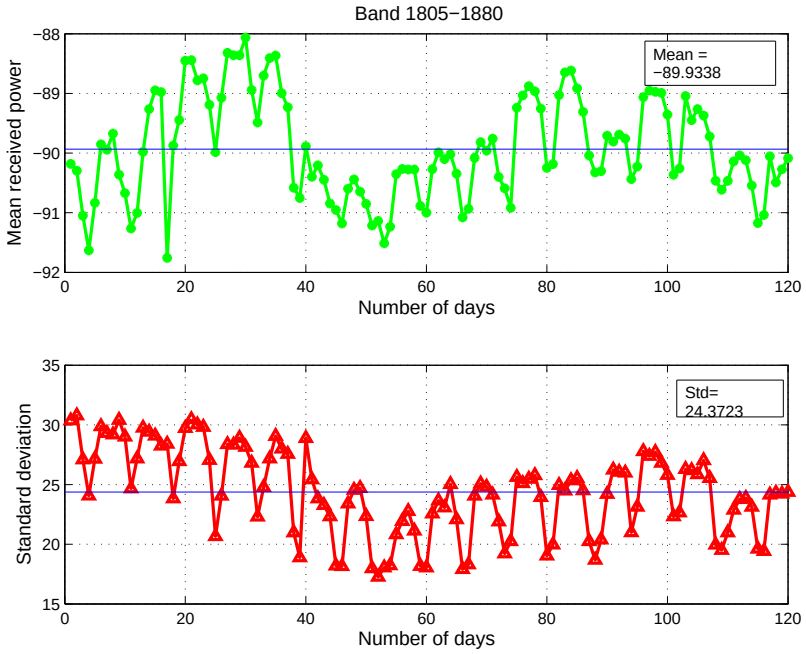


(b)

Figure 4.4: (a) Histogram of 925-960 MHz using $n = 172800$ and $k = 192$. (b) Mean received Power and σ^i of Band 925-960 MHz over the period of 120 days using $n = 1440$ and $k = 192$.

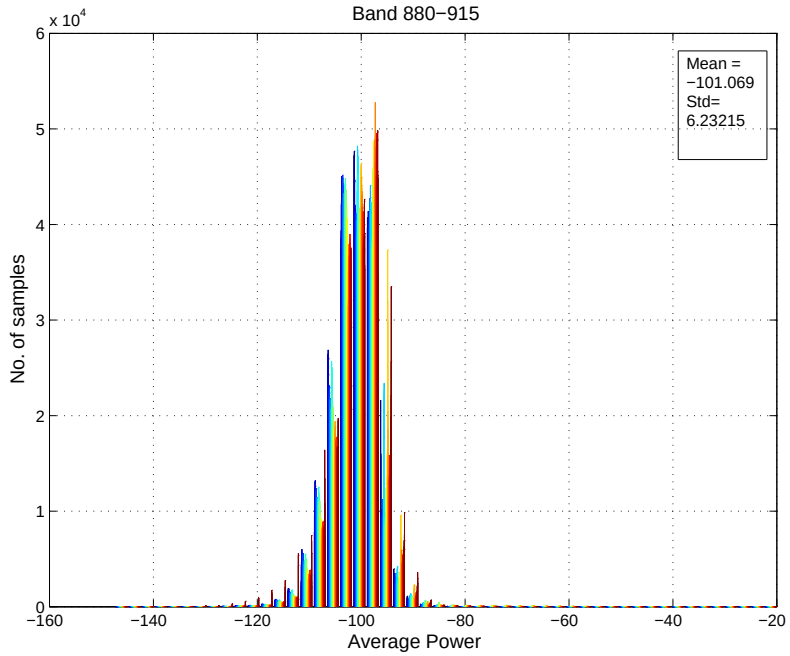


(a)

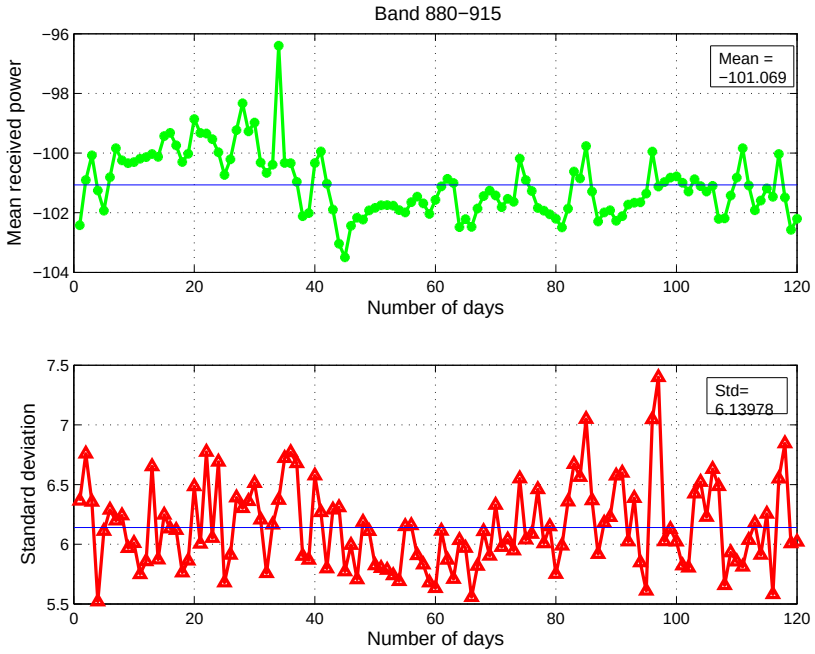


(b)

Figure 4.5: (a) Histogram of 1805-1880 MHz using $n = 172800$ and $k = 448$. (b) Mean received Power and σ^i of Band 1805-1880 MHz over the period of 120 days using $n = 1440$ and $k = 448$.

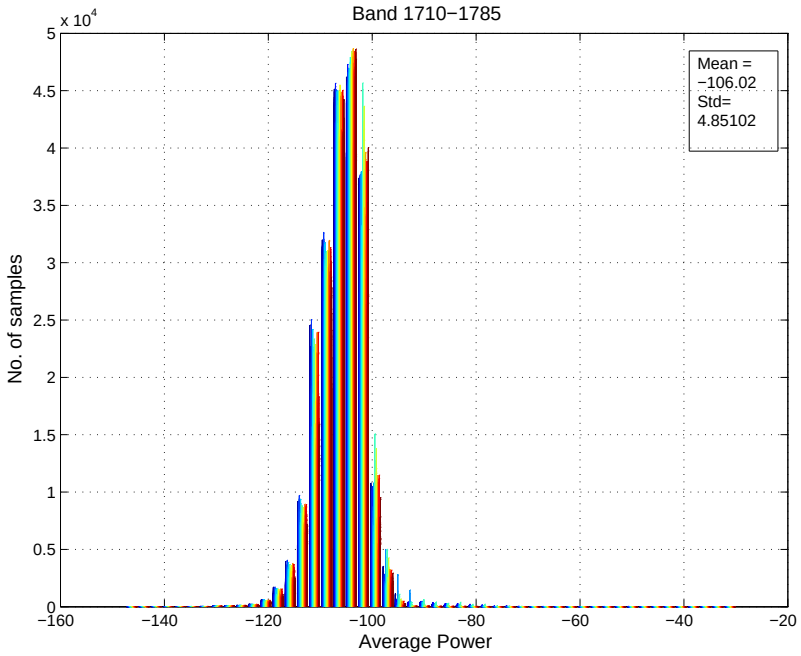


(a)

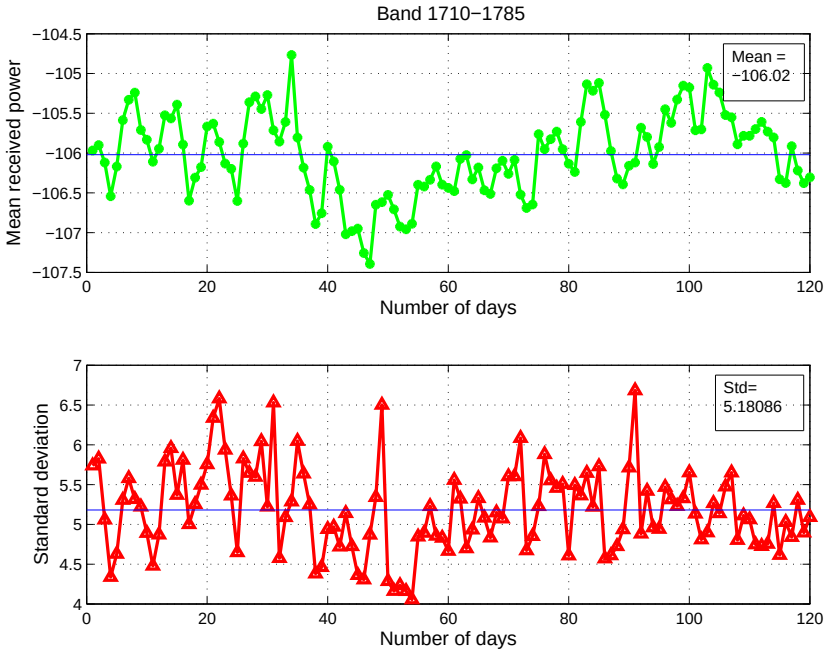


(b)

Figure 4.6: (a) Histogram of 880-915 MHz using $n = 172800$ and $k = 192$. (b) Mean received Power and σ^i of Band 880-915 MHz over the period of 120 days using $n = 1440$ and $k = 192$.



(a)



(b)

Figure 4.7: (a) Histogram of 1710-1785 MHz using $n = 172800$ and $k = 448$. (b) Mean received Power and σ^i of Band 1710-1785 MHz over the period of 120 days using $n = 1440$ and $k = 448$.

Occupancy VS Frequency

Due to stochastic noise and non-linearities, it is impossible to find the pure periodic and stationary signal in the real time. However reasonably periodic structures were noticed in some bands, when the relationship between occupancy and frequency is analysed. The occupancy for each frequency bin (OC^j) is given as ($OC^j = \frac{\sum_{i=1}^{n} S^i(j)}{n}$). It was observed that Group B bands can be classified as periodic bands while Group A bands do not have this property. Fig.4.8 and Fig. 4.9 illustrates the relationship between occupancy and frequency channels for two aperiodic and periodic bands from Group A and Group B respectively.

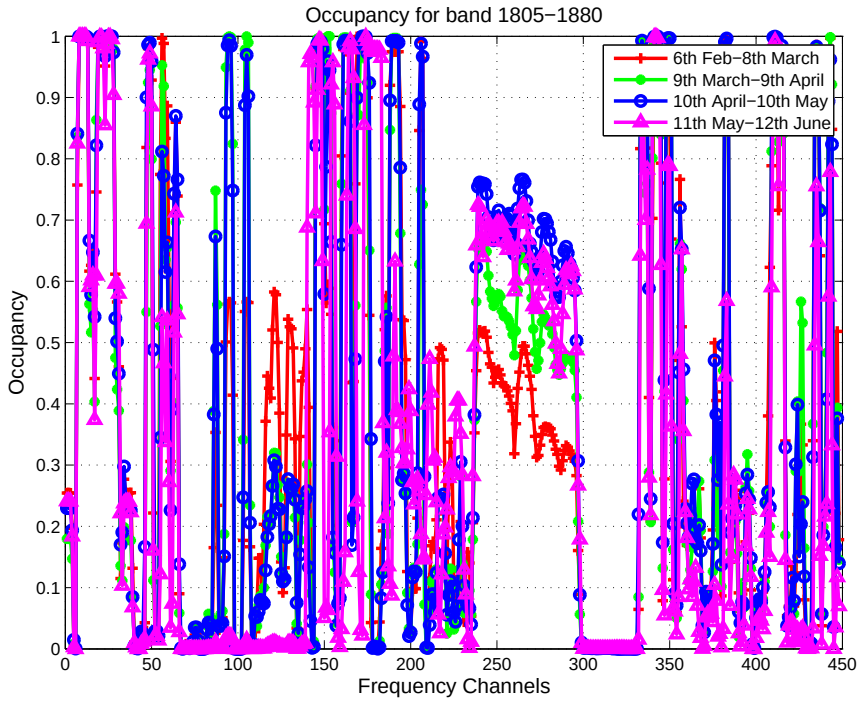
The periodicity may be caused by the usage pattern. For instance, the periodicity in each band lies in their uplink/ downlink usage pattern. For instance, the band 1710-1785 MHz is an uplink band, while the aperiodic band 1805-1880 MHz is the downlink. The uplink transmits data from the mobile user to base station so that its activity is completely determined by mobile user's periodic usage pattern. On the other hand, the downlink transmits the data from base station to the mobile user so that its activity is also affected by control and broadcast channels, making it less or non-periodic.

Occupancy VS Threshold

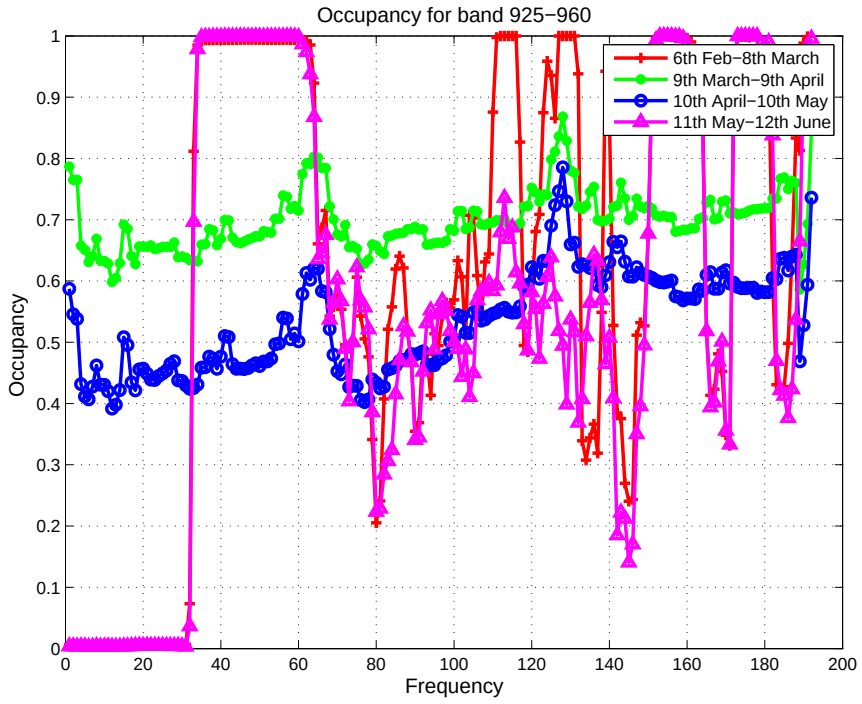
The results in Fig. 4.8 and Fig. 4.9 are computed using λ evaluated from otsu-thresholding method. The effect of setting a fixed threshold ($\lambda = -76$ dbm) on OC^j is also analysed in Fig. 4.10 and Fig.4.11 respectively. The fixed threshold has affected occupancy pattern in both aperiodic and periodic bands. It was observed in Fig. 4.8(a) that average occupancy between $k = [250, 300]$ for Band 1805-1880 MHz is $OC^j = 0.5$ reduces to average $OC^j = 0.3$ in Fig. 4.10(a), when fixed threshold is used. It was also observed in Fig.

4.11, that occupancy patterns for both periodic bands are adversely affected with $\lambda = -76$ dBm, compared to Fig 4.9 when λ was evaluated using otsu's thresholding method. This is because the value of average power for periodic bands: 880-915 MHz and 1710-1785 MHz is around -101.06 dBm and -106.02 dBm, respectively. By setting a higher value of fixed threshold than the average power, allows less number of signals to be sensed as primary signals, which eventually decreases the percentage occupancy.

From this, it was concluded that threshold selection plays an important role for analysing occupancy. Therefore, the minimum and the maximum values of the power for each band are considered and tested using seven values of thresholds in this range. It was noticed for all bands that occupancy monotonically decreases when the value of threshold increases. These results have proved that larger value of threshold will classify less samples as occupied. The effect of different threshold levels on mean occupancy for one band from Group A (1805-1880 MHz) and Group B (880-915 MHz) is shown in Fig. 4.12.

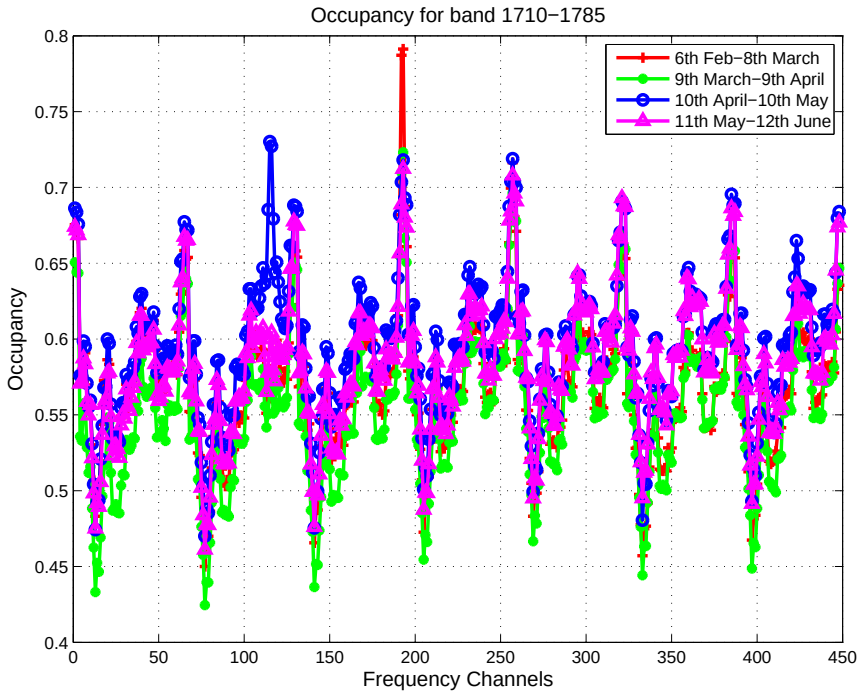


(a)

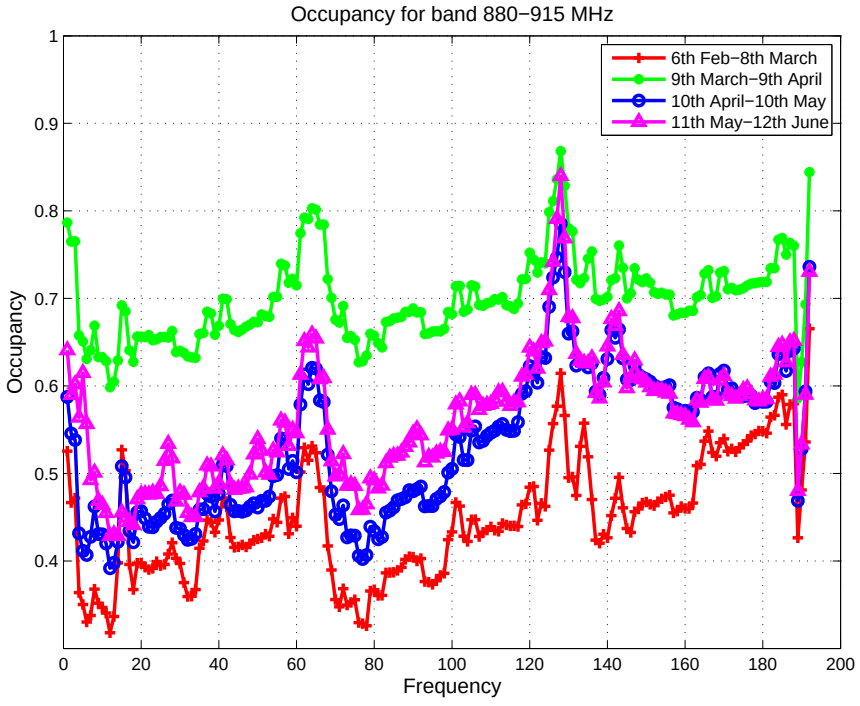


(b)

Figure 4.8: Occupancy VS spectrum frequency for aperiodic bands (a) 1805-1880 MHz (b) 925-960 MHz.

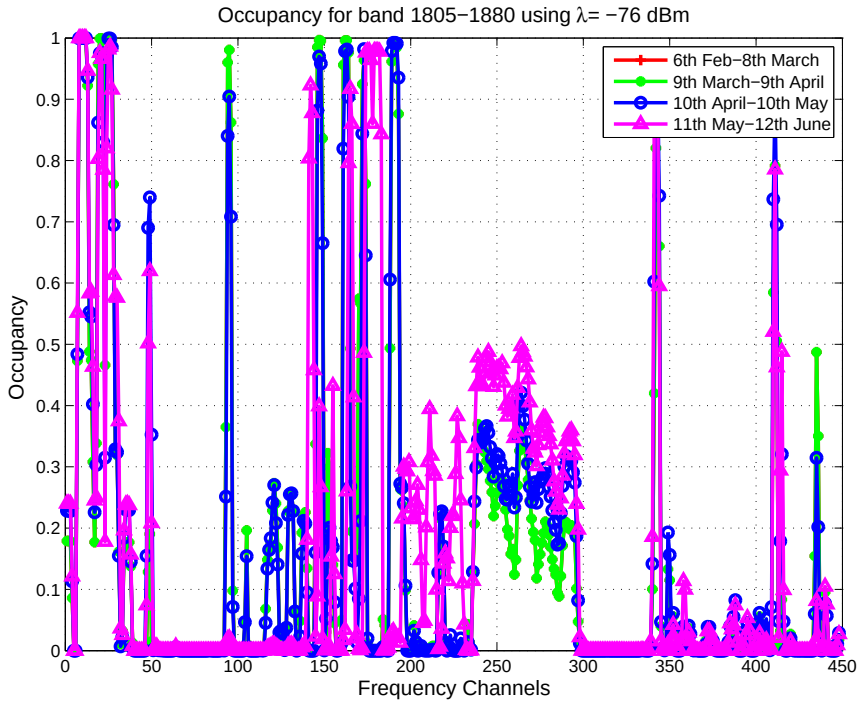


(a)

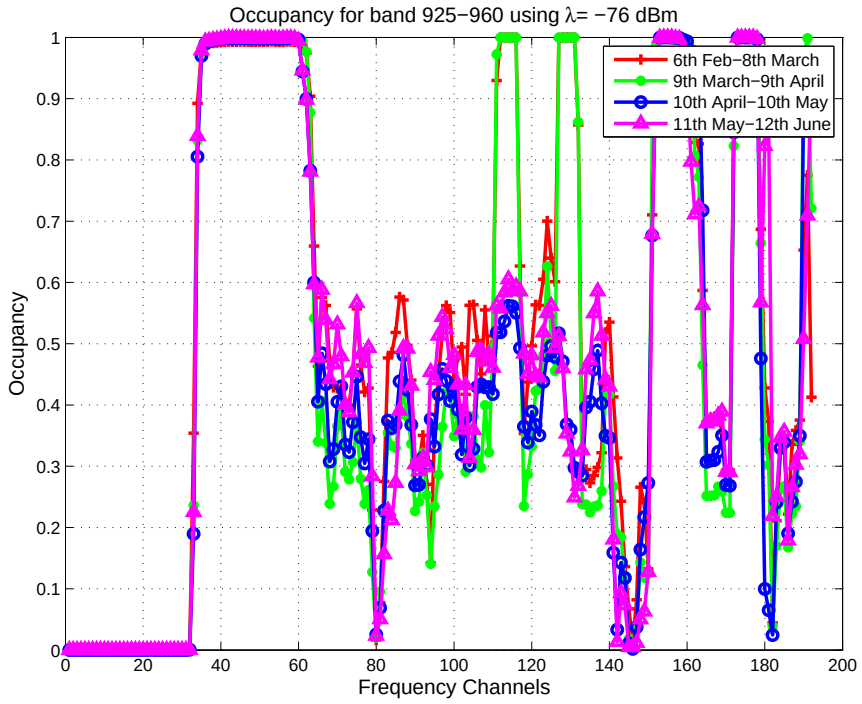


(b)

Figure 4.9: Occupancy VS spectrum frequency for periodic bands (a) 1710-1785 MHz (b) 880-915 MHz

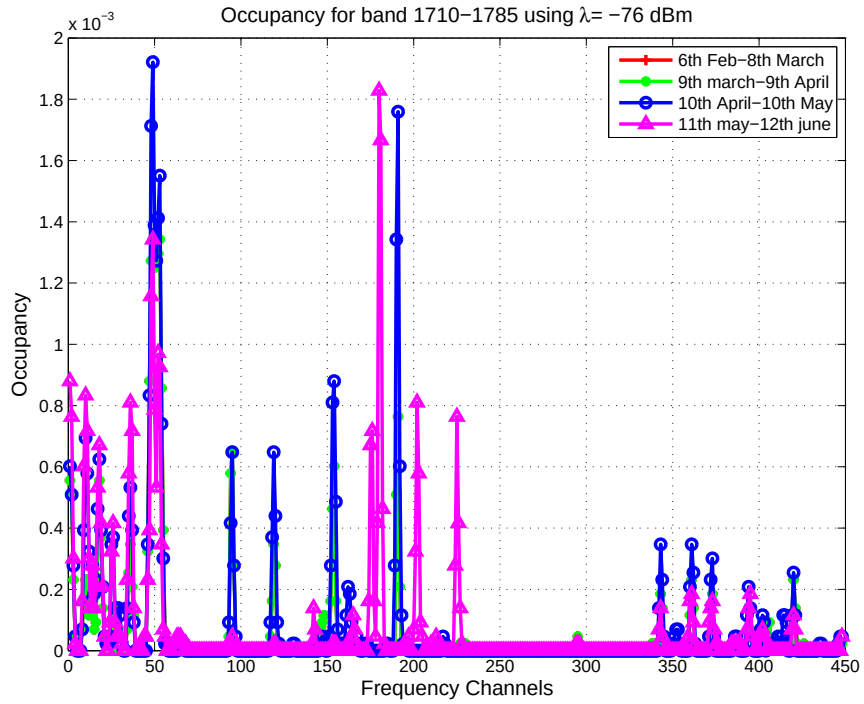


(a)

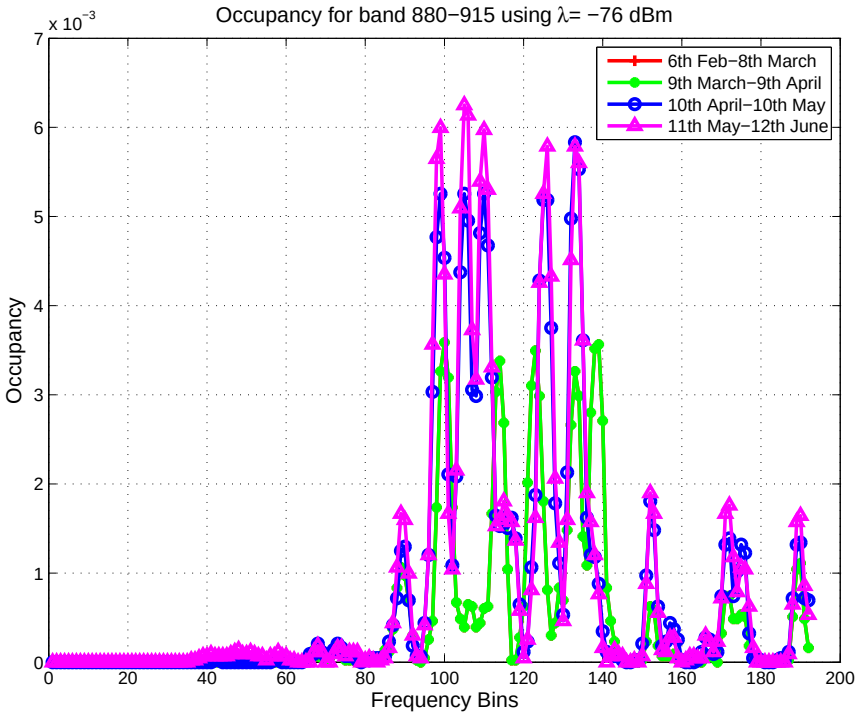


(b)

Figure 4.10: Occupancy VS spectrum frequency for aperiodic bands using -76dBm (a) 1805-1880 MHz (b) 925-960 MHz.



(a)



(b)

Figure 4.11: Occupancy VS spectrum frequency for periodic bands using -76dBm (a) 1710-1785 MHz (b) 880-915 MHz.

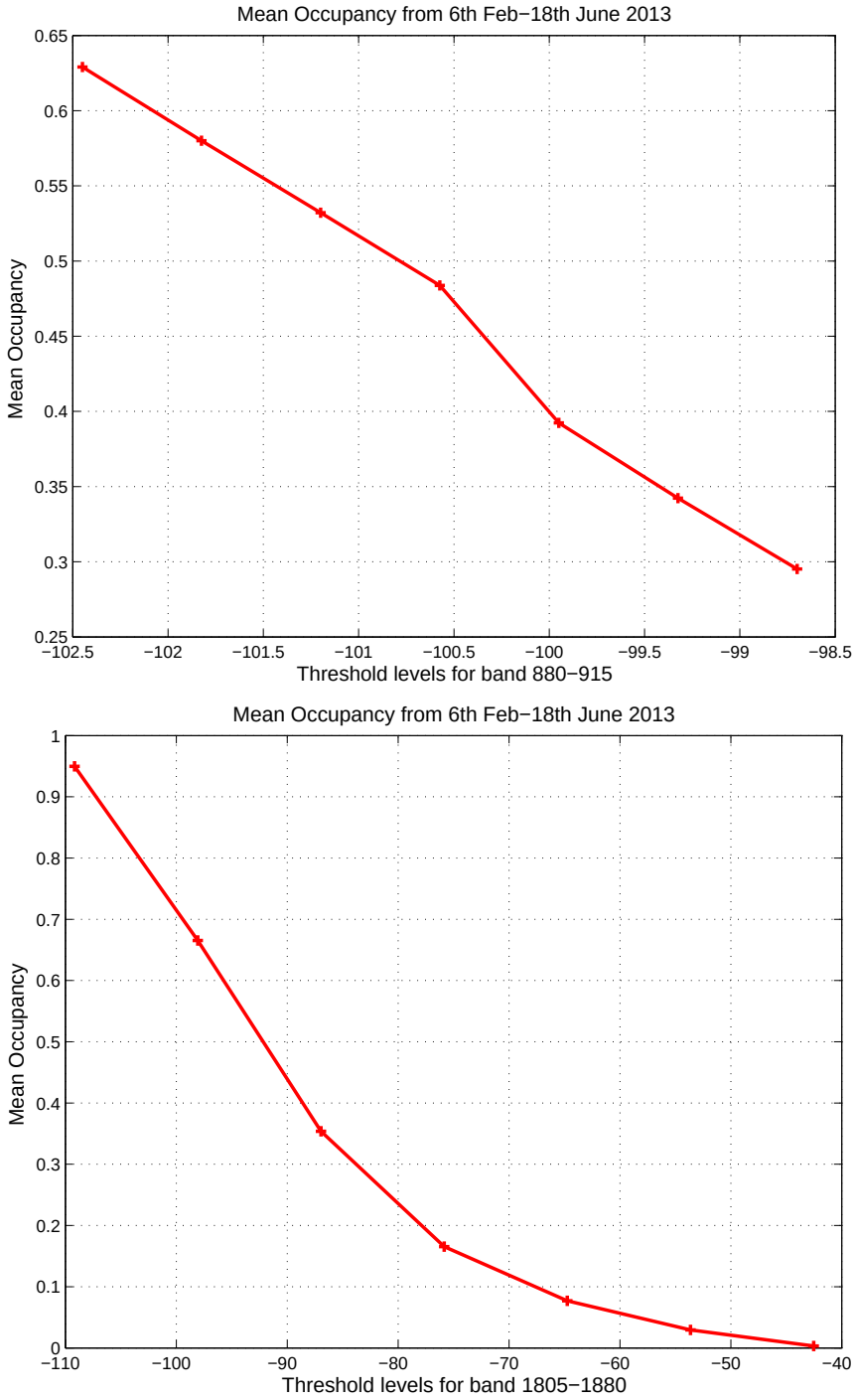


Figure 4.12: Effect of different threshold levels on mean occupancy for (a) 880-915 MHz (b) 1805-1880 MHz.

Occupancy VS Time

The main motivation of calculating occupancy w.r.t time as it can help to predict future variations in occupancy for each day. The occupancy in each *ith* time slot is calculated using (4.3), given as ($OC^i = \frac{\sum_{j=1}^{k} S^i(j)}{k}$), which determines the occupancy for each minute. The occupancy of each day is evaluated by determining the occupancy of each minute in 24 hours of the day and then taking average of them. In Fig. 4.13, the occupancy for each day is evaluated for the Band 880-915 MHz and Band 925-960 MHz over the period of 120 days. It was observed that average occupancy for Band 880-915 MHz is 57%, while it is 46% for the Band 925-960 MHz. The average occupancy for Band 925-960 MHz is smaller in value than Band 880-915 MHz because it was also observed that occupancy of 925-960 MHz decreases from 18th March (40th day) to 7th May (80th day). The occupancy for Band 925-960 MHz shows periodic patterns on weekly basis while Band 880-915 does not represent any specific pattern.

The statistics of frequency bands has been presented in Table 4.2. It was observed that that frequency bands can be categorized into two groups (A and B) on the basis of mean power, standard deviation and their periodicity. The average occupancy for each frequency bin (OC^j) is compared with the average occupancy in each time slot (OC^i) for all eight frequency bands. It was observed in Table 4.2 that OC^j is greater than OC^i for all bands.

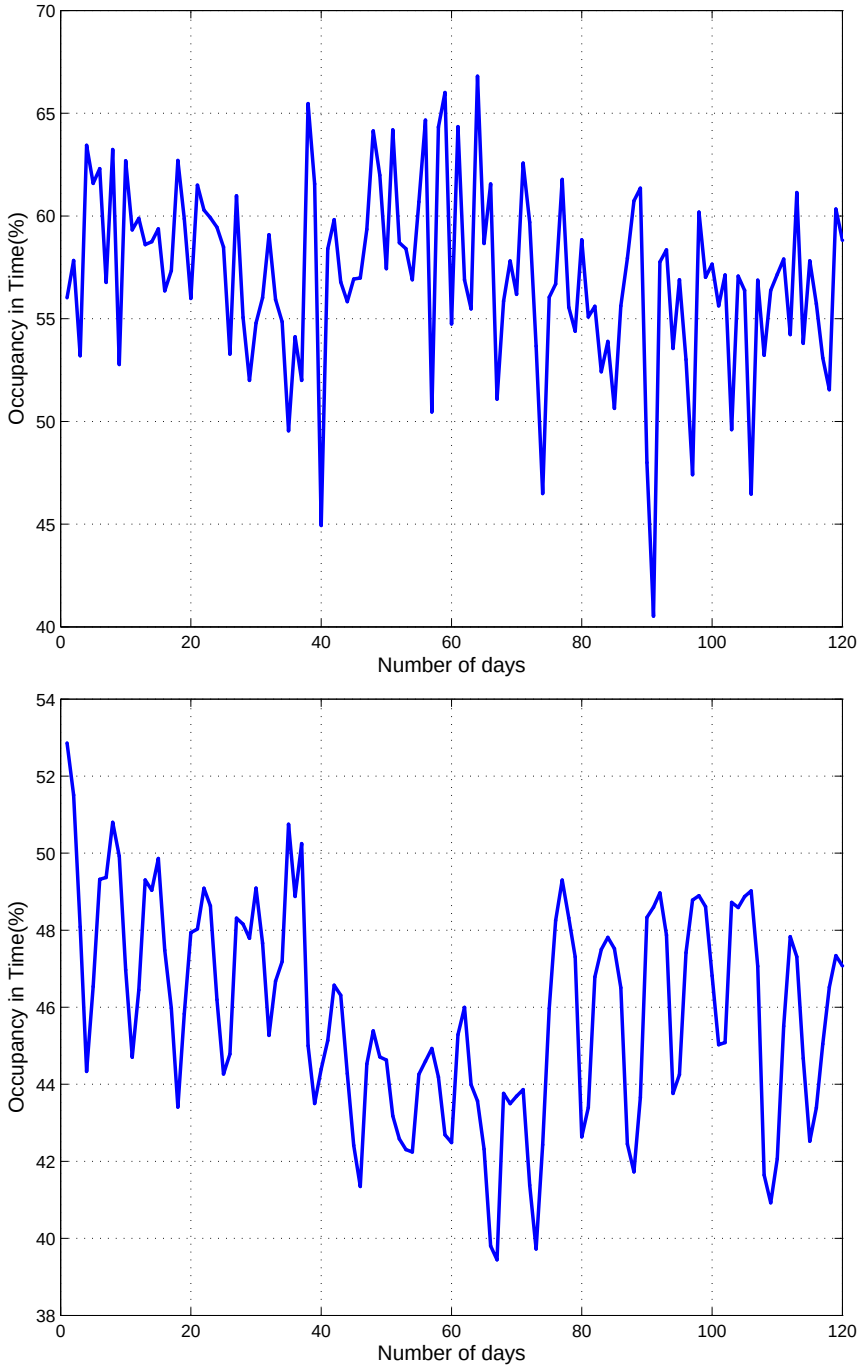


Figure 4.13: Occupany VS time using $k = 192$ and $n = 172800$ (a) 880-915 MHz (b) 925-960 MHz.

Group	Band (MHz)	Mean Power (dBm)	σ^j	σ^i	OC^i (%)	OC^j (%)	Periodic
A	925-960	-75.07	23.43	24.93	45.8	65.8	x
	1805-1880	-89.9338	21.054	23.64	32.45	86.21	x
	2110-2170	-71.1649	23.66	12.097	53.90	59.41	x
B	880-915	-101.0693	16.67	6.2322	57.06	79.1	✓
	1710-1785	-106.0201	9.2423	4.8510	57.71	66.52	✓
	1900-1920	-103.8796	6.4199	4.0611	62.25	75.17	✓
	1920-1980	-103.7396	15.58	4.18	61.48	72.12	✓
	2400 -2500	-103.5026	16.8873	7.9524	31.17	36.02	✓

Table 4.2: Statistics of Frequency Bands measured at the University of Warwick between 880-2500 MHz.

4.4.2 Classification Criteria

Following the discussion in Section 4.2.3, this subsection studies the choice of U_{oc}, L_{oc}, con^i and B.

Optimal Occupancy Range $[L_{oc}, U_{oc}]$

The data of Day1 (1-1440 min), Day 2 (1441-2448 min) and Day 5 (7200-8640 min) for Band 880-915 MHz is analysed in this section by setting four different values of threshold: $\lambda = [-102, -104, -106, -108]$ dBm . The parameters U_{oc} and L_{oc} will be selected by M_s, which represents the occupancy split that divides the data into occupied and idle classes. It varies from 0.1 to 0.9 with a step size of 0.1. It is observed in Fig. 4.14, that the value of CA depends on day and the value of threshold.

The actual value of OC^i always lies in a certain range, $[L_s, U_s]$, where L_s represents the lowest value of OC^i and U_s represents the maximum value of OC^i. When $L_s <= M_s <= U_s$, two groups of classes $P^i = 0$ (available class) and $P^i = 1$ (occupied class) can be classified correctly. When $M_s > U_s$ or $M_s < L_s$, all the samples will be classified as one class because OC^i is a closed set whose values do not lie outside the range $[L_s, U_s]$. This explains why $CA = 1$ for $[L_{oc}, U_{oc}] = [0.1, 0.2]$ and $[L_{oc}, U_{oc}] = [0.65, 0.9]$ respectively, in Fig. 4.14 when $\lambda = -102$ dBm is used for Day 1 data. However $CA < 1$ for $[L_{oc}, U_{oc}] = [0.2, 0.65]$. Thus, the classification cannot be performed when $M_s > U_s$ or $M_s < L_s$. Therefore optimal occupancy range is $[L_{oc}, U_{oc}] = [0.2, 0.65]$ for $CA < 1$. It was also observed in Fig. 4.14 using Day 1 data that there are four different choices of thresholds available for the case when $CA < 1$. In the proposed approach, the specific value of threshold is selected that contains the largest number of values between L_{oc} and U_{oc}. Following this, $\lambda = -102$ dBm for Day1, Day2 and Day5 is selected as the optimal threshold which ensures

the maximum amount of samples between L_{oc} and U_{oc}. Furthermore, the L_{oc} and U_{oc} for three days can be selected using knowledge of optimal threshold, for instance $[L_{oc}, U_{oc}] = [0.2, 0.65]$ for Day 1, $[L_{oc}, U_{oc}] = [0.4, 0.85]$ for Day2 and $[L_{oc}, U_{oc}] = [0.2, 0.80]$ for Day 5 are selected respectively. The optimal values of λ, U_{oc} and L_{oc} are further used for finding B for each day.

Optimal Free Consecutive Slots (B)

As discussed in Section 4.2.3, the PU status (P^i) can be evaluated for Condition 1 and Condition 4 using L_{oc} and U_{oc}, where L_{oc} and U_{oc} represents the optimal occupancy range as discussed above. However when $L_{oc} <= OC^i <= U_{oc}$ for Condition 2 and Condition 3, then it is difficult to evaluate P^i without determining the value of B. For evaluating B, several values of con^i are assumed in Fig. 4.15 and the specific value of con^i that ensures highest CA, will be selected as the optimal value. The relationship between CA and con^i is shown in Fig. 4.15 using the optimal occupancy range ($[U_{oc}, L_{oc}]$) and the threshold evaluated using Fig. 4.14. For the proposed case, $con^i = 40$ is selected as B, because it ensures the highest CA for all three days.

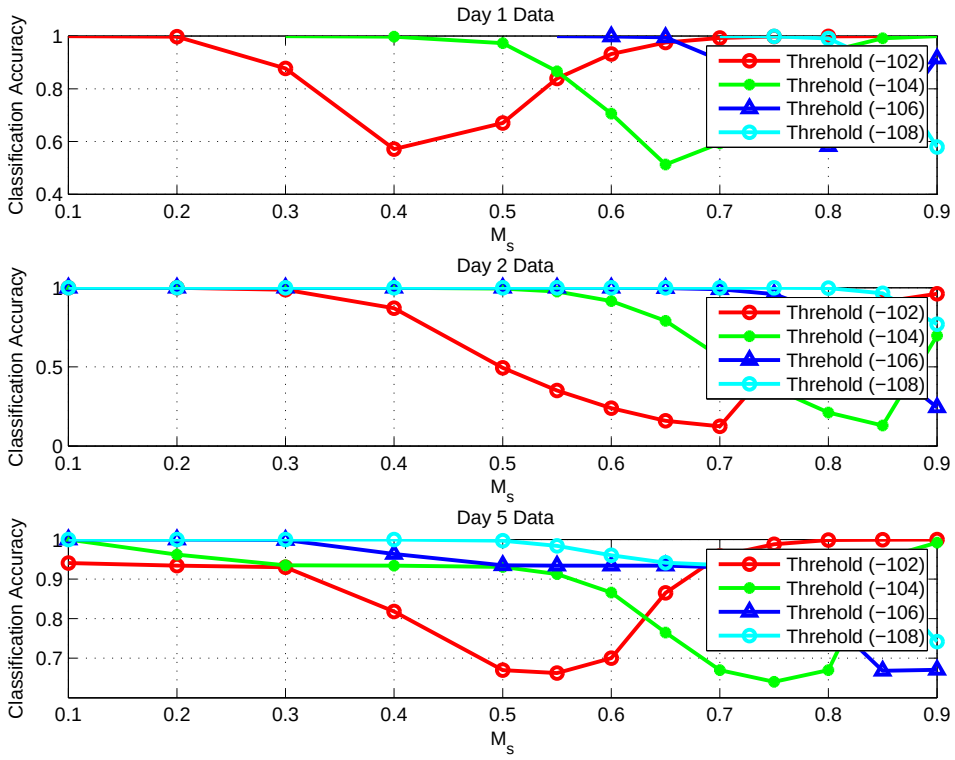


Figure 4.14: Selection of optimal threshold (λ) and optimal splitting range ($[U_{oc}, L_{oc}]$) for determining the classification criteria of three days data.

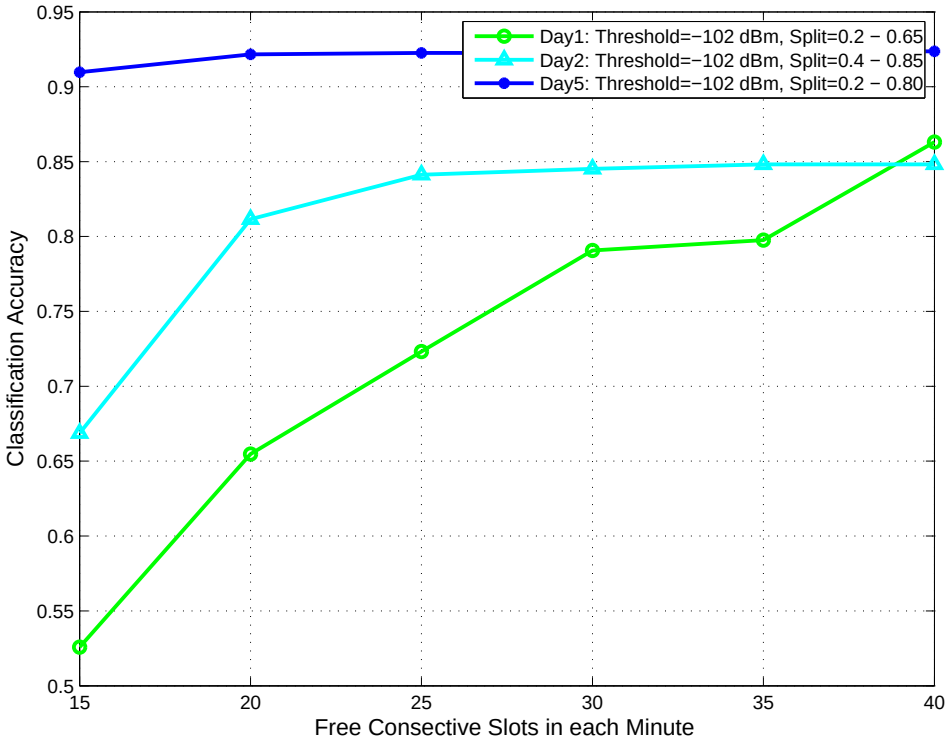


Figure 4.15: Selection of optimal number of free frequency slots using optimal splitting threshold and optimal occupancy range attained using Fig. 4.14.

4.4.3 Parameter Selection

The selection of parameters for the supervised models affects the performance of the classifiers. Considering the classification criteria in Section 4.4.2, the parameter selection for each algorithm is discussed as follows

NBC

As discussed in Section 4.3.1 that NBC is sensitive to the kernel selection, therefore four different kernel functions: normal, epanechnikov, triangular and uniform are used in order to select the best NBC kernel for classification in Fig. 4.16 for Band 880-915 MHz. It was observed in Fig. 4.16 that normal

kernel has outperformed all other kernel functions. The Gaussian kernel has attained the highest classification accuracy and it can be utilized for further analysis using NBC.

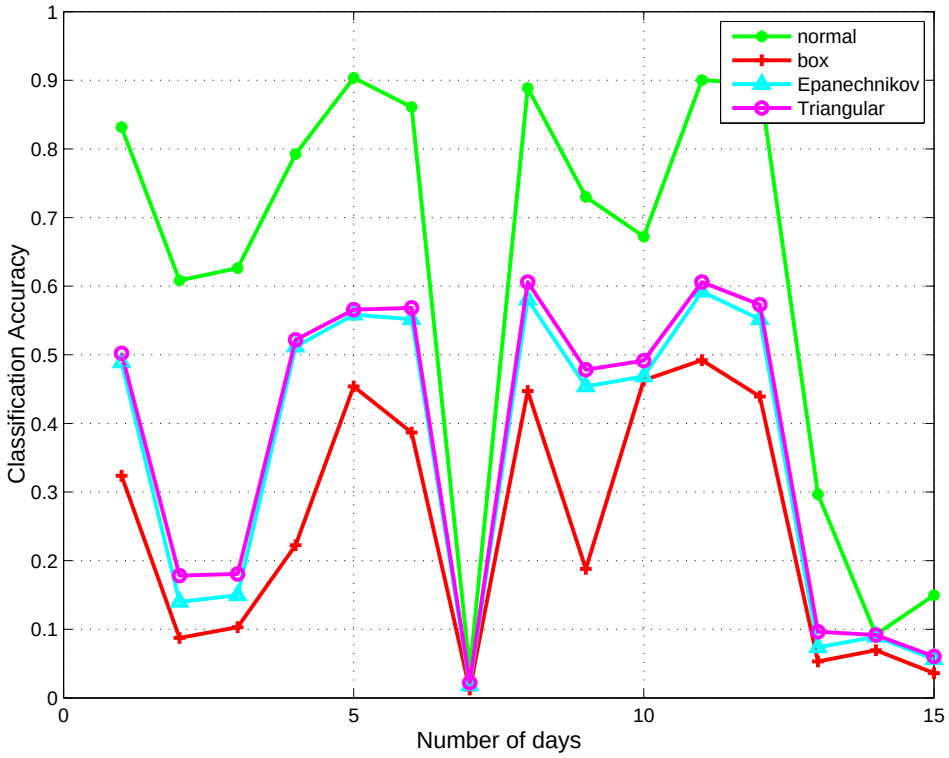


Figure 4.16: Effect of different kernel functions on classification accuracy using NBC

Decision Trees

The performance of decision trees also depend upon the splitting criteria. The criteria for splitting DT can be one of the following [124]:

The performance of DT is affected by the splitting criteria chosen for the analysis. The effect of using two different splitting criteria: entropy and gini diversity index (as explained in Section 4.3.2) is illustrated for DT in Fig.4.17. The minimum number of observation per node is varied between 1 to 50. It was observed in Fig.4.17 that gdi attains higher classification accu-

racy, when observation per node between 1 to 37 are considered and entropy outperforms gdi, when observation per node increases from 38 to 50. It was also observed that size of tree changes when minimum number of observation per node changes. With the decrease in number of minimum observations per node, the size of decision tree expands.

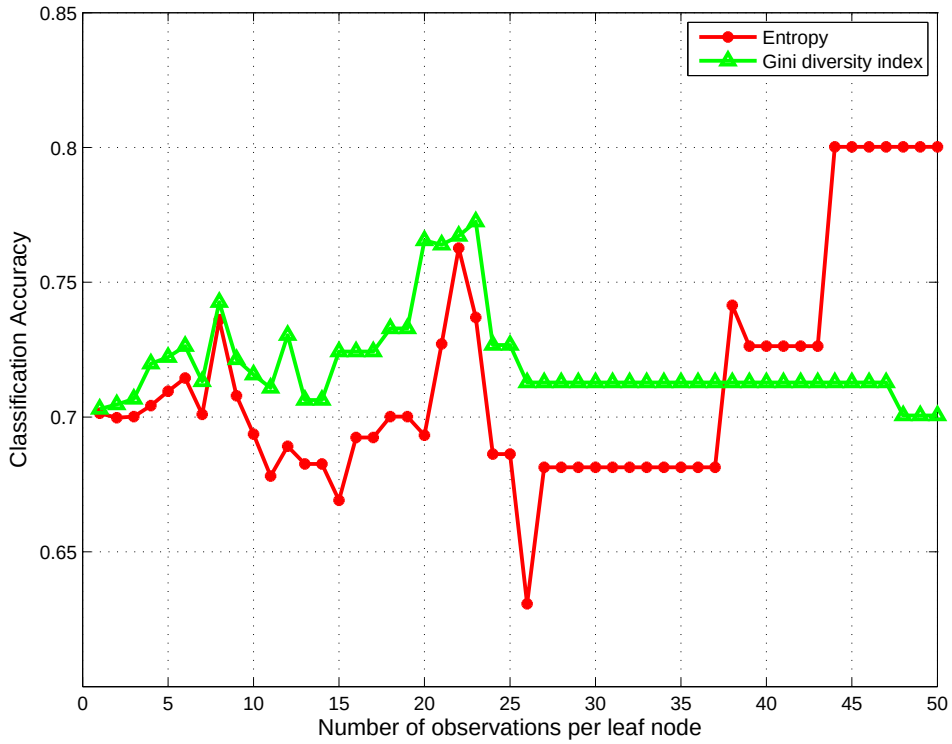


Figure 4.17: Effect of changing splitting criteria on classification accuracy, when the number of observations per leaf node ϵ [1,50].

SVM

In SVM, kernel selection is the most important task that affects the classification accuracy depending on the nature of the problem. Five kernel functions: polynomial, Gaussian, radial basis function, multi perceptron and linear are utilized in order to select the best kernel for classification using SVM. It was

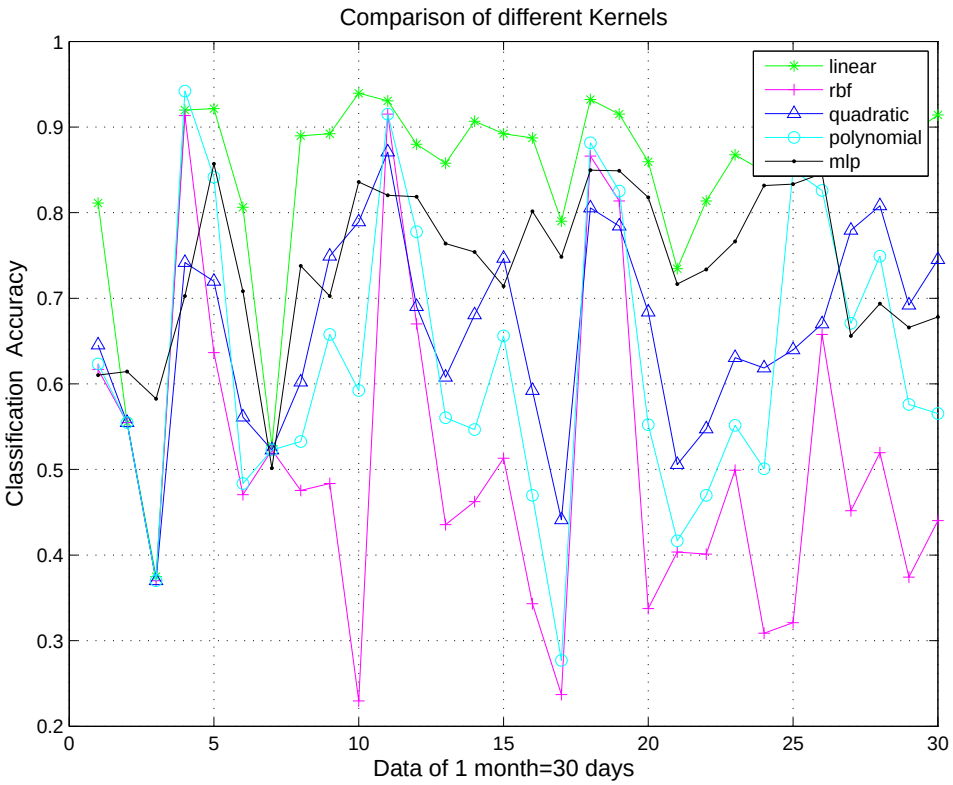


Figure 4.18: Effect of different kernel functions on classification accuracy using SVM

observed in Fig. 4.18 that linear kernel has attained the highest classification accuracy compared to others, when data of one month of Band 880 -915 MHz is analysed. This is because linear SVM performs best for two class case system and there are mainly two classes in the proposed system that needs to be separated: occupied and available classes. In general, the RBF kernel is a reasonable first choice for classification problems. RBF kernel nonlinearly maps samples into a higher dimensional space so it, unlike the linear kernel, can handle the case when the relation between class labels and attributes is nonlinear.

SVM with fire fly algorithm

The performance of SVM can be improved by optimizing the performance of the kernel function. The performance of the kernel is directly affected by the change in the value of the box constraints Box_{ct} (as explained in Section 4.3.4). Five different values of Box_{ct} are considered in Fig. 4.19 using linear kernel. It was observed in Fig. 4.19 that change in value of box constraints has brought change in the performance of the linear kernel. In proposed approach, FFA is utilized to find the optimal value of Box_{ct}.

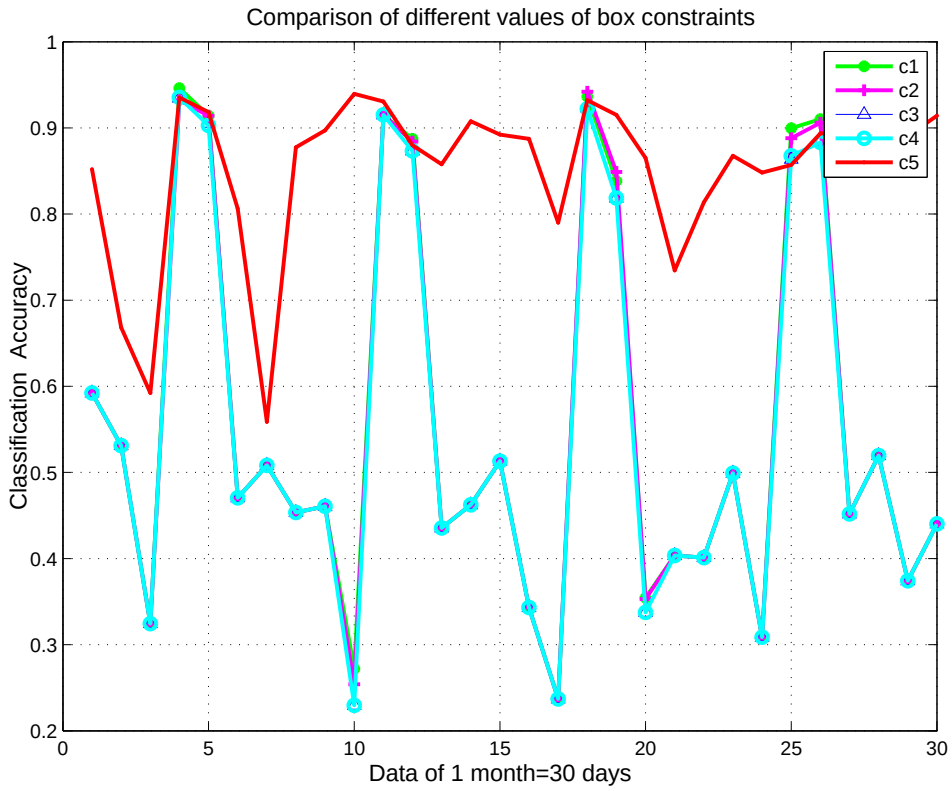


Figure 4.19: Effect of different values of box constraints on classification accuracy.

4.4.4 Model Performance Comparison

Following the discussion above, the performance of the ML algorithms is compared in this section using one month data of Band 880-915 MHz. The parameter selection tests (as discussed above in Section 4.4.3) can be utilized for selecting the optimal variables for each classifier. The optimal splitting range, optimal threshold and B will be selected corresponding to the data of each day as explained in Section 4.4.2.

Supervised vs Unsupervised Algorithms

The performance comparison of HMM, trained HMM, SVM, DT and NBC is illustrated in Fig.4.20 for 30 days using k=192. Each iteration represents 1 day. It was observed that trained HMM performed better than HMM, but worse than DT, NBC and SVM. The mean CA attained by trained HMM, HMM, SVM, DT and NBC is 0.6816, 0.4887, 0.8528, 0.8392, 0.7970 while the computational time consumed for each iteration using trained HMM, HMM, SVM, DT and NBC is 0.0205, 0.09066, 0.0135, 0.0163, 0.0095 seconds, respectively. Thus, SVM is the best in this case with highest CA and shortest time.

SVM with Fire Fly Algorithm

So far, the best overall performance is attained using the SVM technique. The performance of SVM is affected by the value of Box_{ct}. The optimized value of Box_{ct} is evaluated after running FFA for specific number of iterations using $\alpha = 1$, $\beta_0 = 2$ and $\psi_{l_1 l_2} = 1.3$. The optimal value the box constraints is selected and further used in SVM classification. Fig. 4.21(a) depicts that 'SVM+FFA' performs better than the conventional SVM in most of the cases. The mean CA attained by SVM+FFA, SVM, DT, NBC and HMM is 0.8728,

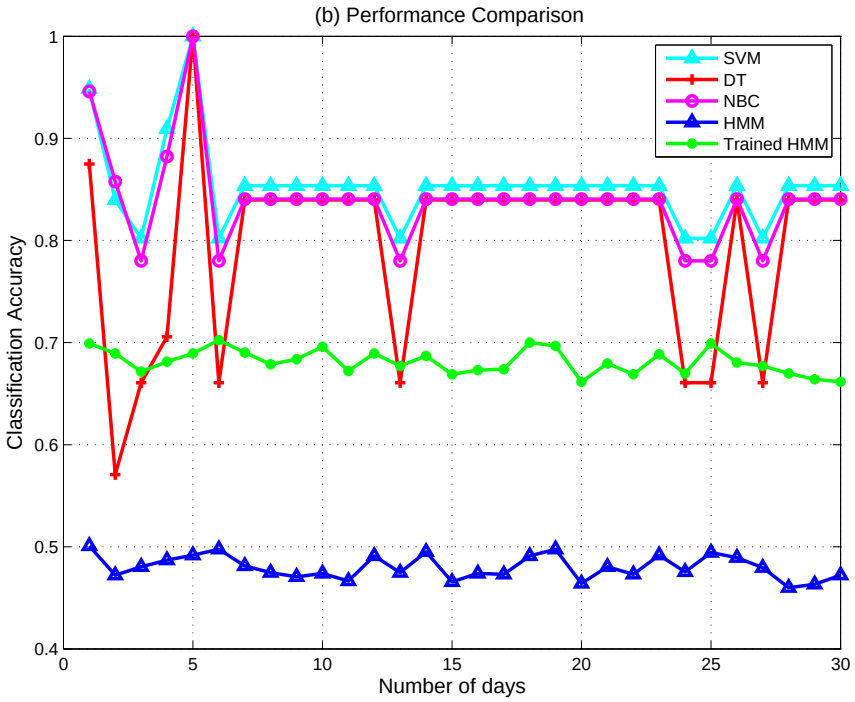


Figure 4.20: Performance Comparison using $k = 192$.

0.8499, 0.7970, 0.8392 and 0.4822, respectively.

Comparison with Statistical/ Probabilistic Model

In this section, the proposed ML classification framework is compared with the model in [112]. The inputs of this model are the statistical parameters extracted from the real time measurements. The outputs obtained from this model are the transmission times t_{ON} and t_{OFF}, where t_{ON} and t_{OFF} represents the busy and idle duration respectively. Based on t_{ON} and t_{OFF}, the model predicts the PUs occupancy for the testing data matrix.

Following [112], the PU occupancy of the testing data matrix is predicted using data of one day ($n = 1440$ and $k = 192$) of Band 880-915 MHz. In order to compare the statistical model with the proposed approach, the occupancy (OC^i) evaluated using the statistical model is transformed to PU

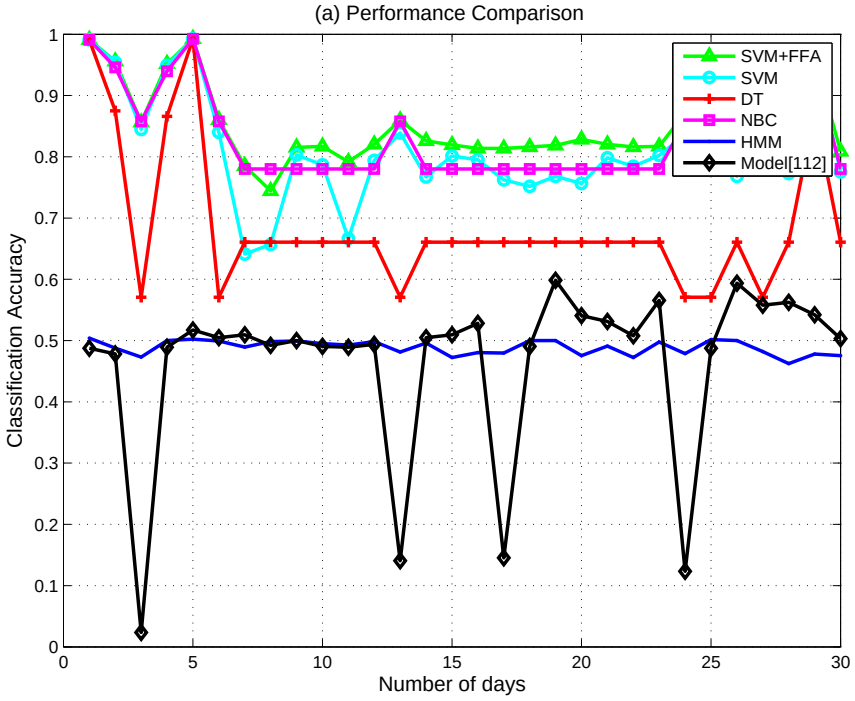
status labels (P^i_{eval}), following the criteria explained in Section 4.2.4. This is because, the validation in the proposed approach is performed using CA, where P^i_{eval} is checked against the value of P^i_{test}. It was observed in Fig. 4.21(a), that ML algorithms have attained higher value of CA than statistical model. The mean CA for the statistical model is only 0.45, much lower than average value of CA attained by DT, NBC, SVM and SVM +FFA.

Blocking Probability

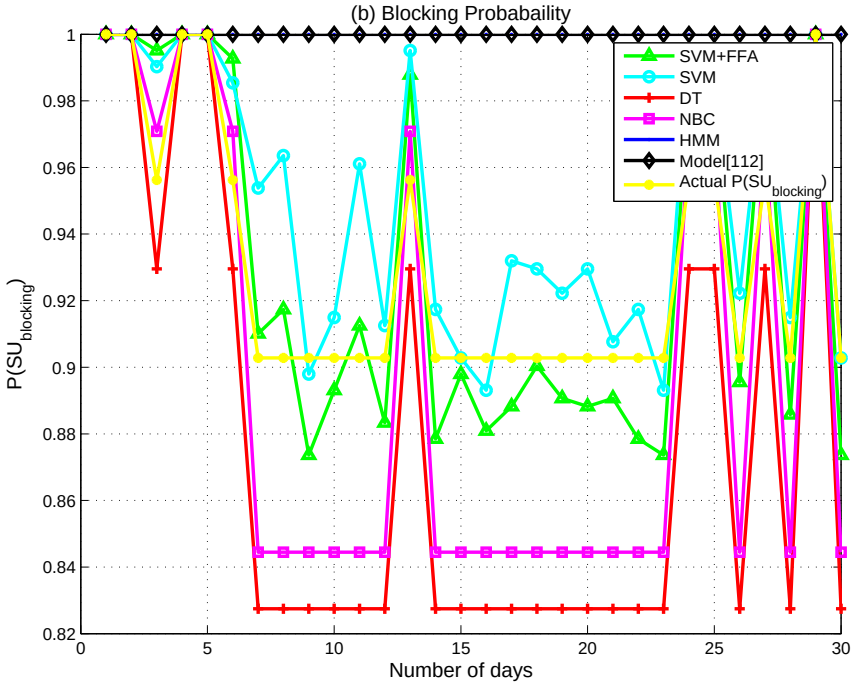
The blocking probability is computed using SVM+FFA, SVM, DT, NBC, HMM and the statistical model [112] in Fig. 4.21(b). It is further compared with the expected $P_r(SU_{blocking})$ to compute the difference between evaluated and expected values. It is evident in Fig. 4.21(b), that SVM+FFA has predicted the $P_r(SU_{blocking})$ with minimum difference and is very close to the expected one. The expected blocking probability is 0.9191 in Fig. 4.21(b) while the predicted $P_r(SU_{blocking})$ using SVM+FFA, SVM, NBC, DT, HMM and statistical model is 0.9264, 0.9322, 0.9638, 0.9577, 1 and 1, respectively. The $P_r(SU_{blocking})$ for HMM and statistical model is always 1, which implies that both HMM and statistical model have failed to find a single block of consecutive free time slots having length out_{su}.

Supervised vs Unsupervised Algorithms using different Training/ Testing Data vectors

The detailed comparison of supervised and unsupervised algorithms is presented in Table 4.3 using the mean classification accuracy (CA) and mean computational time as the performance metric for different sizes of training and testing data. It was observed that computation time for all supervised algorithms increases with an increase in the size of the training data. SVM+FFA



(a)



(b)

Figure 4.21: (a) Performance comparison of ML algorithms: SVM, DT, NBC, HMM, 'SVM+FFA' and statistical model [112] using $k = 192$ for 30 days. (b) Comparison of expected and evalauted $P(SU_{blocking})$ using SVM, DT, NBC, HMM, 'SVM+FFA' and statistical model [112] using $k = 192$ for 30 days.

has attained the highest CA but with the longest computation time in most cases.

4.5 Conclusion

The spectrum measurements have been used to evaluate the spectral opportunities for cognitive radio. Insightful relationships between the spectrum occupancy and its parameters: time, frequency channels and decision threshold have been derived. The occupancy has been modelled as a function of these three parameters. The occupancy model has been further used for classification of the idle and busy channels using the supervised and the unsupervised algorithms. It has been shown from the results that SVM is the best classification algorithm which has attained the highest classification accuracy. The performance of SVM has been further improved by using fire fly algorithm.

Number of Iterations/ Number of days =5			
Training data, Testing data	Technique	Mean CA	Mean Computational Time(s)
15 %, 85 %	DT	0.7612	0.0132
	SVM	0.8945	0.0128
	SVM + FFA	0.9034	3.0412
	HMM	0.4925	0.0241
	NBC	0.8714	0.0084
23 %, 77 %	DT	0.8034	0.0178
	SVM	0.9124	0.0129
	SVM + FFA	0.9148	3.5701
	HMM	0.4903	0.0343
	NBC	0.8960	0.0090
30 %, 70 %	DT	0.8028	0.0198
	SVM	0.9143	0.0153
	SVM + FFA	0.9189	3.8947
	HMM	0.4841	0.0191
	NBC	0.9064	0.0098
Number of Iterations/ Number of days =10			
Training data, Testing data	Technique	Mean CA	Mean Computational Time(s)
15 %, 85 %	DT	0.7823	0.0140
	SVM	0.8556	0.0117
	SVM + FFA	0.8840	7.5390
	HMM	0.4892	0.0260
	NBC	0.8499	0.088
23 %, 77 %	DT	0.7923	0.0342
	SVM	0.8928	0.0168
	SVM + FFA	0.9097	9.9629
	HMM	0.4893	0.0224
	NBC	0.8828	0.0090
30 %, 70 %	DT	0.7924	0.3400
	SVM	0.8894	0.0272
	SVM + FFA	0.9148	12.7505
	HMM	0.4855	0.0219
	NBC	0.9158	0.0112

Table 4.3: Performance Comparison of five ML algorithms for several iterations using different sizes of Training/Testing data.

Chapter 5

Predictive Modelling for Energy Harvesting

5.1 Introduction

The heterogeneous wireless devices and applications pose big challenges to operators in terms of energy consumption. Energy harvesting devices could be a potential source of energy. In particular, radio frequency (RF) energy harvesting is a technology that allows ambient RF signals to be collected by an antenna and converted into DC power using a rectifier [89]. Compared with other natural sources of energy (solar, vibration, wind and acoustic), the RF energy is more reliable as it does not depend on nature. Apart from its controllable and deterministic nature, RF signals can be used for simultaneous transmission of power and information [91] for far-field communication that makes it better than magnetic induction schemes utilized for near-field transmissions. Thus, the adoption of RF energy harvesting technology is plausible.

On the other hand, the amount of RF energy that could be harvested changes with time and frequency. For example, there are more mobile signals

during the day than during the night time in the commercial areas. There is also more RF power in the TV band than in the mobile band, due to the higher TV transmission power. Thus, it is very important for RF energy harvesters to choose the right operating time and frequency for attaining maximum energy as well as scheduling the transmission time and power for the efficient communications.

To achieve this, energy prediction is a key component because it equips the harvesting node with better knowledge of the energy availability in the future. Most current works assume a theoretical model for energy as discussed in Chapter 1. For example, a queuing model is proposed in [135] that analyses energy profile in solar powered wireless sensor network. This model would be useful for designing the energy efficiency protocols while ensuring QoS constraints. Similarly in [136], the performance of an energy harvesting node is analysed as a function of energy profile and outage probability. In [137], a Markov chain model is presented to illustrate energy renewal process for replensiable sensors. Also the optimal transmission policies for managing sensors with different energy budgets is presented. Similarly, in [138], a weather-conditioned moving average (WCMA) method was used to model energy in solar harvesting systems. In [139], exponentially weighted moving average (EWMA) method was proposed to improve WCMA. All these models can be utilized for the management of energy efficient networks. However, none of them used real time measurements from practice. After in depth literature review, it was concluded that there has been no work on the predictive modelling of the harvested RF energy using real data.

In this chapter, real-time power measurements taken at the University of Warwick from 880 -2500 MHz (as discussed in Chapter 4) are used for predicting the RF energy, that could be harvested. The main contributions of

this chapter can be summarized as follows:

- A machine learning based predictive framework is presented that evaluates a specific time of the day and a frequency channel, where more energy can be harvested. The amplitude probability distribution (APD) is also evaluated for the harvested power that estimates the percentage of attaining a specific amount of power in a frequency band.
- Two machine learning techniques: linear regression (LR) and decision trees (DT), are employed for predicting the harvested energy by considering low and medium-efficiency RF energy harvesters Numerical results show that LR outperforms DT by attaining minimum 85% prediction accuracy.
- The performance comparison of LR is also evaluated with a conventional prediction scheme in literature: moving average method. It was observed that LR outperformed moving average method as well.

The remainder of the chapter is organized as follows. Section 5.2 introduces the system model including the details of RF energy harvesters, machine learning framework and APD. Section 5.3 discusses the methodology that how to utilize linear regression and decision trees for predicting the harvested power. Numerical results are presented in Section 5.4. Finally, concluding remarks are given in Section 5.5.

5.2 System Model

5.2.1 RF Energy Harvesters

The key component for converting radio waves into electrical energy in an energy harvester is a rectenna, comprised of an antenna and a RF-DC converter.

The DC output power of RF energy harvester is typically in the micro watt to milli watt range, depending on a number of factors, such as the distance, power conversion efficiency (PCE) and the input power of the source [140]. The efficiency of a harvester greatly depends on the design of the rectenna. The PCE is the most critical parameter for a rectanna. It depends on the circuit topology, diode-device parameters and the input RF amplitude. Given PCE (η), the harvested power is calculated as:

$$H^{i,j} = \eta * P^{i,j}. \tag{5.1}$$

where $P^{i,j}$ represents the input power incident at the RF energy harvester in the ith time slot of the jth frequency bin and $H^{i,j}$ represents the harvested power. The input power $P^{i,j}$ represents the same data matrix $y^i(j)$ (as discussed in Section 4.2.1), where $i = 1, 2, .., n$ and $j = 1, 2, .., k$.

5.2.2 Machine learning framework

Supervised ML is considered in this chapter, where feature vectors are trained corresponding to their defined target labels to attain the trained model. The trained model is further tested using the unseen testing feature vector. The proposed approach in this work is a regression problem, where the target label $H^{i,m}$ (represents the harvested power of the m-th frequency bin in the i-th time slot) is predicted using the feature vector: $\mathbf{H}^{i,m-1:m-q} = [H^{i,m-1}, .., H^{i,m-q}]^T$ (represents the harvested power of q frequency bins). In this approach, m represents the frequency bin to be predicted and q represents the number of frequency bins used before m to estimate the power of the m-th frequency bin. The parameter q is a subset of k and it represents the length of training and testing feature vectors.

5.2.3 Amplitude Probability Distribution (APD)

The probability of occurrence of a specific amount of harvested power level is evaluated using amplitude probability distribution (APD) in our approach [141]. The APD is given as:

$$APD(H_{thres}, j) = P_r(H^{i,j} < H_{thres}) = \frac{1}{n}\sum_{i=1}^{n} F^{ij} \tag{5.2}$$

where H_{thres} represents the specific threshold for which the probability of occurrence is calculated and F^{ij} represents the status of the ith time slot in the jth frequency bin, given as as [141]

$$F^{ij} = \begin{cases} 1, & H^{i,j} > H_{thres} \\ 0, & H^{i,j} < H_{thres}. \end{cases}$$

Using APD, one can find out what amount of energy is most likely to be harvested.

5.3 Proposed Methodology

5.3.1 Linear Regression

The LR model for the training phase of the i-th time slot is given as:

$$\begin{aligned} H_{train}^{i,m} &= b_0 + b_{m-1}H_{train}^{i,m-1} + b_{m-2}H_{train}^{i,m-2} + ... + b_{m-q}H_{train}^{i,m-q} \\ &= b_0 + \sum_{j=m-1}^{m-q} b_j H_{train}^{i,j} \end{aligned} \tag{5.3}$$

Equation (5.3) can written as

$$H_{train}^{i,m} = b_0 + [b_{m-1}, b_{m-2}, .., b_{m-q}]^T \mathbf{H}_{train}^{i,m-1:m-q} \tag{5.4}$$

where $\mathbf{H}_{train}^{i,m-1:m-q}$ represents the i-th training feature vector, $H_{train}^{i,m}$ represents the target label and $b_{m-1}, b_{m-2}, .., b_{m-q}$ represents the trained model

parameters. The trained model is further tested to predict the target label of the unseen testing feature vector $\mathbf{H}_{test}^{i,m-1:m-q}$.

5.3.2 Decision Trees

Decision trees work on divide-and-conquer algorithm. It divides the data into subsets, builds a tree for each of them and then combines those subtrees into a single tree [142]. Data comes in the form

$$(\mathbf{H}_{train}^{i,m-1:m-q}, H_{train}^{i,m}) = (H_{train}^{i,m-1}, H_{train}^{i,m-2}, H_{train}^{i,m-3}, .., H_{train}^{i,m-q}, H_{train}^{i,m}). \quad (5.5)$$

The data space ($\mathbf{H}_{train}^{i,m-1:m-q}$) is partitioned into smaller regions (leaves) to make the interactions more manageable. The choice of partition is guided by a least squares error criterion . Binary trees consider two way splits at each tree node. The best split at each node t is the split s that maximizes

$$\triangle E(s,t) = E(t) - P_L E(t_L) - P_R E(t_R) \quad (5.6)$$

where $E(t)$ is the mean squared error at node t, P_L and P_R represents the proportions of the instances that fall to the left and right branch of the node t respectively, $E(t_L)$ and $E(t_R)$ are the errors of the left and right branches [143]. Once the DT model is trained, it is further tested to predict the target label of the unseen testing feature vector.

The prediction error of LR/ DT model for the i-th time slot is given as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(H_{test}^{i,m} - \widehat{H_{test}^{i,m}})^2}{N}} \quad (5.7)$$

where $\widehat{H_{test}^{i,m}}$ represents the predicted target label of the m-th frequency bin evaluated using LR/ DT trained model and $H_{test}^{i,m}$ represents the expected target label of the harvested power. The normalized RMSE (NRMSE) is evaluated

in this approach given as

$$NRMSE = \frac{RMSE}{max(H_{test}^{i,m}) - min(H_{test}^{i,m})}. \tag{5.8}$$

5.4 Numerical Results and Discussion

For evaluating the proposed prediction framework, $n = 100$ time instants of the Band 880-915 MHz are considered, where $n = 100$ represents the data of around two hours. The data normalization is performed for standardizing the range of all feature vectors using the relation $x' = \frac{x-\bar{x}}{\sigma}$, where x, x', $\bar{x}$ and σ represents the original value, normalized value, mean and standard deviation of the feature vector. Using $n = 100$ and $k = 192$, it was observed in Fig. 5.1 that the input power of the band 880-915 MHz varies between [0.2167, 6.2695] mW after normalization. It was also observed from Fig. 5.1 that average input power of the Band 880-915 MHz is $P^{i,j} = 1.0436$ mW. As the input power is represented as a function of time and frequency in Fig. 5.1, the peak power at specific frequency bin and time slot can also be extracted from the Fig.5.1. It was observed that peak input power for Band 880-915 MHz can be extracted using frequency bin $k = 912$ MHz and time slot $n = 60$.

The input power reaching harvester is dependant on many factors such as frequency, distance and source power. As discussed in Section 2.4.1, that harvested energy attained using isotropic RF transmitter with a source power of 1.78W for frequency 868 MHz at a distance of 25m produces 2.3 μW. However the harvested power decreases to 2μW, when distance is increased to 27m using the same source power and frequency. This shows that harvested power decreases when distance between RF energy harvester and source transmitter increases. Th Friis equation for free space loss [144] also endorses this observation, where the power density decreases at the rate of $\frac{1}{d^2}$ and d rep-

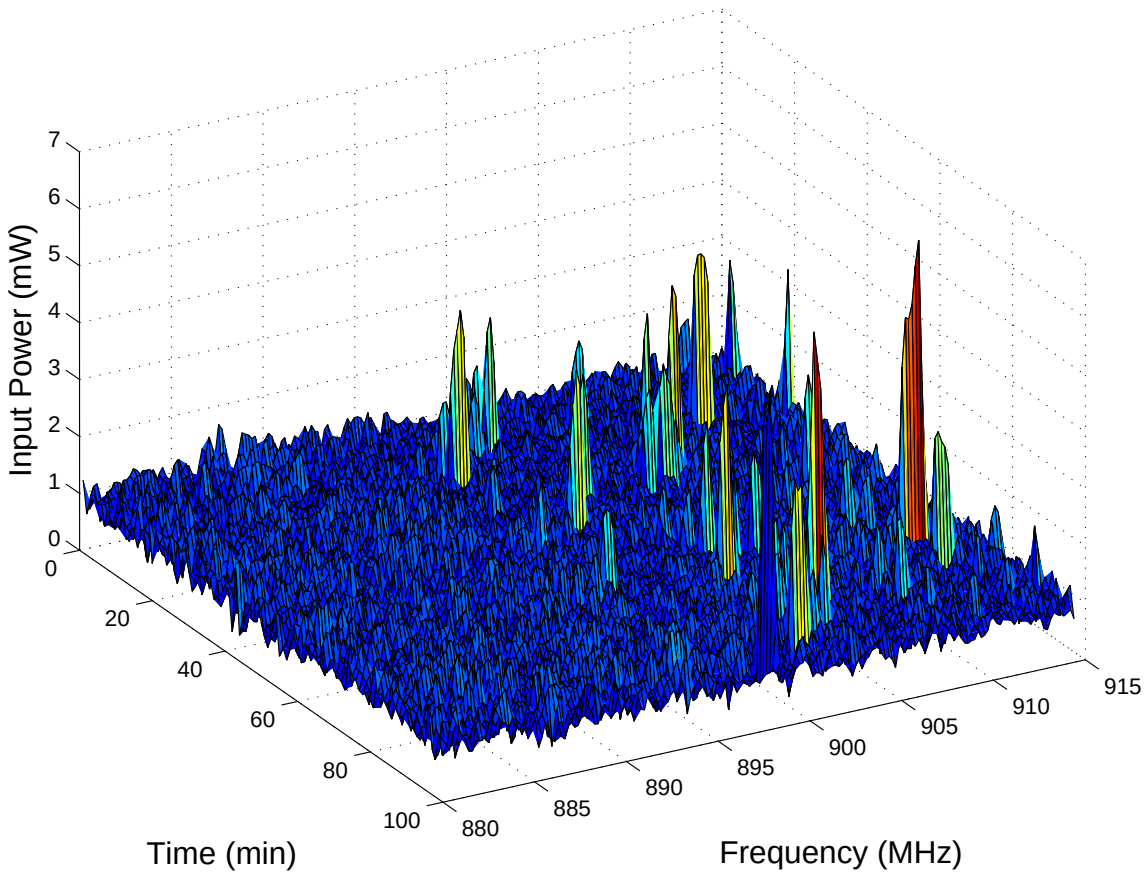


Figure 5.1: Input Power(mW) after normalization using $n = 100$ and $k = 192$ of Band 880-915 MHz.

resents the distance. Therefore, it can be concluded that harvested power is directly affected by the distance. The harvested power for the proposed model is explained below.

5.4.1 Harvested Power

For harvesting, energy, two types of harvesters are used in this approach as shown in Fig. 5.2, where low efficiency harvester (LEH) exhibits a maximum η of 29% at $P^{i,j} = 0.1$ mW [145], while medium efficiency harvester (MEH) produces a maximum η of 45% at $P^{i,j} = 0.1$ mW [146]. Both harvesters are

different because each one of them exhibits a different value of PCE for the same input power. As discussed in Section 6.2.1, PCE is the most critical parameter of an harvester and depends on the circuit topology and diode-device parameters. It was observed in Fig. 5.2 that the PCE for both harvestors decreases when RF input power increases. It was observed in Fig. 5.3(a), that maximum harvested power attained using LEH is $H^{i,j} = 0.13$ mW ($H^{i,j} = 2.3\% * 6.2695$, following (5.1)) and using MEH, it is $H^{i,j} = 0.69$ mW ($H^{i,j} = 11\% * 6.2695$) in Fig. 5.3(b). It was also observed from Fig. 5.3 that average harvested power using LEH and MEH is $H^{i,j} = 0.0774$ mW and $H^{i,j} = 0.3342$ mW respectively.

5.4.2 Amplitude Probability Distribution

The APD of the harvested power is evaluated by setting 13 different values of H_{thres} ranging between $H_{thres} = [0.01, 0.13]$ mW using LEH in Fig. 5.4(a) and $H_{thres} = [0.1, 0.7]$ mW using MEH in Fig. 5.4(b). It was observed in Fig. 5.4(a) that probability of harvesting power level greater than $H^{i,j} = 0.04$ is 1. This can be illustrated by analysing the power samples in Fig. 5.3(a), where it is evident that minimum harvested power attained using low efficiency harvester is $H^{i,j} = 0.04$ mW. It was also observed in Fig. 5.4(a) that probability reduces to 0.5, when $H_{thres} = 0.07$ mW is used for LEH. This is because one half of the total harvested signals lies below the average harvested power ($H^{i,j} = 0.07$ mW) and the other half exists above the average harvested power. The similar behaviour is observed for MEH in Fig. 5.4(b) that probability reduces to 0.5, when $H_{thres} = 0.33$ mW is used (since the average harvested power is $H^{i,j} = 0.33$ mW for MEH as mentioned in Section 5.4.1).

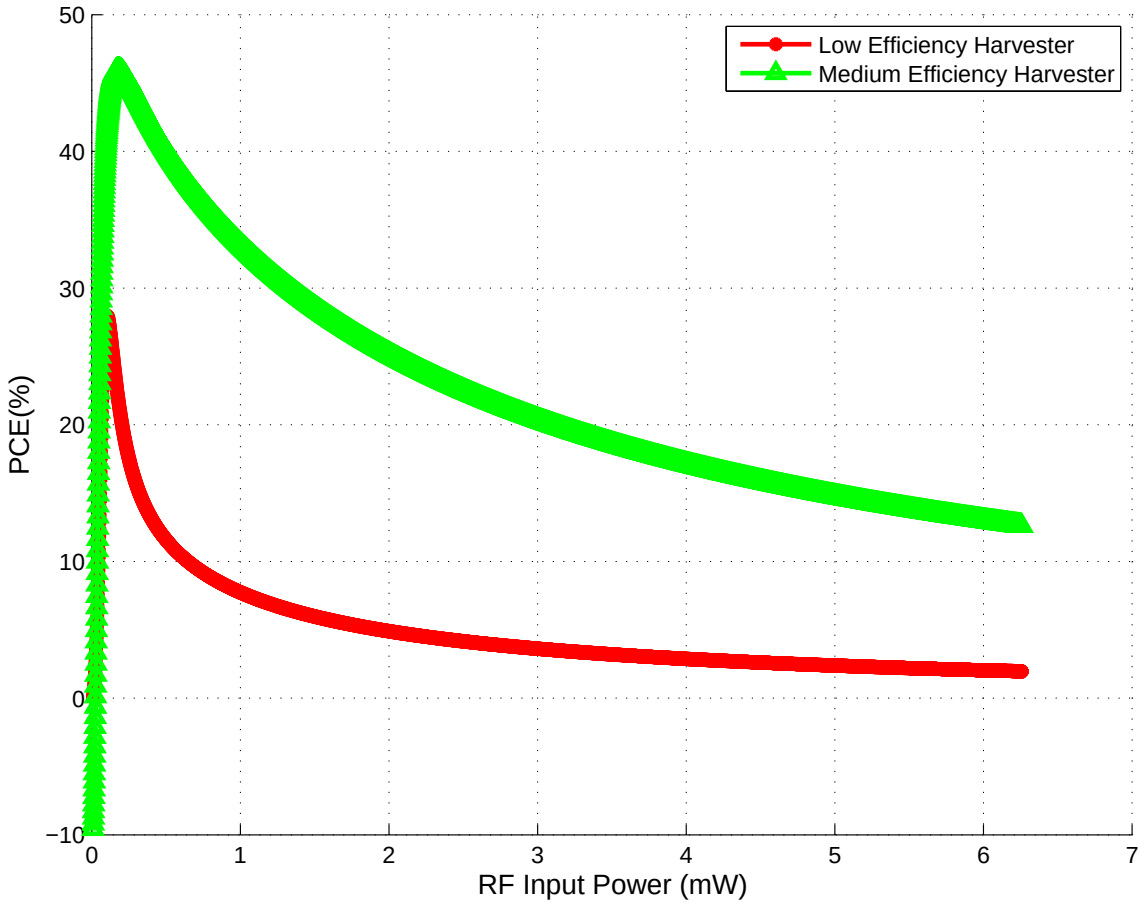
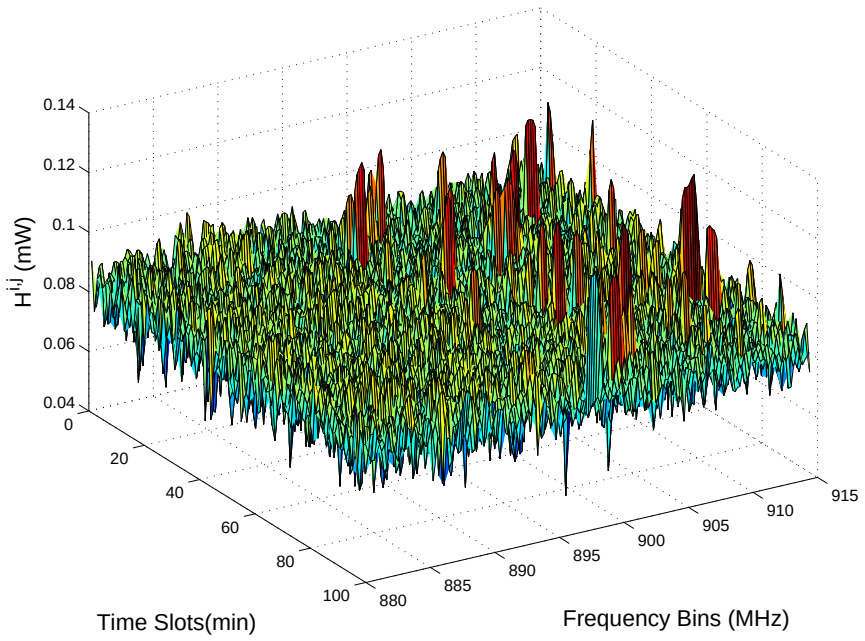
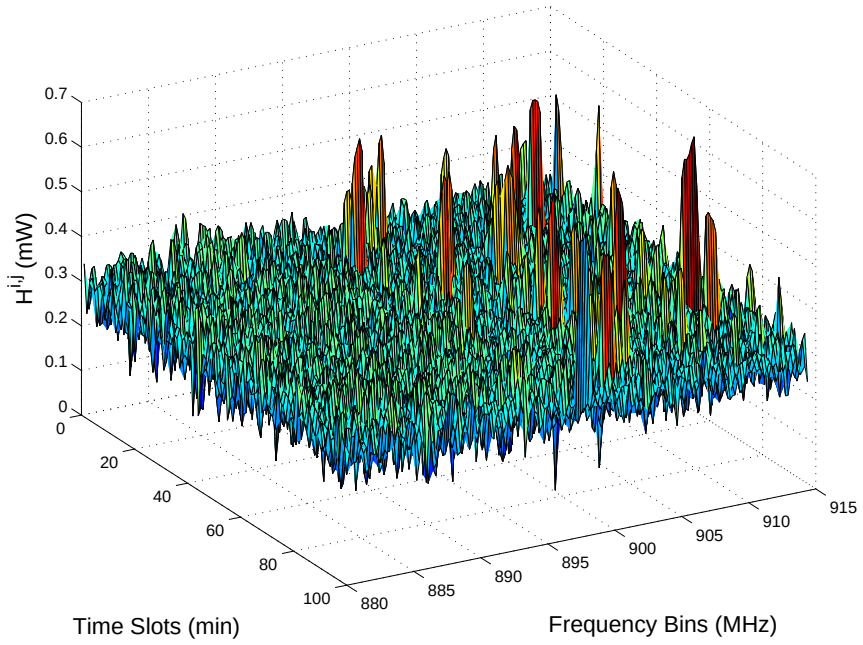


Figure 5.2: PCE plots of LEH [145] and MEH [146] for Band 880-915 MHz.

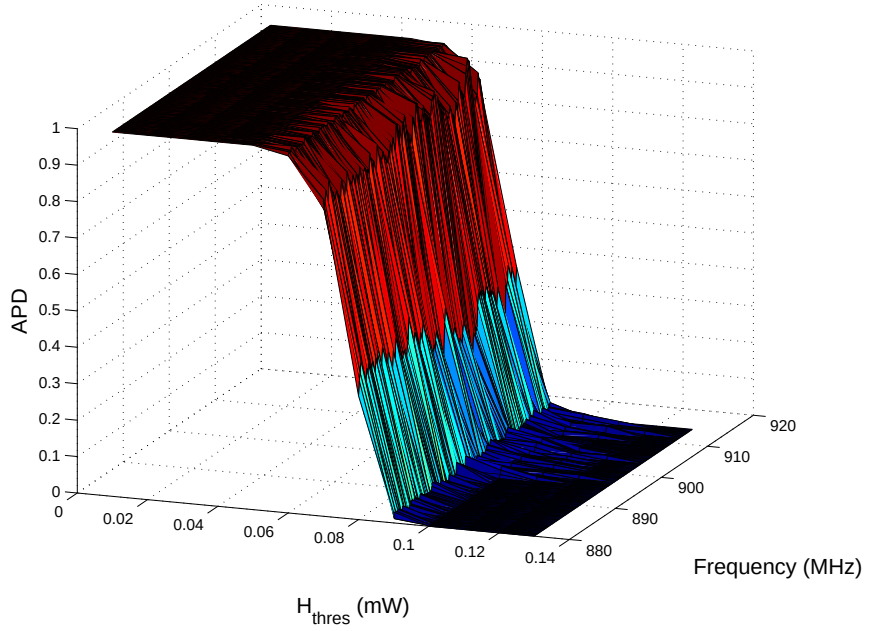


(a)

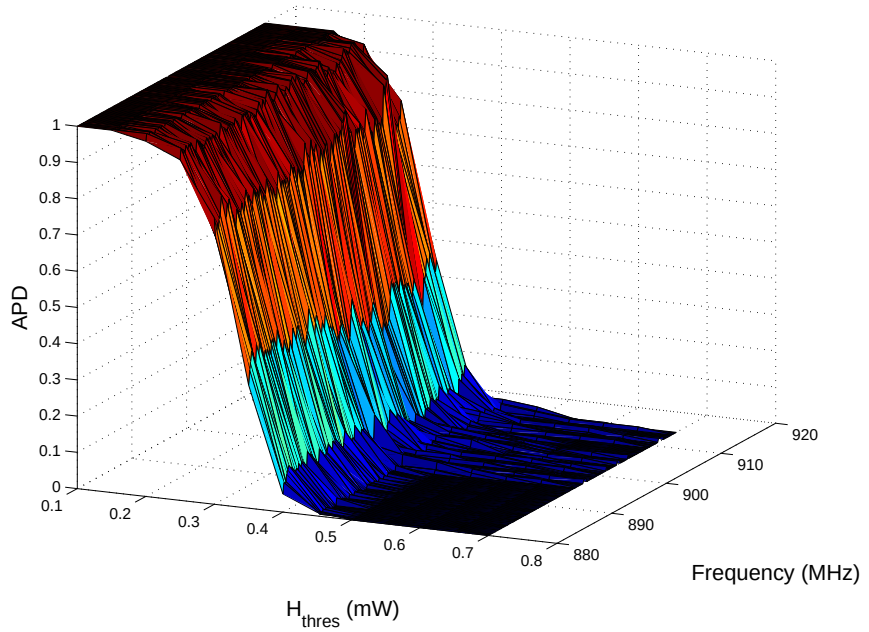


(b)

Figure 5.3: Harvested power of Band 880-915 MHz using $n = 100$ and $k = 192$ for (a) low efficiency harvester (b) medium efficiency harvester.



(a)



(b)

Figure 5.4: Amplitude probability distribution of Band 880-915 MHz for $n = 100$ and $k = 192$ using (a) LEH (b) MEH.

5.4.3 Predictive Energy model

The harvested power is predicted in Fig. 5.5 for frequency bins between 896-906 MHz using both LEH and MEH. It was observed that LR has outperformed DT by attaining a smaller value of NRMSE for both LEH and MEH. The maximum value of NRMSE attained by LR and DT using LEH is 0.15 and 0.2 respectively, which depicts that predictions are at least 85 % and 80% accurate using LR and DT, respectively. It was observed that harvested power ($\widehat{H_{test}^{i,m}}$) attained by LR and DT for LEH is 0.118 mW and 0.116 mw, respectively, while it is 0.6347 mW and 0.5655 mW, respectively, for the case of MEH. Hence, it can be concluded that $\widehat{H_{test}^{i,m}}$ attained using LR is larger than DT.

So far, the proposed framework is tested using data of Band 880-915 MHz, that belongs to Group B (periodic group), therefore in this section, the performance of the proposed predictive model is also evaluated using data of an aperiodic band that belongs to Group A. For that, $n = 100$ time slots of Band 2110-2170 MHz are used. It was observed in Fig. 5.6 that LR has outperformed DT again by attaining the maximum NRMSE of 0.23 using LEH and 0.2 using MEH respectively, however DT attained the maximum error of 0.24 for both LEH and MEH. It was observed that prediction error is increased in case of Band 2110-2170 MHz compared to the Band 880-915 MHz. Perhaps, this may have happened because 880-915 MHz is a periodic band, where the pattern of the input power repeats periodically and can be predicted more accurately. However Band 2110-2170 MHz is an aperiodic band, where input power fluctuates without any defined pattern.

It was also observed in Fig. 5.6 that $\widehat{H_{test}^{i,m}}$ attained using both LEH and MEH is higher than Band 880-915 MHz. This is because the harvesters used for Band 2110-2170 MHz have higher PCE compared to the harvesters utilized for Band 880-915 MHz in Fig. 5.2. The maximum harvested power attained

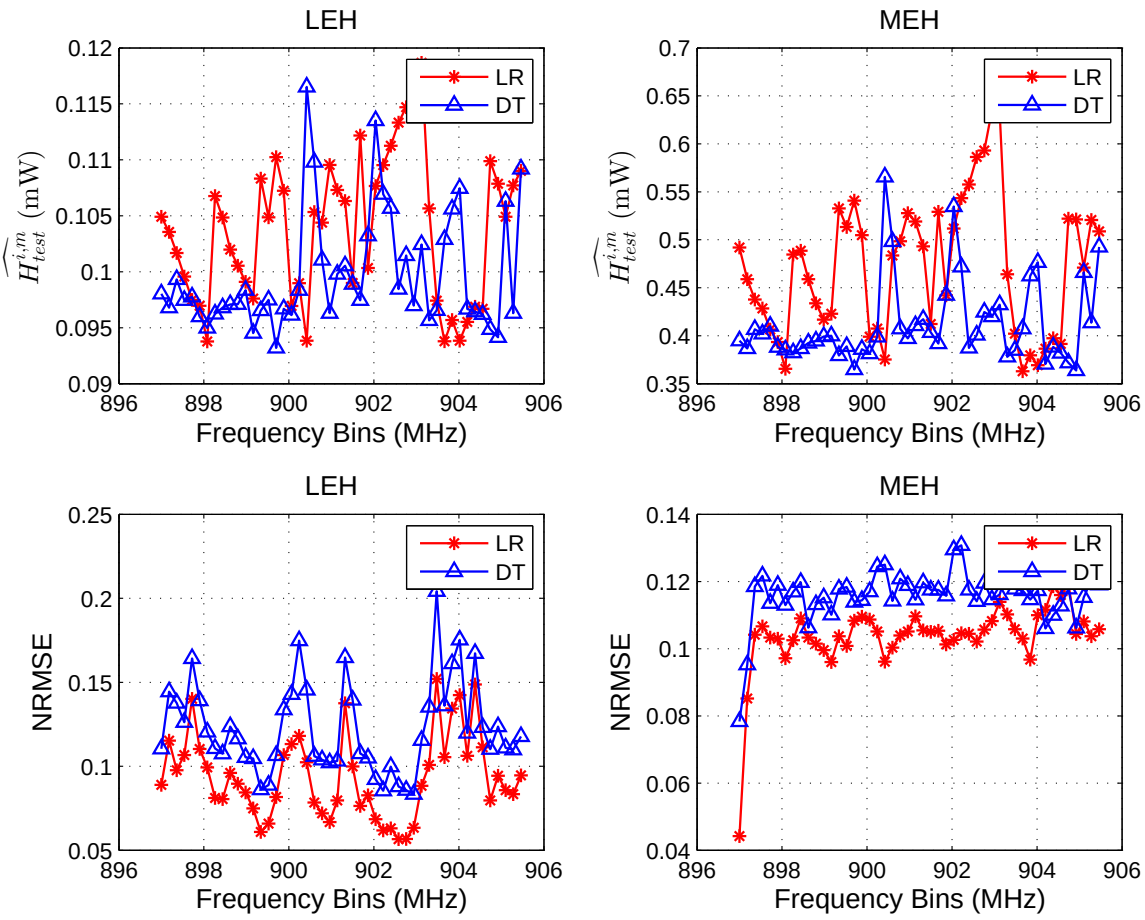


Figure 5.5: The predicted harvested power $\widehat{H^{i,m}_{test}}$ and RMSE using DT and LR for LEH and MEH between 896- 906 MHz.

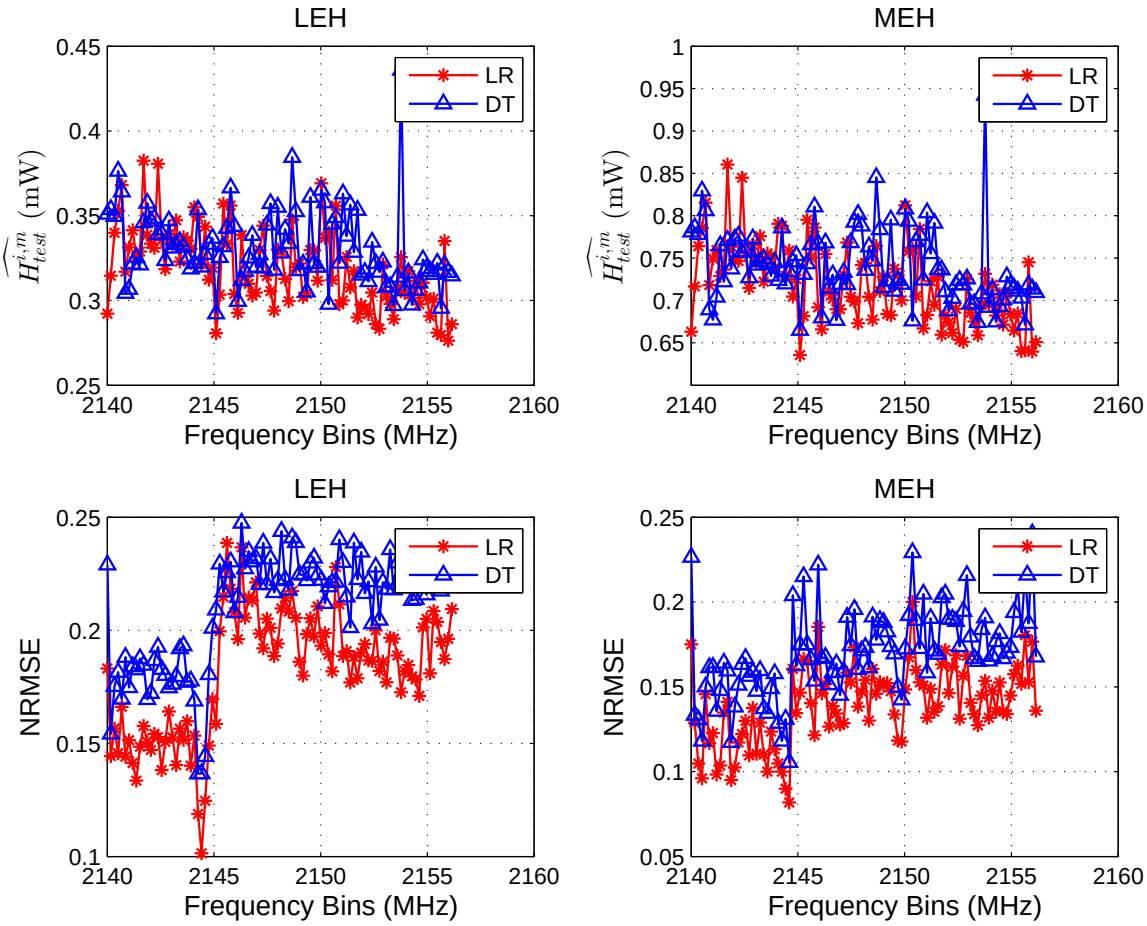


Figure 5.6: The predicted harvested power $\widehat{H_{test}^{i,m}}$ and RMSE using DT and LR for LEH and MEH between 2140- 2160 MHz

by DT for Band 2110-2170 MHz is 0.43 mW using LEH ($H^{i,j} = 22\% * 2.02$, where 2.02 mW is the maximum input power of Band 2110-2170 MHz and 22% is the η attained using [147]). Similarly the maximum harvested power attained by DT is 0.909 mW using MEH ($H^{i,j} = 45\% * 2.02$, where $\eta = 45\%$ is attained using [148]).

Each frequency bin in Fig. 5.5 and Fig. 5.6 contains n time slots, therefore the predicted harvested power in a specific frequency bin eventually means the prediction of the harvested power in n time slots of the selected

frequency bin. In Fig. 5.7(a), $\widehat{H_{test}^{i,m}}$ is evaluated for two frequency bins of Band 880-915 MHz: $m = 897$ MHz and $m = 897.2$ MHz using the period of $n = 100$ time slots. The predicted harvested power is compared with the actual harvested power and it was found that there is positive correlation between actual and predicted value for both LEH and MEH. The $\widehat{H_{test}^{i,m}}$ attained at $m = 897$ MHz and $m = 897.2$ MHz is 0.1049 mW and 0.1036 mW using LEH, while it is 0.4919 mW and 0.4590 mW using MEH, respectively. These values coincide with our results in Fig. 5.5 for $m = 897$ and $m = 897.2$ MHz. It was observed in Fig. 5.7 that $\widehat{H_{test}^{i,m}}$ shows a spike at time slot $i = 34$ and frequency bin $j = 897$ MHz in case of both LEH and MEH. This specific value of time slot and frequency bin predicted by the proposed model determines the specific point, where more power can be harvested. The proposed model can plays a vital role for defining the adaptive transmission policies according to the variation in prediction for the power variations.

.

5.4.4 Performance Comparison of LR and DT

The performance of LR and DT for two bands: 880-915 MHz and 2110-2170 MHz is compared using data of two days (6 -7 Feb, 2013), $n = 500$ time slots and different sizes of the training data set (q) in Table 5.1. It was observed that LR has outperformed DT in all cases by attaining smaller value of NRMSE, however the computation time required by LR in the training phase is higher than DT. The computational time is directly proportional to the size of training data set and it is inversely proportional to the value of NRMSE. It was also observed that values of NRMSE and computational time vary with days. This variation could be due to the change in environmental conditions at a specific time of the day.

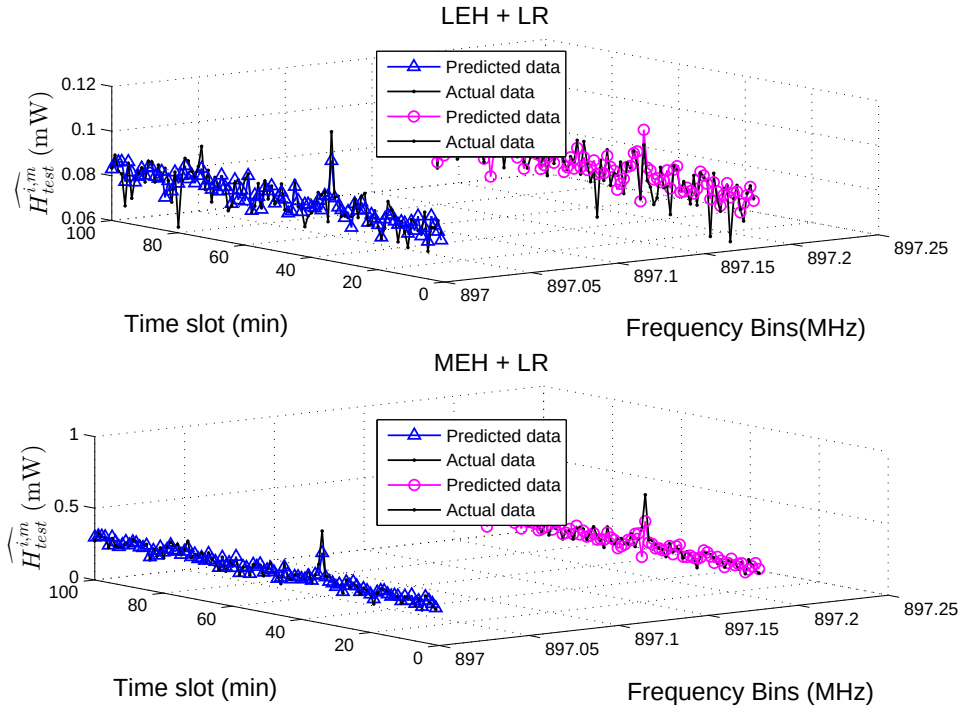


Figure 5.7: The predicted harvested power $\widehat{H_{test}^{i,m}}$ for $m = 897$ MHz and $m = 897.2$ MHz attained using LEH and MEH.

Band, Day	q	LR: NRMSE, Time(s)	DT: NRMSE, Time(s)
880-915, 6th Feb	35	0.1239, 0.8894	0.1606, 0.3830
	30	0.1412, 0.7022	0.1715, 0.4978
880-915, 7th Feb	35	0.1214, 0.9278	0.1446, 0.6538
	30	0.1388, 0.7364	0.1542, 0.4861
2110-2170, 6th Feb	35	0.1739, 0.8770	0.2171, 0.6396
	30	0.1800, 0.7202	0.2195, 0.5042
2110-2170, 7th Feb	35	0.1624, 0.9441	0.2187, 0.6498
	30	0.1766, 0.7535	0.2315, 0.5058

Table 5.1: Performance Comparison of LR and DT for different bands, days and training data sets

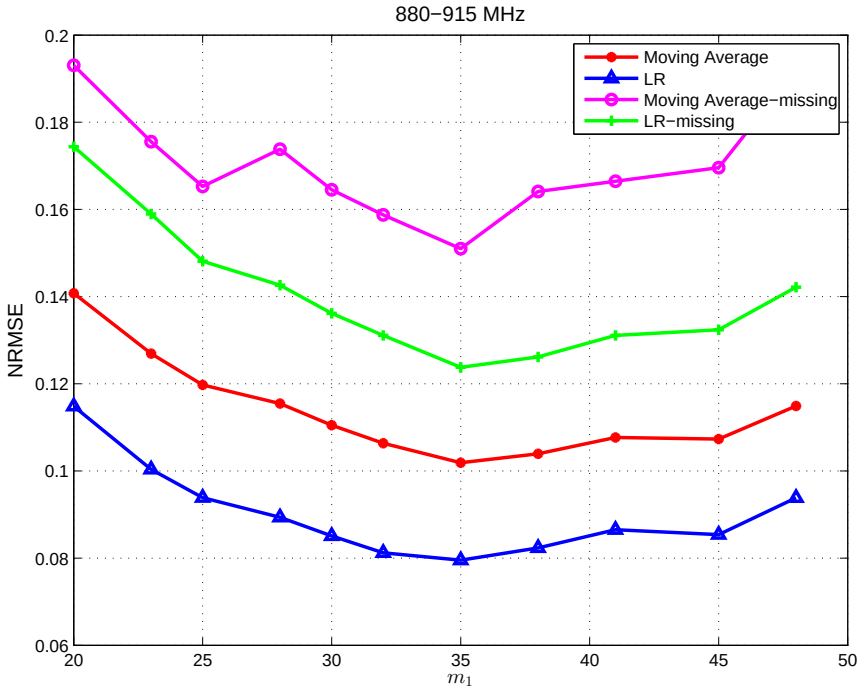
.

5.4.5 Performance Comparison of Moving Average Method and LR

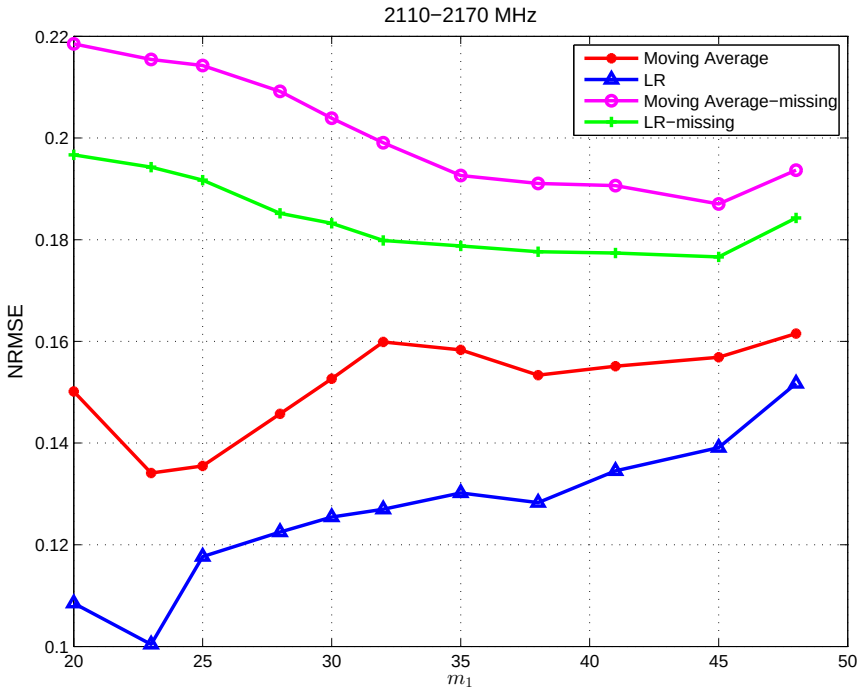
As LR outperforms DT in the considered cases, the performance of LR is compared in this section with a simple forecasting method in the literature i.e. moving average (MA). In MA, the best predicted value for future is evaluated by taking the average of everything that has happened up until now [149]. The observation (m_1) is predicted as the average of the last $m_1 - 1$ observations, where $m_1 - 1$ is the moving average interval. In the proposed approach, the parameter $m_1 - 1$ represents the number of frequency bins utilized for predicting the power of frequency bin m_1 in the i-th time slot given as:

$$\widehat{H^{i,m_1}} = \frac{\sum_{j=1}^{m_1-1} H^{i,j}}{m_1 - 1} \tag{5.9}$$

The parameter $m_1 - 1$ in MA is interpreted as q in LR model. The performance of LR and MA for two frequency bands: 880-915 MHz and 2110-2170 MHz is compared as shown in Fig. 5.8. It was observed that LR has outperformed MA by attaining smaller value of NRMSE for 11 different values of m_1. The stability of the proposed prediction framework is also validated by comparing LR with MA, for the case when missing entries are introduced in the testing data. An amount of 5% missing entries are introduced in the testing data set and fed into the trained model. The predicted values from the LR trained model are further compared with the predicted values evaluated using MA method. It was observed that LR has outperformed MA again by attaining smaller value of NRMSE, when missing entries are introduced. It was also observed that NRMSE is increased by $8 - 10\%$ using LR and MA schemes, when compared with the case, when there are no missing entries in the testing data set. This verifies the observation that error increases with an introduction of noise in the system.



(a)



(b)

Figure 5.8: Comparison of LR with MA using (a) 880-915 MHz (b) 2110-2170 MHz.

5.5 Conclusion

In this chapter the prediction framework is proposed for evaluating the harvested power in frequency channels using two prediction schemes: LR and DT. The results are verified using measurements of different frequency bands for different days. It was observed that LR outperformed DT by attaining smaller value of NRMSE. The minimum prediction accuracy determined by the LR model is 85%. The proposed methodology plays a vital role for defining the adaptive transmission policies because one can find the specific time and frequency, where harvester can generate more power.

Chapter 6

Energy Harvesting Cognitive Radios

6.1 Introduction

Much recent research has been focussed on exploiting the energy harvesting (EH) in cognitive radio networks (CRN) [150]. In [151], the energy harvesting CRN was proposed, where SU harvests energy from the nearby PU or transmits information if the PU is far away. In [152], the achievable throughput of the secondary transmitter that harvests energy while opportunistically accessing the PU licensed spectrum was analysed for improving the energy and spectral efficiency of the CRN. Similarly, [153] focused on determining the optimal spectrum sensing policy of SU, subject to an energy causality constraint and a collision constraint. There are also a few studies which suggested cooperation between PU and SU. For example, in [154], SU was proposed as a relay for PU communication, where SU harvests energy first from the received primary signals and then uses the harvested energy to forward the primary signals. Similarly, in [155, 156], SU cooperated with PU to boost the primary system's

performance.

All these works assume that PU does not have any energy harvesting capability. However, if the PU is a wireless powered system [157], such assumption is not valid any more. As discussed in Section 2.4.2, WPC is referred as the energy harvesting communication. In this chapter, the performance of the WPC enabled CRN is analysed, where the PU is a wireless powered system with energy harvesting capability. Specifically, consider an access point (AP) that transfers energy to all PUs in the downlink (DL) and while PUs harvest energy from the AP, SUs also transmit data to increase its throughput. In the proposed configuration, PU gets chance to harvest energy both from the AP and SU data transmissions. Therefore considering the proposed methodology, PU can take advantage from the cooperation with the SUs without predetermined agreement. The contribution of this chapter can be summarized as follows:

- A wireless powered energy harvesting cognitive radio network is presented in this thesis, where PU gets benefits due to existence of SU without allocating extra resources. The idea of using WPC in energy harvesting cognitive radio networks has not been presented before.

- A three state hypotheses for distinguishing PU statuses (absent, harvesting energy and transmitting data) in WPC enabled CRN is presented. Two decision thresholds have been utilized to differentiate three PU statuses.

- The proposed framework is also compared with the conventional energy harvesting CRN in the literature, where SU is only allowed to harvest energy, when PU is transmitting. Numerical results have shown that proposed strategy outperforms the conventional strategy, where PUs can

attain energy gain from the transmission of SU and SU can attain the throughput gain from the extra transmission time allocated to PUs for energy harvesting.

The rest of the chapter is organized as follows. Section 6.2 introduces the system model that includes the details of the three state hypotheses in wireless powered communication network. Section 6.3 discusses the proposed and the conventional strategies in the literature. Numerical results are presented in Section 6.4. Finally concluding remarks are given in Section 6.5.

6.2 System Model

Consider a wireless powered primary network with one access point (AP) and K PUs, denoted by PU_i, $i = 1, 2, .., K$, as shown in Fig. 6.1. All PUs harvest energy from the AP in the downlink (DL) during time duration τ_0, and use the harvested energy to transmit the information to the AP in the uplink (UL) during τ_i [158], where τ_i represents the time duration allocated to the i^{th} PU for the UL transmission and $\tau_0 + \sum_{i=1}^{K} \tau_i = 1$. It is assumed that all nodes operate in half-duplex. A secondary network is also considered in the same area, having one secondary user transmitter (SUTR) and one secondary user receiver (SUR). The proposed framework works as follows:
(a) When the AP transfers energy to the PUs during τ_0, SUTR transmits data to the SUR. The main rationale behind this method is that AP transmits wireless energy to all PUs in the network during a time slot τ_0 and in order to create equal opportunity for SU to use the network, it has been proposed that SUTR sends data to the SUR during the same time slot τ_0. Due to simultaneous working of both SUTR and AP in τ_0, PU faces interference from SUTR transmission which is further perceived as an advantage by PU because

it can harvest energy from SUTR interference. In this manner, SU and PU both can get equal benefit during time slot τ_0 because SUTR is able send data to SUR that eventually increases throughput of SUTR and PU can harvest energy. Furthermore, it has been assumed that PUs posses less energy than AP because AP is serving as the central controller that provides wireless energy to all PU terminals. that is further used by PUs for sending data back to AP.
(b) When the PUs are detected to send data to the AP during $\tau_1, \tau_2, ..., \tau_K$, then SUTR tunes itself for energy harvesting. In this step, SUTR is not allowed to send data to SUR. This is because maximum protection can be guaranteed to PUs by minimizing interference due to SUTR transmissions.
(c) Also, define τ_{free} as the time during which neither the AP is transferring energy nor the PUs are sending data. The SUTR uses τ_{free} for the data transmission, when both the PUs and the AP are detected absent.

Assume that the primary network has different transmission powers for energy and data transfer, where higher transmission power is assumed for the data transfer than for the energy transfer. Define two decision threshold λ_1 and λ_2, to differentiate the information transmission from the energy transfer, where $\lambda_2 > \lambda_1$. Let $r(t)$ be the sample sensed by SUTR at time t. Using energy detection, the received signal energy at SUTR is measured as: $\sum_{t=0}^{M-1}|r(t)|^2 = T(r)$, where M represents the total number of samples. The PUs are either free or harvest energy or transmit information, denoted as H_0, H_1 and H_2, respectively. Then, the detection by SUTR is performed as

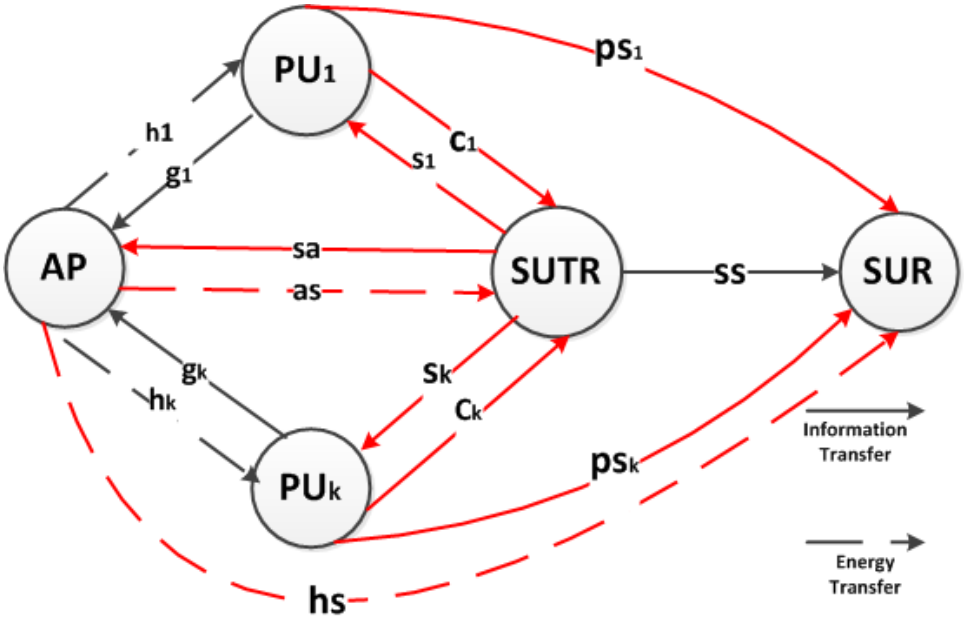


Figure 6.1: A wireless powered cognitive radio framework having K primary users (PU_i, where $i = 1, 2, .., K$), an access point (AP), a secondary user transmitter (SUTR) and a secondary user receiver (SUR). The links in black color represent the communication channels within a primary/ secondary network while the ones in red color represent the communication channels between the primary and the secondary network. The primary users $PU_{i=1}$ and $PU_{i=K}$ are represented in the figure, where i can have any value between 1 and K.

$$\begin{cases} T(r) < \lambda_1 \implies PUs\ are\ detected\ free : H_0 \\ \lambda_1 \leqslant T(r) \leqslant \lambda_2 \implies \\ PUs are\ detected\ to\ harvest\ energy : H_1 \\ T(r) > \lambda_2 \implies \\ PUs\ are\ detected\ to\ transmit\ information : H_2 \end{cases} \tag{6.1}$$

Following the two state hypotheses in [159], the probability of detection and probability of false alarms for three state hypotheses using (6.1) has been derived as

Detection and false alarm probabilities in state H_0

$$P(H_2|H_0) = P(T(r) > \lambda_2|H_0) = \frac{\Gamma(u, \frac{\lambda_2}{2\sigma^2})}{\Gamma(u)} \tag{6.2}$$

$$\begin{aligned} P(H_1|H_0) &= P(\lambda_1 <= T(y) <= \lambda_2|H_0) \\ &= P(T(y) <= \lambda_2|H_0) - P(T(y) <= \lambda_1|H_0) \\ &= (1 - \frac{\Gamma(u, \frac{\lambda_2}{2\sigma^2})}{\Gamma(u)}) - (1 - \frac{\Gamma(u, \frac{\lambda_1}{2\sigma^2})}{\Gamma(u)}) \\ &= \frac{\Gamma(u, \frac{\lambda_1}{2\sigma^2})}{\Gamma(u)} - \frac{\Gamma(u, \frac{\lambda_2}{2\sigma^2})}{\Gamma(u)} \end{aligned} \tag{6.3}$$

$$P(H_0|H_0) = 1 - P(H_1|H_0) - P(H_2|H_0) \tag{6.4}$$

Detection and false alarm probabilities in state H_1

$$P(H_2|H_1) = P(T(y) > \lambda_2|H_1) = Q_u(\sqrt{2\gamma_1}, \sqrt{\lambda_2}) \tag{6.5}$$

$$
\begin{aligned}
P(H_1|H_1) &= P(\lambda_1 <= T(y) <= \lambda_2|H_1) \\
&= P(T(y) < \lambda_2|H_1) - P(T(y) <= \lambda_1|H_1) \\
&= Q_u(\sqrt{2\gamma_1}, \sqrt{\lambda_1}) - Q_u(\sqrt{2\gamma_1}, \sqrt{\lambda_2})
\end{aligned} \tag{6.6}
$$

$$
P(H_0|H_1) = 1 - P(H_1|H_1) - P(H_2|H_1) \tag{6.7}
$$

Detection and false alarm probabilities in state H_2

$$
P(H_2|H_2) = P(T(y) > \lambda_2|H_2) = Q_u(\sqrt{2\gamma_2}, \sqrt{\lambda_2}) \tag{6.8}
$$

$$
\begin{aligned}
P(H_1|H_2) &= P(\lambda_1 <= T(y) <= \lambda_2) \\
&= P(T(y) <= \lambda_2|H_2) - P(T(y) <= \lambda_1|H_2) \\
&= Q_u(\sqrt{2\gamma_2}, \sqrt{\lambda_1}) - Q_u(\sqrt{2\gamma_2}, \sqrt{\lambda_2})
\end{aligned} \tag{6.9}
$$

$$
P(H_0|H_2) = 1 - P(H_2|H_2) - P(H_1|H_2) \tag{6.10}
$$

where $\Gamma(.,.)$, $Q_u(a,b)$, u, γ_1 and γ_2 represents the incomplete Gamma function, the generalized Marcum Q-function, the time-bandwidth product [159], the PU signal to noise ratio (SNR) in H_1 and the PU SNR in H_2, respectively.

6.3 Analysis

In this section, the proposed methodology for energy harvesting in wireless powered CRN is presented. Also the conventional strategy for energy harvesting CRNs in literature is discussed.

6.3.1 Proposed Strategy

The performance of the cognitive radios with wireless powered PUs is analysed in this section, where the SUTR transmits data when PUs are detected to harvest energy. In this case, SUTR transmits data when H_0 and H_1 are detected and harvests energy when H_2 is detected. This allows PUs to achieve energy gain from the SUTR transmission too. Define $H_q|H_j$ as one of the possible conditions following the detection criteria in (6.1), where H_q represents the channel status detected by the SUTR and H_j represents the actual channel status and $q, j = \{0, 1, 2\}$. The received signal at the i^{th} primary user (PU_i) in the DL is given as

$$y_{PU_i} = \begin{cases} (\tilde{h_i}\sqrt{\frac{P_o}{L_{h_i}}})x_a + (\tilde{s_i}\sqrt{\frac{P_s}{L_{s_i}}})x_s + z_i, & H_0|H_1 \\ (\tilde{h_i}\sqrt{\frac{P_o}{L_{h_i}}})x_a + (\tilde{s_i}\sqrt{\frac{P_s}{L_{s_i}}}x_s) + z_i, & H_1|H_1 \\ (\tilde{h_i}\sqrt{\frac{P_o}{L_{h_i}}})x_a + z_i, & H_2|H_1 \end{cases} \tag{6.11}$$

where y_{PU_i} and z_i represent the received signal and the noise at PU_i, respectively, P_0 and P_s represents the transmitted power of the AP and the SUTR, respectively, $\tilde{h_i}$ and $\tilde{s_i}$ are complex random variables between $AP \rightarrow PU_i$ and $SUTR \rightarrow PU_i$ with channel power gains $h_i = |\tilde{h_i}|^2$ and $s_i = |\tilde{s_i}|^2$ respectively. The same definition applies to all other channel power gains in the model. Also in (6.11), L_{h_i} and L_{s_i} represents the path loss between $AP \rightarrow PU_i$ and $SUTR \rightarrow PU_i$ respectively, x_a and x_s represents the baseband signal transmitted by AP and SUTR, respectively. It is assumed that x_a is a complex random signal, satisfying $\mathbf{E}[|x_a|^2] = 1$, and the same applies to all other baseband signals considered in the model. The condition $H_0|H_1$ in (6.11) represents the case, when the SUTR detects the channel status as H_0 and sends data to the SUR, while the actual channel status is H_1. The same definition applies to all other conditions considered in the model. Using (6.11), the expected harvested

energy at PU_i can be derived as

$$\mathbf{E}[PE] = \eta_i \tau_o (\frac{h_i P_o}{L_{h_i}} + \frac{s_i P_s}{L_{s_i}})(1 - P(H_2|H_1)) + \eta_i \tau_o (\frac{h_i P_o}{L_{h_i}}) P(H_2|H_1) \quad (6.12)$$

where η_i represents the energy harvesting efficiency at PU_i and $0 < \eta_i < 1$. Once PU_i has harvested the energy in the DL, it will transmit the information to the AP in the UL during time slot τ_i. If the PUs are detected to send data to the AP, the SUTR can harvest energy from the PU transmission. This means that the SUTR harvests energy when H_2 is detected. Thus, the received signal at the SUTR for energy harvesting is given by

$$y_{SUTR} = \begin{cases} z_{TR}, & H_2|H_0 \\ (\tilde{as}\sqrt{\frac{P_o}{L_{as}}}x_a) + z_{TR}, & H_2|H_1 \\ (\tilde{c_i}\sqrt{\frac{P_i}{L_{c_i}}}x_{pi}) + z_{TR}, & H_2|H_2 \end{cases} \quad (6.13)$$

where z_{TR} represents the noise at SUTR, x_{pi} and P_i represents the baseband signal and the power transmitted by PU_i respectively, $\tilde{as}$ and $\tilde{c_i}$ are complex random variables between $AP \rightarrow SUTR$ and $PU_i \rightarrow SUTR$ respectively with channel power gains $as = |\tilde{as}|^2$ and $c_i = |\tilde{c_i}|^2$ respectively, L_{as} and L_{c_i} represents the path loss between $AP \rightarrow SUTR$ and $PU_i \rightarrow SUTR$ respectively. Using (6.12), the expected value of P_i is given by, $\mathbf{E}[P_i] = \frac{\zeta_i \mathbf{E}[PE]}{\tau_i}$, where ζ represents a fixed portion of the harvested energy utilized for the UL information transmission by each PU_i and $\zeta \in (0, 1)$. The expected harvested energy at the SUTR is given as

$$\mathbf{E}[SE] = \eta_s \tau_0 \frac{as P_0}{L_{as}} P(H_2|H_1) + \eta_s \tau_i \frac{c_i P_i}{L_{c_i}} P(H_2|H_2) \quad (6.14)$$

where η_s represents the energy harvesting efficiency of the SUTR. On the other hand, the SUR receives data from the SUTR, when H_0 or H_1 are detected. Thus, the received signal at SUR (y_{SUR}), is given as

$$y_{SUR} = \begin{cases} (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + z_s, \;\; H_0|H_0 \\ (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + z_s, \;\; H_1|H_0 \\ (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + (\tilde{hs}\sqrt{\frac{P_o}{L_{hs}}})x_a + z_s \;\; H_0|H_1 \\ (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + (\tilde{hs}\sqrt{\frac{P_o}{L_{hs}}})x_a + z_s, \;\; H_1|H_1 \\ (\tilde{ps}_i\sqrt{\frac{P_i}{L_{ps_i}}})x_{pi} + (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + z_s, \;\; H_0|H_2 \\ (\tilde{ps}_i\sqrt{\frac{P_i}{L_{ps_i}}})x_{p_i} + (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}}x_s) + z_s, \;\; H_1|H_2 \end{cases} \tag{6.15}$$

where z_s represents the noise at the SUR with variance σ^2, ss, hs and ps_i represents the channel gains between $SUTR \rightarrow SUR$, $AP \rightarrow SUR$ and $PU_i \rightarrow SUR$ respectively, L_{ss}, L_{hs} and L_{ps_i} represents the path loss between $SUTR \rightarrow SUR$, $AP \rightarrow SUR$ and $PU_i \rightarrow SUR$ respectively. The total expected throughput of SUR, can be derived as:

$$\begin{aligned} \mathbf{E}[R_{SUR}] = \tau_{free} log_2(1 + \frac{\frac{ssP_s}{L_{ss}}}{\sigma^2})(1 - P(H_2|H_0)) + \tau_0 log_2(1 + \frac{\frac{ssP_s}{L_{ss}}}{\frac{hsP_0}{L_{hs}} + \sigma^2})(1 \\ - P(H_2|H_1)) + \tau_i log_2(1 + \frac{\frac{ssP_s}{L_{ss}}}{\frac{ps_iP_i}{L_{ps_i}} + \sigma^2})(1 - P(H_2|H_2)) \end{aligned} \tag{6.16}$$

6.3.2 Conventional Strategy

As discussed in Section 2.4.1, the conventional RF energy harvesting CRN works in a manner that SUs not only identify spectrum spaces for the data transmission but also searches the occupied spectrum to harvest maximum energy. The conventional RF energy harvesting CRN works using two state hypothesis, where PU is either present (H_1) or its absent (H_0). Therefore SU gets chance to harvest energy only when PU is transmitting i.e. during H_1 state. However, the proposed approach works using three state hypothesis, where PU is present in two states: H_1 and H_2 while it is absent in one state

i.e. H_0. Therefore, following three state hypothesis, SU gets chance to harvest energy from AP in both H_1 and H_2 states.

The conventional energy harvesting cognitive radio framework for the proposed approach works in a manner that SUTR does not transmit any data as along as PU activity is detected, or H_1 and H_2 are detected. This means that PUs harvest energy when H_1 is true and transmits data when H_2 is true. The SUTR harvests energy when either H_1 or H_2 is detected, while the SUR only receives data from the SUTR, when H_0 is detected. Following a similar analysis, the received signal at PU_i during DL for energy harvesting is given as

$$y_{PU_i} = \begin{cases} (\tilde{h_i}\sqrt{\frac{P_0}{L_{h_i}}})x_a + (\tilde{s_i}\sqrt{\frac{P_s}{L_{s_i}}})x_s + z_i, & H_0|H_1 \\ (\tilde{h_i}\sqrt{\frac{P_o}{L_{h_i}}})x_a + z_i, & H_1|H_1 \\ (\tilde{h_i}\sqrt{\frac{P_o}{L_{h_i}}})x_a + z_i, & H_2|H_1 \end{cases} \tag{6.17}$$

Using (6.17), the energy harvested at PU_i for conventional strategy is given as

$$\mathbf{E}[PE] = \eta_i\tau_0(\frac{h_iP_0}{L_{h_i}} + \frac{s_iP_s}{L_{s_i}})P(H_0|H_1) + \eta_i\tau_o(\frac{h_iP_0}{L_{h_i}})(1 - P(H_0|H_1)) \tag{6.18}$$

In this case, SUTR does not transmit in H_1 and H_2. It harvests energy in both H_1 and H_2, therefore the received signal at SUTR, is given as

$$y_{SUTR} = \begin{cases} z_{TR}, & H_2|H_0 \\ (\tilde{as}\sqrt{\frac{P_0}{L_{as}}})x_a + z_{TR}, & H_2|H_1 \\ (\tilde{c_i}\sqrt{\frac{P_i}{L_{c_i}}})x_{p_i} + z_{TR}, & H_2|H_2 \\ z_{TR}, & H_1|H_0 \\ (\tilde{as}\sqrt{\frac{P_0}{L_{as}}})x_a + z_{TR}, & H_1|H_1 \\ (\tilde{c_i}\sqrt{\frac{P_i}{L_{c_i}}})x_{p_i} + z_{TR}, & H_1|H_2 \end{cases} \tag{6.19}$$

The energy harvested by SUTR in H_2 is given as

$$\begin{aligned}\mathbf{E}[SE] = \eta_s\tau_0\frac{asP_0}{L_{as}}P(H_2|H_1) + \eta_s\tau_i\frac{c_iP_i}{L_{c_i}}P(H_2|H_2) \\ + \eta_s\tau_o\frac{asP_o}{L_{as}}P(H_1|H_1) + \eta_s\tau_i\frac{c_iP_i}{L_{c_i}}P(H_1|H_2)\end{aligned} \tag{6.20}$$

which can be written as

$$\mathbf{E}[SE] = \eta_s\tau_o\frac{asP_0}{L_{as}}(1 - P(H_0|H_1)) + \eta_s\tau_i\frac{c_iP_i}{L_{c_i}}(1 - P(H_0|H_2)) \tag{6.21}$$

In this case, SUR receives the data from SUTR in H_0 only. The received signal at SUR is given as

$$y_{SUR} = \begin{cases}(\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + z_s, \;\; H_0|H_0 \\ (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + (\tilde{hs}\sqrt{\frac{P_o}{L_{hs}}})x_a + z_s, \;\; H_0|H_1 \\ (\tilde{ps}_i\sqrt{\frac{P_i}{L_{ps_i}}})x_{p_i} + (\tilde{ss}\sqrt{\frac{P_s}{L_{ss}}})x_s + z_s; \;\; H_0|H_2\end{cases} \tag{6.22}$$

Using (6.22), the expected throughput at SUR for the conventional strategy is given as

$$\begin{aligned}\mathbf{E}[R_{SUR}] = \tau_{free}log_2(1 + \frac{\frac{ssP_s}{L_{ss}}}{\sigma^2})P(H_0|H_0) + \tau_0log_2(1 \\ + \frac{\frac{ssP_s}{L_{ss}}}{\frac{hsP_o}{L_{hs}} + \sigma^2})P(H_0|H_1) + \tau_ilog_2(1 + \frac{\frac{ssP_s}{L_{ss}}}{\frac{ps_iP_i}{L_{ps_i}} + \sigma^2})P(H_0|H_2)\end{aligned} \tag{6.23}$$

6.4 Numerical Results and Discussion

In this section, the amount of energy harvested by PU_i, the throughout gained by the SUR and the energy harvested by the SUTR are compared for cognitive radios using the new and conventional strategies. It is assumed that $P_s = 5$ dB, $\eta_i = \eta_s = 0.4$. Also, $h_i = g_i = ss = 3$ dB, $s_i = c_i = 2$ dB, $as = sa = 1.4$ dB, $ps_i = 1.3$ dB and $hs = 1.2$ dB are set. As the distances between $PU_i \rightarrow SUR$, $AP \rightarrow SUR$, and $AP \rightarrow SUTR$ is comparatively larger than other nodes in

Fig. (6.1), therefore L_{as}, L_{sa}, L_{ps_i} and L_{hs} are assumed to be larger than other links. It is assumed that $L_{hi} = L_{g_i} = L_{ss} = -1.023$ dB, $L_{si} = L_{c_i} = -1.5$ dB, $L_{as} = L_{sa} = 0.484$ dB, $L_{ps_i} = 0.773$ dB and $L_{hs} = 0.569$ dB. Further, $\gamma_1 = 1.12$ dB, $\gamma_2 = 1.17$ dB are set, where it is assumed that SNR in H_2 is higher than that in H_1 for both strategies. Similarly λ_2 is assumed to be larger than λ_1, therefore $\lambda_1 = 5$ and $\lambda_2 = 8$.

6.4.1 Effect of changing P_0

Fig. 6.2 shows the relationship between $\mathbf{E}[PE]$ and P_0 for both new and the conventional strategies. Two scenarios are considered for both strategies: low channel gain and high channel gain. The high channel gain scenario represents the condition, where higher values of the channel gains are assumed for the links: $PU_i \leftrightarrow SUTR$ (s_i, c_i) and $SUTR \rightarrow SUR$ (ss) links, while the rest of the parameters are kept same as before. It was observed that $\mathbf{E}[PE]$ increases with an increase in the value of P_0, as expected. It was also observed that $\mathbf{E}[PE]$ increases with an increase in the value of channel gains between $PU_i \leftrightarrow SUTR$ and $SUTR \rightarrow SUR$. The mean value of $\mathbf{E}[PE]$ attained using the new and conventional strategies in the high channel gain scenario is 6.27 dB and 5.83 dB, respectively, while it is 4.22 dB and 3.89 dB, respectively in the low channel gain scenario. It is concluded that the new strategy outperforms the conventional strategy for both high and low channel gain scenarios. This is because PU_i has the interference from the SUTR transmissions during H_1 in new strategy. This interference is dealt as an extra source of energy by the SUTR. However, SUTR is not allowed to transmit during H_1 in the conventional strategy that effectively reduces the harvested energy at PU_i.

Fig. 6.3 shows the relationship between $\mathbf{E}[R_{SUR}]$ and P_0 for the new and conventional strategies using the same set of parameters as considered

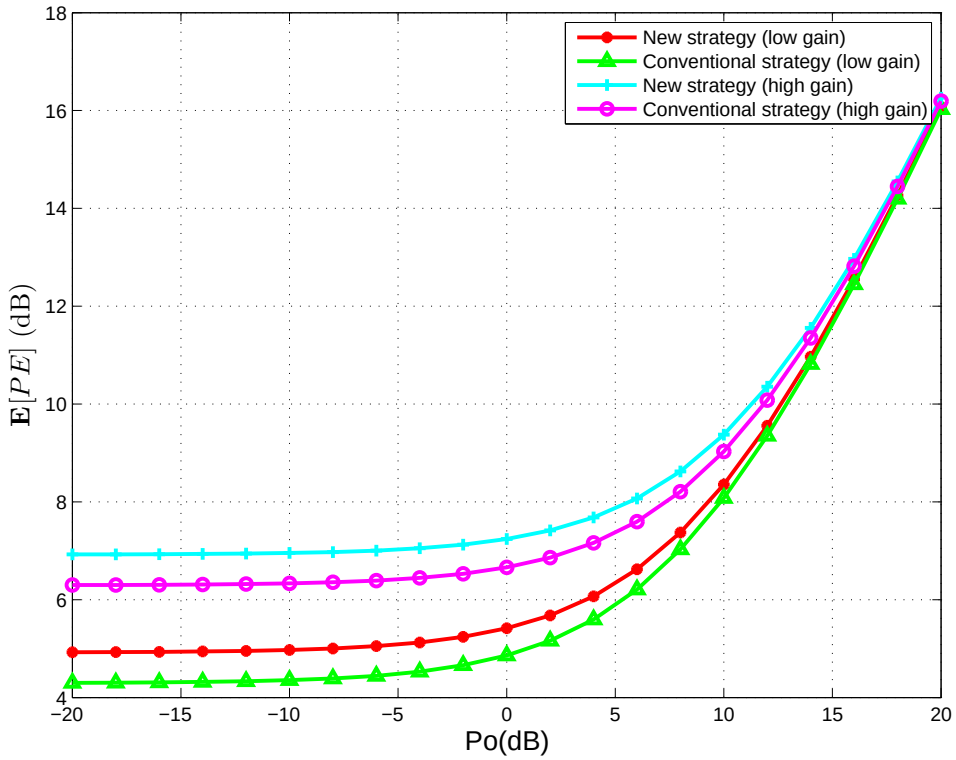


Figure 6.2: $\mathbf{E}[PE]$ Vs P_o.

above for Fig. 6.2. Two different values of P_i are considered in both high and low channel gain scenarios using the new and conventional strategies, where the first assumes that PU_i utilizes the fixed transmission power in the UL i. e. $P_i = 3.3$ dB, however the second assumes $\mathbf{E}[P_i] = \frac{\zeta \mathbf{E}[PE]}{\tau_i}$ in (6.16) and (6.23) respectively, where $\zeta = 1$. It was observed in Fig. 6.3, that the average throughput decreases with an increase in the value of P_0. This is because, $\mathbf{E}[R_{SUR}]$ and P_0 are inversely related to each other in (6.16) and (6.23). It was also observed in both sub-plots that the new strategy outperforms the conventional strategy for both high and low channel gain scenarios. The throughput attained using $\mathbf{E}[P_i] = \frac{\zeta \mathbf{E}[PE]}{\tau_i}$ in the second subplot is considerably smaller in value compared to the case, when $P_i = 3.3$ dB is used in the first sub-plot. This is because, the value of $\mathbf{E}[P_i]$ in the second sub-plot is attained using $\mathbf{E}[PE]$ in Fig. 6.2, where it is clearly depicted that the value of $\mathbf{E}[PE]$ increases exponentially between $P_0 = 5$ dB to $P_0 = 20$ dB using both the new and conventional strategies. As $\mathbf{E}[R_{SUR}]$ and $\mathbf{E}[P_i]$ are inversely related, therefore increasing the value of $\mathbf{E}[PE]$ eventually decreases the value of $\mathbf{E}[R_{SUR}]$.

Fig. 6.4 shows the relationship between $\mathbf{E}[SE]$ and P_0 using the same parameters as considered above for Fig. 6.3. Two different values of P_i are assumed as considered above for Fig. 6.3. It is observed in both sub-plots, that $\mathbf{E}[SE]$ increases with an increase in the value of P_0. It is also observed that conventional strategy outperforms new strategy in both sub-plots for both high and low channel gain scenarios. This is because SUT gets chance to harvest in both H_1 and H_2 in the conventional strategy. However SUT harvests only during H_2 in the new strategy.

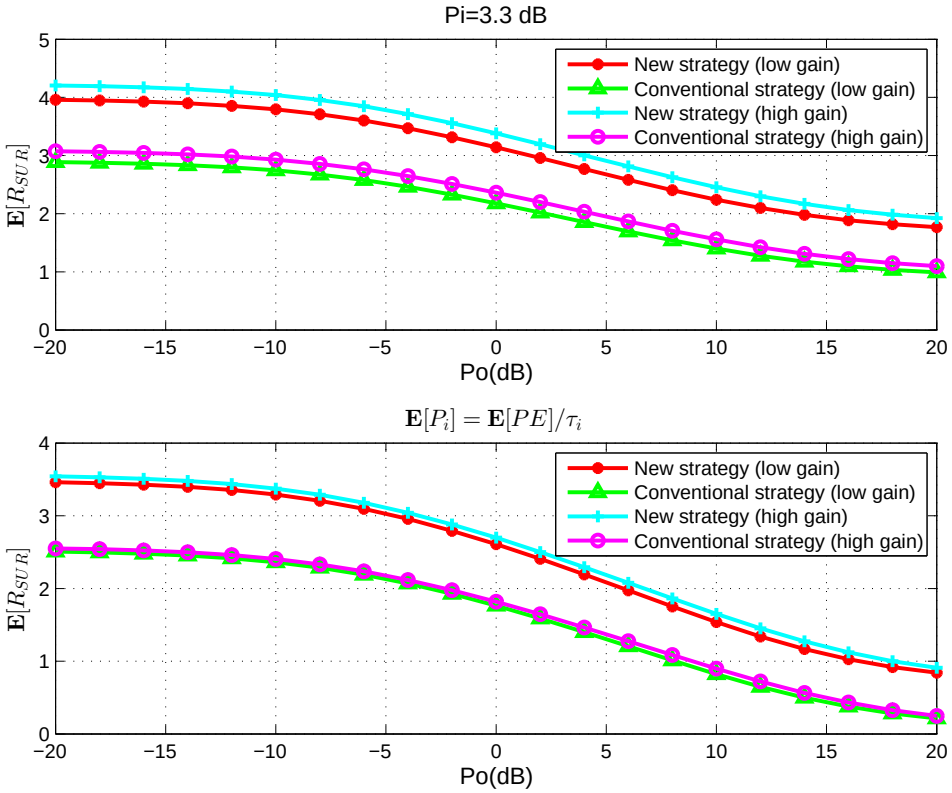


Figure 6.3: $\mathbf{E}[R_{SUR}]$ Vs P_o for (a) P_i=3.3 dB (b) $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$.

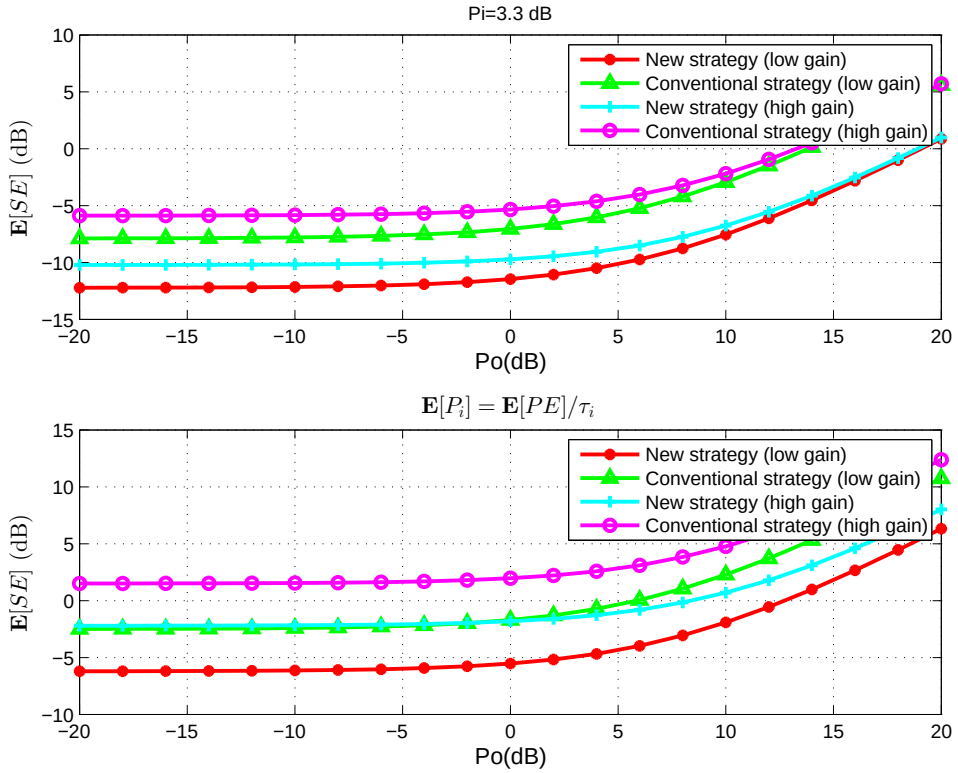


Figure 6.4: $\mathbf{E}[SE]$ Vs P_o for conventional method when SUTR harvests in both H_1 and H_2 using (a) P_i=3.3 dB (b) $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$.

6.4.2 Effect of changing P_s

The relationship between $E[PE]$ and P_s is analysed in Fig 6.5 by fixing the value of $P_0 = 10$ dB and using $P_s = [-20, 20]$ dB. The remaining parameters are assumed to be the same as in section 6.4.1. It was observed that $E[PE]$ increases with an increase in value of P_s. This is because $E[PE]$ is directly to P_s in (6.12) and (6.18). It was also observed that new strategy outperformed the conventional strategy for both high and low channel gain scenarios. This is because SUTR sends data to SUR during H_1 in the new strategy that acts as interference to PU. However PU harvests energy from the interference signals and eventually attain energy gain.

The relationship between $E[R_{SUR}]$ and P_s is analysed in Fig. 6.6. It was observed that the new strategy outperforms the conventional strategy for both high and low channel gain scenarios. This is because SUR receives data from the SUTR during both H_0 and H_1 states in the new strategy. However SUTR is only allowed to transmit to the SUR during H_0 in the conventional strategy.

When SUTR transmits data to SUR, the throughput of the secondary system can significantly be decreased due to the interference caused due to the AP transmission. This effect is analysed using both new and conventional strategies in high channel gain scenario using $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$, by considering two conditions: with interference (where it is assumed that SUR receive the signals from the AP in H_1 along with SUTR) and without interference (where it is assumed that SUR receive the signals from SUTR only in H_1). It was observed in Fig. 6.7, that maximum $E[R_{SUR}]$ attained using the new strategy without interference is larger than the value of the $E[R_{SUR}]$ attained, when interference is considered. Though AP transmission causes interference but it is evident in Fig 6.7 that new strategy outperformed the conventional strategy

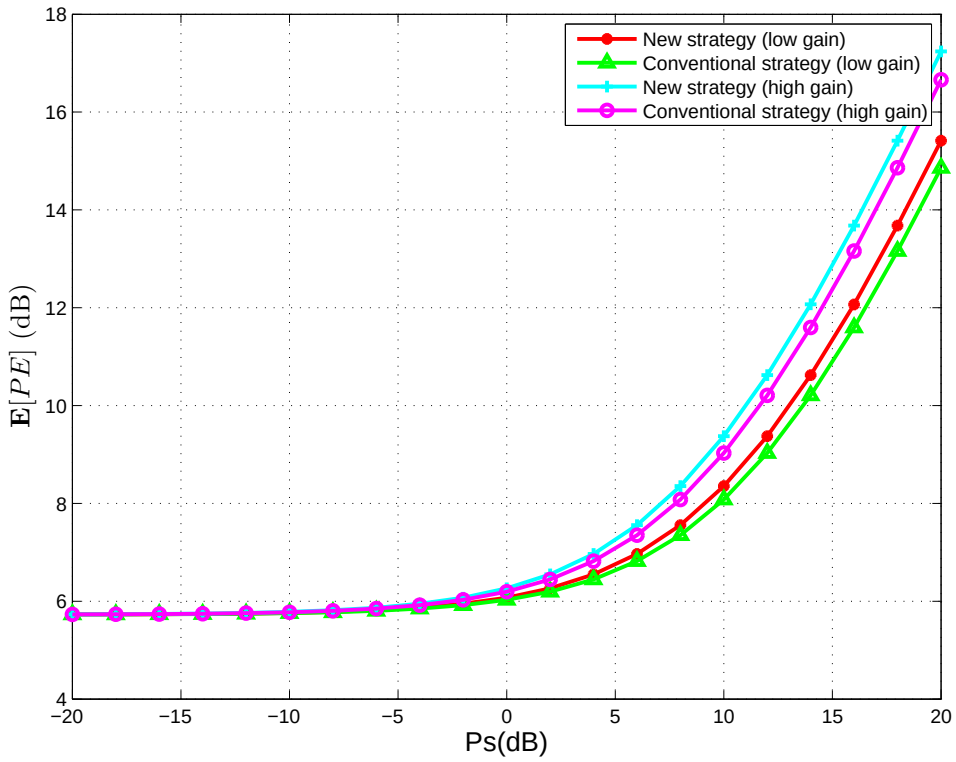


Figure 6.5: $\mathbf{E}[PE]$ Vs P_s.

in both conditions.

The relationship between $E[SE]$ and P_s is analysed in Fig. 6.8. It was observed that conventional strategy outperforms the new strategy again in this case. This is because SUTR harvests energy during both H_1 and H_2 in the conventional strategy compared to the new strategy, where SUTR harvests energy only in H_2.

6.4.3 Effect of changing τ_i

The effect of changing τ_i is analysed by assuming one PU in the framework that has time slot τ_1 for the uplink transmission, which implies: $\tau_0 + \sum_{i=1}^{K} \tau_i = 1$ would reduce to $\tau_0 + \tau_1 = 1$. The relationship between $E[PE]$ and τ_1 is analysed in Fig.6.9 by assuming range of values of τ_1, such as $\tau_1 = [0, 1]$ for

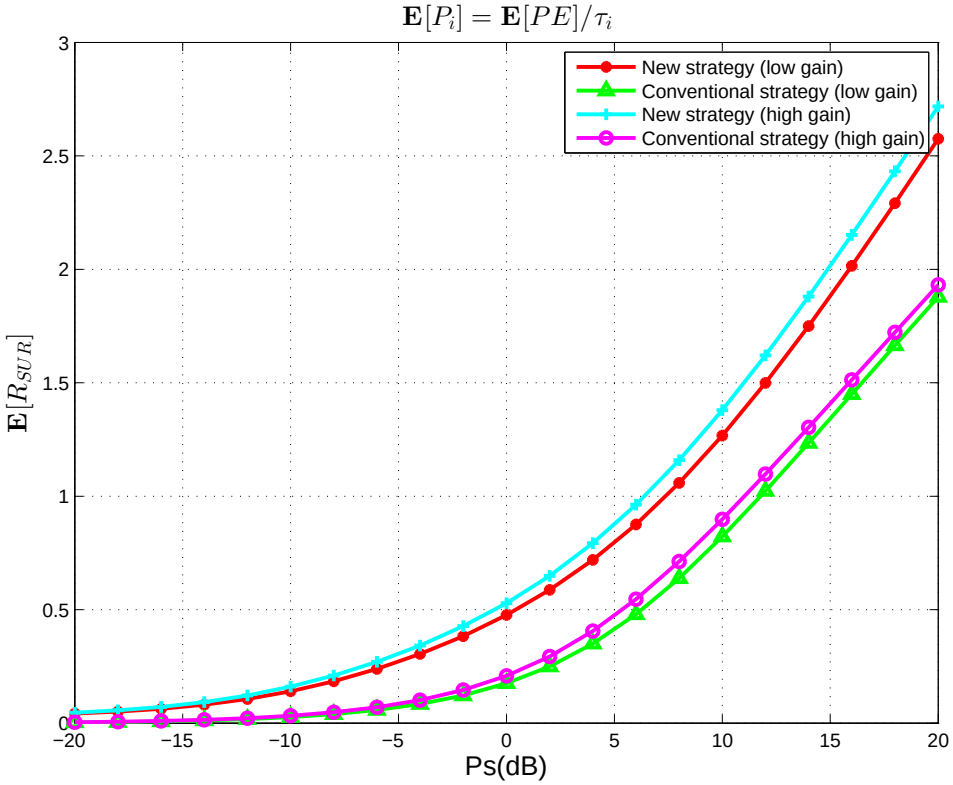


Figure 6.6: $\mathbf{E}[R_{SUR}]$ Vs P_s using $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$.

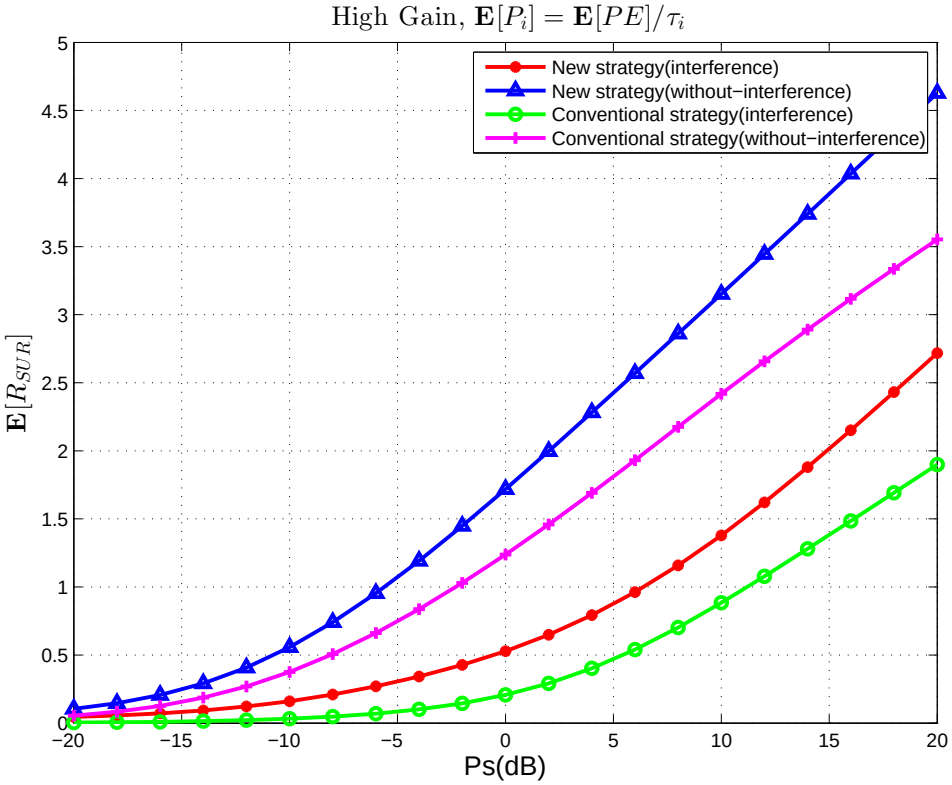


Figure 6.7: $\mathbf{E}[R_{SUR}]$ Vs P_s considering interference due to AP.

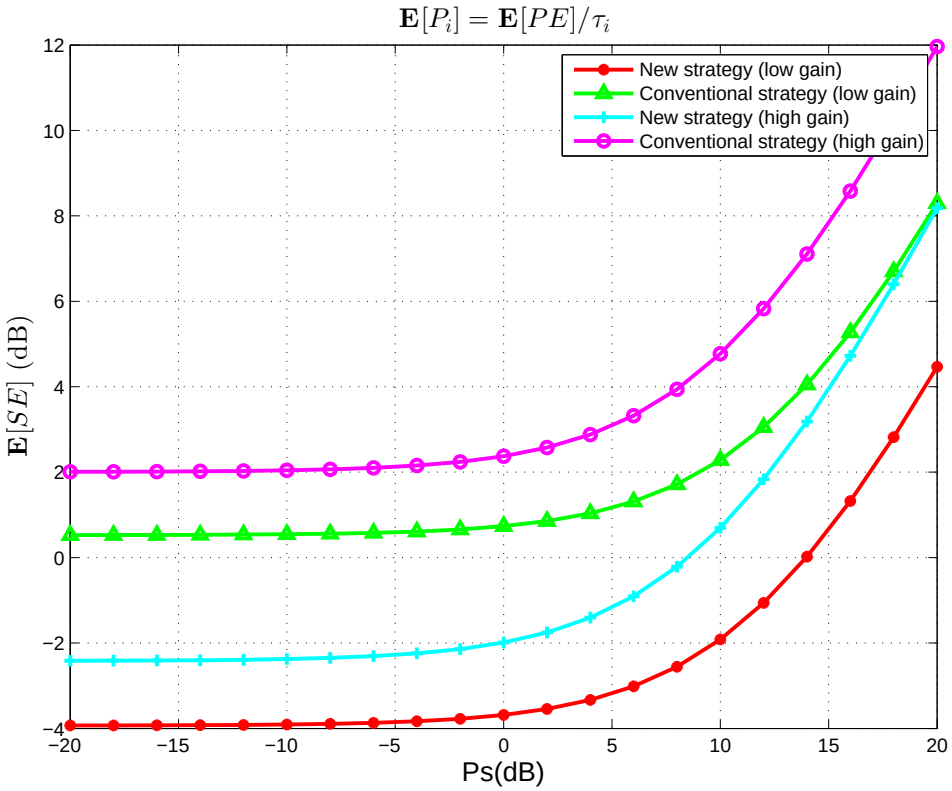


Figure 6.8: $\mathbf{E}[SE]$ Vs P_s for $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$.

high and low channel gain scenarios. It was observed in Fig. 6.9 that $E[PE]$ decreases with an increase in the value of τ_1. This is because $E[PE]$ is directly proportional to τ_0 in (6.12) and (6.18) for the new and conventional strategies respectively. The increase in the value of τ_1 causes decrease in value of τ_0 following $\tau_0 + \tau_1 = 1$, which implies that PU gets shorter duration to harvest energy.

The relationship between $E[R_{SUR}]$ and τ_1 is analysed in Fig. 6.10, assuming a range of values of τ_1 for high and low channel gain scenarios. It was observed that the new strategy outperformed the conventional strategy again. It was also observed in Fig. 6.10 that $E[R_{SUR}]$ increases with an increase in the value of τ_1. This is because $\mathbf{E}[P_i]$ decreases with an increase in the value of τ_1, following the relationship:$\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$. As $E[P_i]$ and $E[R_{SUR}]$ are inversely related to each other in (6.16) and (6.23), so decrease in the value of $E[P_i]$ increases $E[R_{SUR}]$ eventually.

The relationship between $E[SE]$ and τ_1 investigated in Fig. 6.11, assuming a range of values of τ_1 for high and low channel gain scenarios. It was observed that conventional strategy outperformed the new strategy because SUTR can harvest energy during both H_1 and H_2. It was observed that $E[SE]$ decrease with an increase in the value of τ_1. This is because of two reasons: (a) τ_0 is directly proportional to $E[SE]$ in (6.14) and (6.21) so decrease in τ_0 decreases the value of $E[SE]$ (b) P_i is also directly proportional to $E[SE]$ in (6.14) and (6.21). Therefore when the value of P_i decreases the value of $E[SE]$ decreases as well.

6.5 Conclusion

In this chapter, an energy harvesting cognitive radio network (CRN) is considered, where the primary users (PUs) are wirelessly powered with energy

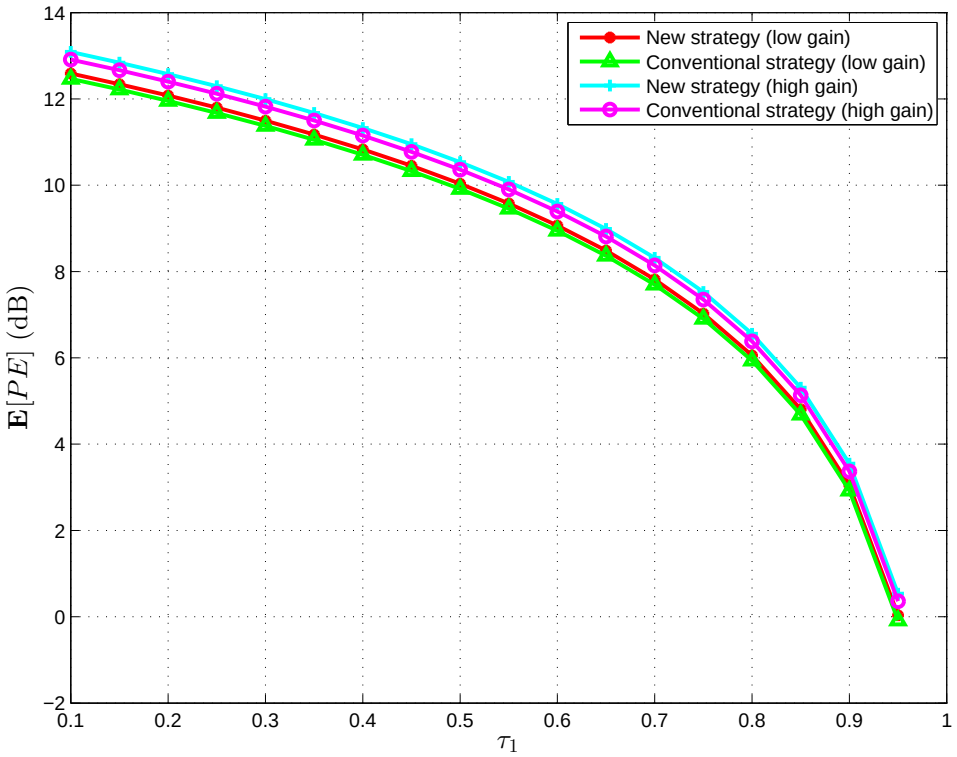


Figure 6.9: $\mathbf{E}[PE]$ Vs τ_1.

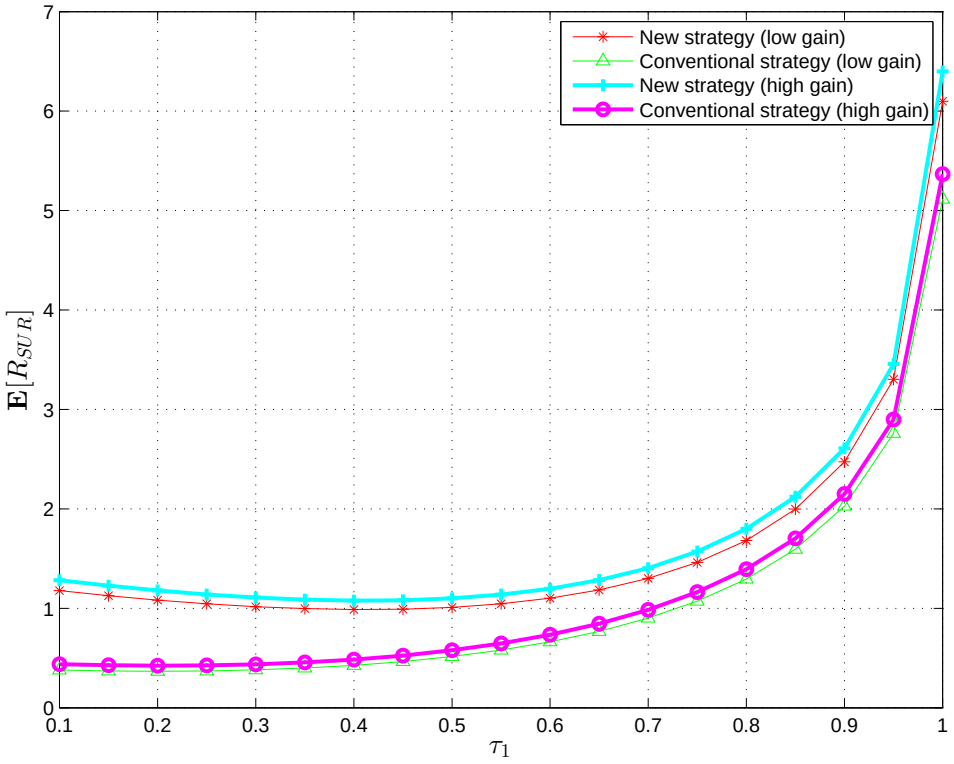


Figure 6.10: $\mathbf{E}[R_{SUR}]$ Vs τ_1 for (a) P_i=3.3 dB (b) $\mathbf{E}[P_i] = \mathbf{E}[PE]/\tau_i$.

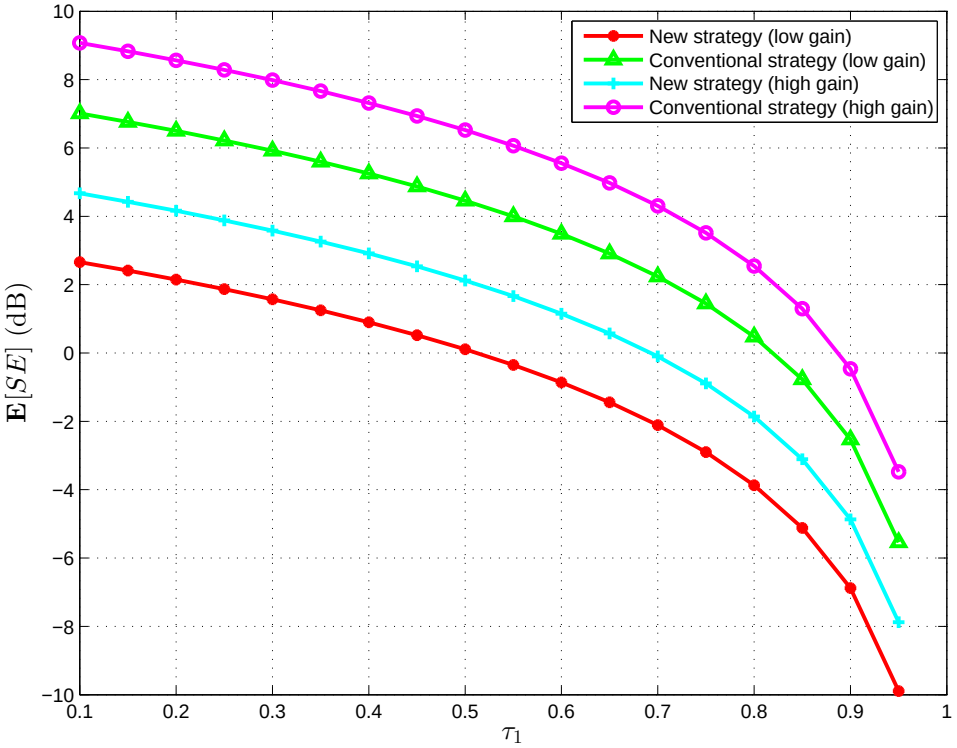


Figure 6.11: $\mathbf{E}[SE]$ Vs τ_1.

harvesting capability. A framework has been proposed, where PUs can get advantage from the presence of SU. The proposed framework is compared with the conventional energy harvesting CRN. Numerical results have shown that PU attains energy gain from the transmission of SU and SU attains throughput gain from the energy harvesting capability of PUs.

Chapter 7

Conclusions and Future Work

The work presented in this thesis has focussed on the performance analysis of cognitive radio using machine learning and bio-inspired algorithms. Although adaptation to the environment, self-organization and self-configuration independently is very challenging task for any communication system, however machine learning and bio-inspired algorithms have the ability to adapt according to the changing environmental circumstances and are able to evolve under new conditions. In this chapter, the main contributions and findings are summarized. Also the potential research directions are provided. The rest of the chapter is organized as follows. The main contributions and findings are presented in Section 7.1 and the future works are suggested in Section 7.2.

7.1 Conclusions and Contributions

Cognitive radio is the key solution for resolving the conflict between spectrum scarcity and under-utilisation using dynamic spectrum access. The main aim of this work is to present an analytical performance evaluation of the machine learning and the energy efficient cognitive radio systems while taking some realistic conditions into account.

Firstly, in Chapter 1, the overview of the cognitive radio networks including the spectrum sharing and access modes, the specifications of IEEE 802.22 standard and the potential applications of the cognitive radio systems is presented. Then, a detailed background about spectrum sensing methods, bio-inspired techniques and machine learning algorithms is presented in Chapter 2. A comprehensive overview about the energy harvesting cognitive radios and the wireless power transfer methods is also presented.

Third, as mentioned in the introduction of this thesis, collaboration between cognitive radios is encouraged for attaining the accurate spectrum sensing results in low SNR conditions. In Chapter 3, bio-inspired algorithms are utilized for the collaborative spectrum sensing and allocation. Both linear and non-linear PU signals are considered and their performance is compared using three bio-inspired techniques. It is observed that the performance of the non-linear signals is degraded more compared to the linear signals. Based on the analysis, it is also observed that bio-inspired algorithms have outperformed the conventional techniques utilized in the literature for collaborative spectrum sensing and allocation. Among three bio-inspired techniques, it is noticed that FFA outperforms PSO and FSS. The performance of the proposed model is also investigated in different SNR conditions and using different number of the collaborating radios. The spectrum allocation framework based on the spectrum sensing results is also presented in Chapter 3. Three spectrum allocation objective functions are considered and it is observed that among all bio-inspired algorithms, FFA has outperformed FSS and PSO by attaining the highest maximum sum reward, maximum minimum reward and the maximum proportional fair reward.

Fourth, as discussed earlier, the selection of the frequency, the time duration, modulation technique and the transmission power for CRs is depen-

dent on the accurate spectrum sensing results. However, the computation of spectrum sensing requires a lot of resources and consumes time. The spectrum sensing performance can be improved using spectrum occupancy analysis. In Chapter 4, four machine learning algorithms are utilized for analysing spectrum occupancy. The performance comparison of the machine learning algorithms has shown that SVM performs the best among all supervised and unsupervised algorithms. It is noticed that performance of the SVM is affected by the change in the value of the box constraints. A bio-inspired technique, FFA is utilized to evaluate the optimal value of the box constraints. It is observed that SVM+FFA outperform other algorithms including SVM. The performance of the machine learning algorithms is also compared with a statistical approach used for classification in the literature. The proposed framework also presents a method for selecting the optimal parameters for all supervised and unsupervised algorithms that helps to achieve higher value of the classification accuracy. Based on the analysis, two distinct categories of the spectrum bands were observed in the data. The categorization of the bands was done on the basis of their mean input power, the maximum standard deviation and their periodicity.

Fifth, the wireless networks are the main consumers of the energy in the telecommunication infrastructure and their current power requirements are increasing because of an exponential increase in the mobile data traffic. Energy harvesting is one of the key solutions to achieve green communication. However the wireless energy changes dynamically with time, frequency and distance. In Chapter 5, a machine learning based energy predictive model is presented. The optimal time slots and the frequency channels, where more energy can be harvested is evaluated using the proposed framework. The analysis is verified using two prediction algorithms and it is noticed that both

proposed algorithms outperform the conventional prediction scheme in the literature. The harvested power using low and medium efficiency harvesters has been evaluated. It is noticed that harvested power using MEH is greater than LEH, when data of different frequency bands was used. The prediction error for different bands has also been investigated. Importantly, it has been noticed that prediction error for the aperiodic bands is larger than periodic bands.

Finally, a novel energy harvesting wireless powered cognitive radio framework is presented in Chapter 6. In the proposed framework, energy harvesting capabilities are proposed for both PUs and SU, where PUs harvest energy when SU transmits data and vice versa. Different power levels for both PU and SU have been utilized to test the model. Further, both low and high channel gain scenarios are investigated for the proposed framework. Based on the analysis, it is observed that the proposed framework outperform the conventional techniques because PUs attain energy gain from the SU transmission. Moreover SU achieves throughput gain because SU gets extra transmission time to send data while PUs are harvesting energy.

7.2 Future Research Directions

Based on the summary and findings about the conducted research in this thesis, it can be concluded that there are still a lot of improvements that can be done. Some of the guidelines and suggestions for the future research are presented as follows.

The proposed framework in Chapter 3 has presented the comparison of different bio-inspired techniques, where it was observed that each technique has its own pros/ cons. Although FFA has shown the best performance but it has some disadvantages. For example, it can be trapped into local optima

while performing local search. Similarly the computational time for FFA is higher than PSO. For combating this in future, novel and hybrid bio-inspired algorithms can be developed, where the good features of the best algorithms can be integrated. The hybrid bio-inspired algorithms would be useful for globally optimizing the problem space while gaining convergence at the same time.

The proposed intelligent framework in Chapter 3 and Chapter 4 deals with bio-inspired and machine learning algorithms that are utilized to learn from previous experiences and take optimal decisions. The intelligent framework presented in this work can be improved by deploying human neo-cortical learning architecture, called as, Hierarchical Temporal memory (HTMs). HTM is a machine learning intelligence technology that aims to capture the structural and algorithmic properties of the neocortex [160]. HTM is specifically designed for learning in an unsupervised manner from the environment and uses its previous knowledge to make future predictions. It learns the time based patterns in data, predicts future values, and detects anomalies. It has been reported in [161] that this technology lays the groundwork for the new era of machine intelligence. HTMs can help to use cognitive radio as one of the key enabler for next generation networks (NGNs) because the intelligence is required at each node in NGNs for executing the computations at run time.

Similarly the predictive modelling framework presented in Chapter 5 analyses the real time power measurements attained using spectrum analyser, where the harvested energy is evaluated using the RF energy harvesters in [145], [146]. The proposed framework in Chapter 5 can be improved in future by deploying the RF energy harvesters at the University of Warwick. In this scenario, the data attained using harvesters will give a better insight about the variations of harvested energy with time and frequency.

Furthermore, the complexity of the linear regression model in Chapter 5 linearly increases with an increase in the number of predictors. Sophisticated techniques, such as principal component analysis (PCA) can be utilized in future for evaluating the appropriate number of predictors that can keep maximum variance in the data and can ensure satisfactory performance at the same time.

In Chapter 6, the energy harvested by PU and the throughput attained by SU are computed without considering collision and energy constraints. In reality, there are constraints in the networks. Therefore in future, the analysis of the unconstrained network in Chapter 6 can be improved by considering collision constraint, that indicates the maximum permissible probability of interference which can be tolerated by the PU network and energy causality constraint, that determines the balance between the harvested and the consumed energy [162]. By considering the collision and energy causality constraints at the same time, the energy and spectrum efficiency can be jointly optimized.

Furthermore in Chapter 6, information and energy signals are distinguished using different power levels. However this still needs to be addressed that how information and energy signals can be separated if they have the same power level. For addressing the mentioned scenario in future, different features of the information and the energy signals can be extracted using sophisticated feature detection techniques.

Bibliography

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

[2] M. A. McHenry, "NSF spectrum occupancy measurements project summary", *shared spectrum co. report*, December 2005.

[3] M. Wellens, J. Wu and P. Mahonen, "Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio", *Proc. CrownCom'07*, pp. 420 - 427, Orlando, USA, Aug. 2007.

[4] M. H. Islam, C. L. Koh, S. W. Oh, et all, "Spectrum survey in Singapore: occupancy measurements and analyses", Proc. CrownCom'08, pp. 1 - 7, May 2008.

[5] R. Rajbanshi, "OFDM-based cognitive radio for DSA networks", *Technical Report ITTC-FY2008-TR-31620-05*, September 2007.

[6] M. L. Bentez, A. Umbert and F. Casadevall, "Evaluation of spectrum occupancy in spain for cognitive radio applications", *Proc. VTC-Spring09*, pp. 1 - 5, Apr. 2009.

[7] J. Mitola, "Cognitive radio for flexible mobile multimedia communications", *Proc. Mo-MuC 1999*, pp. 3 - 10.

[8] S. Haykin, "Cognitive radio: brain-empowered wireless communications", *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201 - 220, 2005.

[9] F. K. Jondral, "Software-defined radio: basic and evolution to cognitive radio", *EURASIP Journal on Wireless Communication and Networking*, vol. 2005, no. 3, pp. 275-283, Aug 2005.

[10] J. Xiang, Y. Zhang and T. Skeie, "Medium access control protocols in cognitive radio networks", *Wiley Wireless communications and mobile computing*, vol. 10, pp. 31 - 49, Dec 2009.

[11] D. Cavalcanti and M. Ghosh, "Cognitive radio networks: enableing new wireless broadband opportunities", *Proc. CrownCom' 08*, pp. 1 - 6, May 2008.

[12] A. Ghasemi and E. S. Sousa, "Collaborative spectrum sensing for opportunistic access in fading environments", *Proc. DySPAN'05*, pp. 131 - 136, Nov. 8 - 11 2005, Baltimore.

[13] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation issues in spectrum sensing for cognitive radios", *Proc. Asilomar Conference*, pp. 772 - 776, Nov. 7 - 10 2004, Pacific Grove, CA.

[14] P. Leaves et al., "Dynamic spectrum allocation in a multiradio environment concept and algorithm", *Second International Conf. 3G Mobile Commun. Technologies*, pp. 53 - 57, March 2001, London, UK.

[15] C. X. Wang, X. Hong, H. H. Chen and J. Thompson, "On capacity of cognitive radio networks with average interference power constraints", *IEEE Trans. Wireless Communications*, vol. 8, no. 4, pp. 1620-1625, April 2009.

[16] A. Ghasemi and E. Sousa, "Fundamental limits of spectrum sharing in fading environments", *IEEE Trans. on Wireless Communication*, vol. 6, no. 2, pp. 649 - 658, 2007.

[17] Federal Communications commission, "Establishment of an interference temperature metric to quantify and manage interference and to expand available unlicensed operation in certain fixed, mobile and satellite frequency bands", *ET Docket No. 03-237*, Aug. 2004.

[18] Y. Han, A. Pandharipande and S. H. Ting, "Cooperative decode-and-forward relaying for secondary spectrum access," *IEEE Transactions on Wireless Communication*, vol. 8, no. 10, pp. 4945 - 4950, Oct. 2009.

[19] S. Shellhammer and G. Chouinard, "Spectrum sensing requirements summary", July 2006, Available Online: http://www.ieee802.org.

[20] C. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. Shellhammer, and W. Caldwell, "IEEE 802.22: the first cognitive radio wireless regional area network standard", *IEEE Communications Magazine*, vol. 47, pp. 130 - 138, Jan. 2009.

[21] M. Sherman, A. N. Mody, R. Martinez and C. Rodriguez, "IEEE standards supporting cognitive radio and networks, dynamic spectrum access and coexistence", *IEEE Communications Magazine*, vol. 46, no. 7, pp. 72-79, July 2008.

[22] H. R. Kim, H. Venkatesan, "ECEN 5692-Principals of digital communication project on IEEE 802.22", Available Online: http://ecee.colorado.edu/ ecen4242/802˙22/.

[23] E. Biglieri, A. J. Goldsmith, L. J. Greenstein, N. B. Mandayam and H. V. Poor, "Principles of cognitive radio", *Cambridge university press*, ISBN: 9781107028753, Nov 2012.

[24] T. Li, N. Mandayam and A. Reznik, "A framework for resource allocation in cognitive digital home", *Proc. GLOBECOM'10*, pp. 1 - 5, Dec 2010.

[25] T. Li, N. Mandayam and A. Reznik, "Distributed algorithms for joint channel and RAT allocation in a cognitive digital home", *Proc. WiOpt*, pp. 213 - 219, May 2011.

[26] G. Feng, Y. Wei, L. Wei, C. Wenqing, W. Shu, "Pipelined cooperative spectrum sensing in cognitive radio networks", *Proc. IEEE wireless communications and networking conference*, pp. 1 - 5, 2009.

[27] Y. Chen, "Collaborative spectrum sensing in the presence of secondary user interference for lognormal shadowing", *Wiley Journal of Wireless Communications and Mobile Computing*, vol. 12, no. 10, pp. 463 - 472, 2012.

[28] V. G. Chavali, C. R. Dasilva, "Collaborative spectrum sensing based on new SNR estimation and energy combining method", *IEEE Transactions on Vehicular technology*, vol. 60, no. 8, pp. 4024 - 4029, 2011.

[29] Z. Quan, A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks", *IEEE Journal of selected topics in signal processing*, vol. 2, no. 1, 2008.

[30] C. Peng, H. Zheng, B. Zhao, "Utilization and fairness in spectrum assignment for opportunistic spectrum", *ACM Journal of mobile networks and application*, vol. 11, no. 4, pp. 555 - 576, 2006.

[31] Z. Zhao, Z. Peng, S. Zheng and J. Shang ,"Cognitive radio spectrum allocation using evolutionary Algorithms", *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4421-4425, Sept 2009.

[32] V. Blaschke, H. Jaekel, T. Renk, C. Kloeck, F. K. Jondral, "Occupation measurements supporting dynamic spectrum allocation for cognitive radio design", *Proc. CrownCom'07*, pp. 50 - 57, Orlando, Florida, Aug. 2007.

[33] S. Kaneko, S. Nomoto, T. Ueda, S. Nomura and K. Takeuchi, "Predicting radio resource availability in cognitive radio - an experimental examination", *Proc. CrownCom'08*, Singapore, May. 2008.

[34] M. Hoyhtya, S. Pollin, A. Mammela, "Classification - based predictive channel selection for cognitive radios", *Proc. ICC'10*, pp. 1 - 6, Cape town, South Africa, May. 2010.

[35] X. Zhou, J.Ma, Y. Li, Y. H. Kwon, A. C. K. Soong, G. Zhao, "Probability-based transmit power control for dynamic spectrum access", *Proc. DySPAN'08*, pp. 1 - 5, Chicago, USA, Oct. 2008.

[36] X. Zhou, J. Ma, Y. Li, Y. H. Kwon, A. C. K. Soong, "Probability-based optimization of inter-sensing duration and power control in cognitive radio", *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4922 - 4927, Oct. 2009.

[37] Smart2020, "Enabling the low carbon economy in the information age", *The Climate Group SMART 2020 Report*, Available online: http://www.theclimategroup.org.

[38] A. Dejonghe, B. Bougard, S. Pollin, J. Craninckx, A. Bourdoux, L. V. der Perre, et al, "Green reconfigurable radio systems", *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 90 - 101, May 2007.

[39] M. Pinuela, P. D. Mitcheson, S. Lucyszyn, "Ambient RF energy Harvesting in Urban and Semi-Urban environments", *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 7, pp. 2715 - 2726, July 2013.

[40] S. Park, H. Kim, and D. Hong, "Cognitive radio networks with energy harvesting", *IEEE Transactions on Wireless Communications*, vol. 12, no. 3, pp. 1386-1397, March 2013.

[41] S. Lee, R. Zhang and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks", *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, September 2013.

[42] T. Nguyen, I. Koo, "Throughout of primary user with cognitive radio function", *Proc. ISCIT*, 2014.

[43] Z. Wang, Z. Chen, Y. Yao, B. Xia and H. Liu, "Wireless energy harvesting and information transfer in cognitive two-way relay networks", *Proc. Globecom'14*, Dec 2014.

[44] G. Zheng, Z. Ho, E. A. Jorswieck, B. Ottersten, "Information and energy cooperation in cognitive radio networks", *IEEE Transactions on Signal Processing*, vol. 62, no. 9, May 2014.

[45] E. Matskani, N. Chatzidiamantis, L. Georgiadis, I. Koutsopoulos and L. Tassiulas, "The mutual benfits of primary-secondary user cooperation in wireless cognitive networks", *Proc. WiOpt*, 2014.

[46] C. Clancy, J. Hecker, E. Stuntebeck, and T. O. Shea, "Application of machine learning to cognitive radio networks", *IEEE Wireless Communications*, vol. 14, no. 4, pp. 47 - 52, 2007.

[47] S. M. Kay, "Fundamentals of statistical signal processing: detection theory", *Prentice- Hall PTR*, 1998.

[48] Z. Lei and F. Chin, "OFDM signal sensing for cognitive radios", *Proc. PIMRC'08*, France, Sept 2008.

[49] H. Urkowitz, "Energy detection of unknown deterministic signals", *Proceedings of IEEE*, vol. 55, no. 4, pp. 523-531, April 1967.

[50] T. Yucek, H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications", *IEEE Communications Surveys and Tutorials*, vol. 11, no. 1, pp. 16-30, 2009.

[51] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications", Proc. MoMuC99, pp. 3-10. 1999.

[52] H. Tang, "Some physical layer issues of wide-band cognitive radio systems", *Proc. DySPAN'05*, pp. 151-159, Nov 2005.

[53] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio", *Proc. Allerton Conf. Commun. Control, Comput*, Oct. 2004.

[54] W. A. Gardner, "Cyclostationarity in Communications and Signal Processing", IEEE Press, 1994.

[55] W. A. Gardner, "Signal interception: a unifying theoretical framework for feature detection", *IEEE Transactions on Communications*, vol. 36, no. 8, pp. 897-906, Aug 1988.

[56] W. A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals", *IEEE signal processing magazine*, vol. 8, no. 2, pp. 14-36, April 1991.

[57] Y. Zeng and Y. C. Liang, "Maximum-minimum eigenvalue detection for cognitive radio", *Proc. IEEE PIMRC*, Athens, Greece, Sept 2007.

[58] Y. Zeng, C. L. Koh and Y. C. Liang, "Maximum eigenvalue detection: theory and applications", *Proc. ICC'08*, pp. 4160-4164, Beijing, May 2008.

[59] Y. Zeng and Y. C. Liang, "Eigen value-based spectrum sensing algorithms for cognitive radio", *IEEE Transaction on Communications*, vol. 57, no. 6, pp. 1784-1793, June 2009.

[60] Y. Yuan, P. Bahl, R. Chandra, P. A. Chou, J. I. Ferrell, T. Moscibroda, S. Narlanka, and Y. Wu, "KNOWS: Cognitive radio networks over white spaces", *Proc. DySPAN'07*, pp. 416427, April 2007.

[61] Y. Zeng, C. L. Koh, and Y.-C. Liang, "Maximum eigenvalue detection: Theory and application", *Proc. ICC'08*, pp. 4160-4164, May 2008.

[62] M. Ibnkahla, "Cooperative cognitive radio networks", *Boca Raton, FL : CRC Press*, 2015, ISBN 9781466570795.

[63] L. Gavrilovska, V. Atanasovski, I. Macaluso and L. A. DaSilva, "Learning and reasoning in cognitive radio networks", *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, 2013.

[64] L. Gavirlovska, V. Atanasovsjo, I. Macaluso and L. DaSilva, "Learning and reasoning in cognitive radio networks", *IEEE Communications Surveys and Tutorials*, vol. 15, no.4, pp. 1761-1771, 2013.

[65] D. Floreano and C. Mattiussi, "Bio-inspired artificial intelligence", *MIT Press*, ISBN-13: 978-0-262-06271-8, Sept 2008.

[66] F. Dressler, O. B. Akan, "A Survey on Bio-inspired Networking", *Elsevier Computer Networks*, vol. 54, no. 6, pp. 881-900, 2010.

[67] P. D. Lorenzo, S. Barbaroosa, "A Bio-Inspired swarming algorithm for decentralized access in cognitive radio", *IEEE Transactions on signal Processing*, vol. 59, no 12, pp. 6160 - 6174, Dec 2011.

[68] A. Ghasemi, M. A. Mehrzad, "Spectrum allocation based on artificial bee colony in cognitive radio networks", *Proc. IST'12*, pp. 182-187, 6-8 Nov, 2012, Tehran.

[69] T. Renk, C. Kloeck, D. Burgkhardt, "Bio-Inspired algorithms for dynamic resource allocation in cognitive wireless networks", *Proc. CrownCom'07*, Orlando, USA, Aug. 2007.

[70] S. Zheng,C. Lou and X. Yang, "Co-operative spectrum sensing using particle swarm optimization", *IET Electronic Letters*, vol. 46, no. 22, Oct 2010.

[71] X. Cheng, M. Jiang, "Cognitive radio spectrum assignment based on artificial bee colony optimization", *Proc. ICCT'11*, pp. 161-164, Sept 2011.

[72] F. Camastra and A. Vinciarelli, "Machine learning for audio, image and video analysis". *Wiley publishers*, 2006.

[73] R. Herbrich, "Learning Kernel Classifiers". *MIT Press*, 2003.

[74] T. Kohonen, "The self-organizing map", *Neurocomputing*, vol. 21, pp. 16, November 1998.

[75] W. L. Chao, "Machine learning tutorial", *Graduate institute of communication engineering*, national Taiwan university. Available online: http://disp.ee.ntu.edu.tw/ pujols/Machine

[76] R. Sutton and A. Barto, "Reinforcement Learning: An Introduction". *MIT Press*, 1998.

[77] K. M. Thilina, K. W. Choi, N. Saquib and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks", *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 11, Nov 2013.

[78] J. Oksanen, J. Lunden, V. Koivunen, "Reinforcement learning based sensing policy optimization for energy efficient cognitive radio networks", *Neurocomputing*, vol. 80, pp. 102-110, March 2012.

[79] S. Hu, Y. Yao and Z. Yang, "MAC protocol identification using support vector machines for cognitive radio networks", *IEEE Wireless Communications*, vol. 21, no. 1, Feb 2014.

[80] G. Ding, Q. Wu, Y. Yao, J. Wang, and Y. Chen, "Kernel-based learning for statistical signal processing in cognitive radio networks", *IEEE Signal Processing Magazine*, vol. 30, no. 4, July 2013.

[81] H. Pervaiz, L. Musavian, Q. Ni and Z. Ding, "Energy and spectrum efficient transmission techniques under QoS constraints towards green heterogeneous networks", *IEEE Access*, vol. 3, pp. 1655-1671, Sept 2015.

[82] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access", *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 7989, May 2007.

[83] S. Park, H. Kim, D. Hong, "Cognitive radio networks with energy harvesting", *IEEE transactions on wireless communications*, vol. 12, no. 3, march 2013.

[84] T. Le, K. Mayaram, and T. Fiez, "Efficient far-field radio frequency energy harvesting for passively powered sensor networks", *IEEE J. Solid- State Circuits*, vol. 43, no. 5, pp. 1287-1302, May 2008.

[85] A. M. Zungeru, L. M. Ang, S. Prabaharan, and K. P. Seng, "Radio frequency energy harvesting and management for wireless sensor networks", *Green Mobile Devices and Networks: Energy Optimization and Scavanging Techniques*, CRC Press, pp. 341-368, 2012.

[86] R. J. M. Vullers, R. V. Schaijk, I. Doms, C. V. Hoof, and R. Mertens, "Micropower energy harvesting", *Elsevier Solid-State Circuits*, vol. 53, no. 7, pp. 684 - 693, July 2009.

[87] X. Lu, P. Wang, D. Niyato and E. Hossain, "Dynamic spectrum access in cognitive radio networks with RF energy harvesting", *IEEE Wireless Communications*, vol. 21, no. 3, pp. 102-110, June 2014.

[88] S. Lee, R. Zhang and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks", *IEEE Transactions on Wireless Communications*, vol. 12, No. 9, Sept 2013.

[89] Z. Ding, C. Zhong, D. W. K. Ng, M. Peng, H. A. Suraweera, R. Schober, H. V. Poor, "Application of smart antenna technologies in simultaneous wireless information and power transfer", *IEEE Communications Magazine*, vol. 53, no. 4, pp. 86-93, April 2015.

[90] X. Chen, C. Yuen, and Z. Zhang, "Wireless energy and information transfer trade-off for limited feedback multiantenna systems with energy beamforming", *IEEE Transactions Vehicular Technology*, vol. 63, no. 1, pp. 407-412, Jan.2014.

[91] X. Chen, Z. Zhang, H. Chen, H. Zhang, "Enhancing wireless information and power transfer by exploiting multi-antenna techniques", *IEEE Communications Magazine*, vol. 53, no. 4, pp. 133-141, April 2015.

[92] S. Dikmese, S. Srinivasan, M. Shaat, *et al.*, "Spectrum sensing and resource allocation for multicarrier cognitive radio systems under interference and power constraints", *EURASIP Journal on Advances in Signal Processing*, DOI: 10.1186/1687-6180-2014-68, Dec 2014.

[93] T. S. Ghazaany, "Design and implementation of adaptive base band predistorter for OFDM nonlinear transmitter", PhD Thesis, University of Bradford, U. K., 2011.

[94] D. W. Chi, P. Das, "Effects of nonlinear amplifiers and narrow band interference in MIMO-OFDM with application to 802.11n WLAN", *Proc. ICSPC*, pp 1-7, Dec 2008.

[95] A. A. M. Saleh, "Frequency-independent and frequency-dependent nonlinear models of TWT amplifiers", *IEEE Transactions on Communications*, vol. 29, no. 11, pp. 1715-1720, Nov 1981.

[96] Z. Quan, A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks", *IEEE Journal of selected topics in signal processing*, vol. 2, no. 1, pp. 28-40, Feb 2008.

[97] M. F. Lohrer, "A comparison between the firefly algorithm and particle swarm optimization", PhD Thesis, Oakland University, U. S. A, March 2013.

[98] G. M. Cavalcanti, C. J. A. Bastos-Filho, F. B. Lima-Neto, *et al*, "A hybrid algorithm based on fish school search and particle swarm optimization for dynamic problems", *Advances in Swarm Intelligence*, Lecture Notes in Computer Science, vol. 6729, pp 543-552, 2011.

[99] X. Yang, "Firefly algorithms for multimodal optimization", *Springer Lecture Notes in Computer Science*, vol. 5792, pp. 169-178, 2009.

[100] A. B. Filho, and F. B. De Lima Neto, "A novel search algorithm based on fish school behaviour", *Proc. ICSMC'08*, pp. 2646 - 2651, Oct 2008.

[101] R. Poli, J. Kennedy and T. Blackwell, "Particle Swarm Optimization", *Springer Journal on Swarm Intelligence*, vol. 1, no. 1, pp. 33-57, Aug 2007.

[102] X. Yang, "Nature - inspired metaheuristic algorithms", Luniver Press, U. K, 2nd edn. 2010.

[103] J. Kennedy, R. C. Eberhart, "Particle swarm optimization", *Proc. ICNN*, vol. 1, pp. 42 - 48. May 2001.

[104] Y. Shi, R. Eberhart, "Parameter selection in particle swarm optimization", *Proc'. Evolutionary Programming*, vol. 1447, pp. 591-600, Dec 2005.

[105] R. Eberhart, Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization', *Proc. Congress on Evolutionary Computation*, vol. 1, pp. 8488, July 2000.

[106] M. Erik, H. Pedersen, "Good Parameters for Particle Swarm Optimization", *Hvass Laboratories*, Technical Report no. HL1001, 2010.

[107] D. Tse, P. Viswanath, "Fundamentals of wireless communication", *Cambridge University Press*, 2005.

[108] Y. Chen, H-S. Oh, "A survey of measurement-based spectrum occupancy modelling for cognitive radios", *IEEE Communications Surveys and Tutorials*, vol. PP, no. 99, pp. 1, Oct 2014.

[109] X. Zhou, J. Ma, Y. Li, Y. H. Kwon, A. C. K. Soong, G. Zhao, "Probability-based transmit power control for dynamic spectrum access", *Proc. DySPAN'08*, pp. 1-5, Chicago, USA, Oct. 2008.

[110] X. Zhou, J. Ma, Y. Li, Y. H. Kwon, A. C. K. Soong, "Probability-based optimization of inter-sensing duration and power control in cognitive radio", *IEEE Transactions on Wireless Communications*, vol. 8, pp. 4922 - 4927, Oct. 2009.

[111] Z. Wang, S. Salous, "Spectrum occupancy statistics and time series models for cognitive radio", *Journal of Signal Processing Systems*, vol. 62, Feb. 2011.

[112] Ghosh, S. Pagadarai, D. P. Agrawal, A. M. Wyglinski, "A framework for statistical wireless spectrum occupancy modeling", *IEEE Transactions on Wireless Communications*, vol. 9, No. 1, Jan 2010.

[113] C. Rudin, K. L. Wagstaff, "Machine learning for science and society", *Springer Journal on Machine Learning*, vol. 95, no. 1, pp. 1-9, Nov 2013.

[114] K. W. Choi, E. Hossain, D. I. Kin, "Cooperative spectrum sensing under a random geometric primary user network model", *IEEE Transaction on Wireless Communications*, vol. 10, no. 6, June 2011.

[115] K. M. Thilina, K. W. Choi, N. Saquib, and E. Hossain, "Machine learning techniques for cooperative spectrum sensing in cognitive radio networks", *IEEE Journal on Selected Areas in Communications*, vol. 31 , no. 11, pp. 2209 -2221, Nov 2013.

[116] D. Willkomm, S. Machiraju, J. Bolot, A. Wolisz, "Primary users in cellular networks: A large-scale measurement study", *Proc. DySPAN' 08*, pp. 1-11, 2008.

[117] V. K. Tumuluru, P. Wang, D. Niyato, "Channel status prediction for cognitive radio networks", *Wiley Wireless Communications and Mobile Computing*, vol. 12, no. 10, pp. 862-874, July 2012.

[118] S. Pagadarai and A. M. Wyglinski, "A linear mixed-effects model of wireless spectrum occupancy", *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, no. 203178, Aug. 2010.

[119] A. J. Petain, "Maximizing the utility of radio spectrum: broadband spectrum measurements and occupancy model for use by cognitive radio", *Ph.D. dissertation*, Georgia Institute of Technology, Atlanta, GA, USA, 2005.

[120] RFeye evaluation system, Available online: `https://uk.crfs.com/en/products/nodes/evaluation-system/`.

[121] M. Zhang, B. Li, S. Jiang, "Call blocking probabaility and packet delay in cognitive radio networks" , *Proc. WiCOM' 12*, Sept 2012.

[122] A. Y. N, M. I. Jordan, "On discriminative vs generative classifiers: A comparison of logistic regression and Nave Bayes", *Advances in Neural Information Processing Systems*, pp. 841-848, 2002.

[123] L. Rokach, O. Maimon, "Decision Trees", *Data Mining and Knowledge Discovery Handbook*, Springer Publisher, 2nd ed, 1285, 2010.

[124] F. Berzal, J. C. Cubero, F. Cuenca, M. J. Bautista, "On the quest for easy o understand splitting rules", *Elsevier Data and Knowledge Engineering*, Vol. 44, no. 1, pp. 31-48, Jan 2003.

[125] P. Flach, "Machine Learning: The Art and Science of Algorithms that make sense of data", *Cambridge university press*, Sept 2012.

[126] P. Laskov, C. Gehl, S. Kruger, K.R. Muller, "Incremental support vector learning: analysis, implementation and applicatins", vol. 7, pp. 1909-1936, 2006.

[127] A. b. Hur, J. Weston, "A user's guide to support vector machines", *Data Mining Techniques for the Life Sciences Methods in Molecular Biology*, vol. 609, pp. 223-239, 2010.

[128] A. Gani, K. Mohammadi, S. Shamshirband, T. A. Altameem, "A combined method to estimate wind speed distribution based on integrating the support vector machine with fire fly algorithm", Wiley environmental progress and sustainable energy, Vol. 35, Issue 3, Oct 2015.

[129] L. Olatomiwa, S. Mekhilef, S. Shamshirband, "A support vector machine-firefly algorithm-based model for global solar radiation prediction", Vol. 115, pp. 632-644, May 2015.

[130] S. Ch, S. K. Sohani, D. Kumar, "A support vector machine-firefly algorithms based forecasting model to determine malaria transmission", Elsevier Journal on Neurocomputing, Vol. 129, pp. 279-288, April 2014.

[131] X. Yang, "Firefly algorithms for multimodal optimization", *LNCS 5792*, vol. 5792, pp. 169-178, 2009.

[132] S. Chatterjee, A. S. Hadi, "Simple linear regression", *Regression Analysis by Example*, Fourth Edition, 2006 John Wiley and Sons.

[133] G. A. Fink, "Markov Models for pattern Recognition: From theory to Application", *Springer Advances in Computer Vision and Pattern Recognition*, 2nd edition, 2014.

[134] D. Garrette, J. Baldridge, "Type-supervised hidden Markov models for part-of-speech tagging with incomplete tag dictionaries", *EMNLP - CoNLL'12*, pp. 821-831, Stroudsburg, PA, USA, 2012.

[135] D. Niyato, E. Hossain and A. Fallahi, "Sleep and wakeup strategies in solar powered wireless sensor/Mesh networks: performance analysis and optimization", *IEEE Transactions on mobile computing*, vol. 6, no. 2, Feb 2007.

[136] B. Medepally N. B. Mehta, C. R. Murthy, "Implications of energy profile and storage on energy harvesting sensor link performance, *Proc. IEEE Globecom09*, pp. 1-6, Honolulu, Dec 2009.

[137] J. Lei, R. Yates and L. Greenstein, A generic model for optimizing single-hop transmission policy of replenishable Sensors, *IEEE Transactions on Wireless Communications*, vol. 8, no. 2, Feb 2009.

[138] J. R. Piorno, C. Bergonzini, D. Atienza and T. S. Rosing, "Prediction and management in energy harvested wireless sensor nodes", *Proc. Wireless VITAE'09*, pp.6-10, 17-20 May 2009.

[139] C. Bergonzini, D. Brunelli, L. Benini, "Algorithms for harvested energy prediction in battery less wireless sensor networks", *Proc. IWASI'09*, pp.144-149, 25-26 June 2009.

[140] M. Pinuela, P. D. Mitcheson, S. Lucyszyn, "Ambient RF Energy Harvesting in Urban and Semi-Urban environments", *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, pp. 2715-2726, 2013.

[141] Y. Chen, H. S. Oh, "A Survey of measurement-based spectrum occupancy modelling for cognitive Radios", *IEEE Communications Surveys Tutorials*, vol. pp, no. 99, pp. 1, Oct 2014.

[142] P. Flach, "Machine Learning: The Art and Science of Algorithms that make sense of data", *Cambridge university press*, Sept 2012.

[143] L. Torgo, "Functional Models for regression tree leaves, *Proc. ICML*, 1997.

[144] M. Torlak, " Telecom. switching and transmission", UT Dallas, Available online: https://www.utdallas.edu/ tor-lak/courses/ee4367/lectures/lectureradio.pdf.

[145] K. Kotani, T. Ito, "High efficiency CMOS rectifier circuit with Self-Vth Cancellation and power regulation functions for UHF RFIDs", *Proc. ASSCC '07*, pp. 119-122, 2007.

[146] S. Scorcioni, L. Larcher, A. Bertacchini, "Optimised CMOS RF-DC Converters for remote wireless powering of RFID applications", *2012 IEEE International Conference on RFID*, pp. 47-53, 2012.

[147] E. Khansalee, Y. Zhao, E. Leelarasmee, K. Nuanyai, "A dual-Band rectifier for RF energy harvesting systems", Proc'. *ECTI-CON*, pp. 1-4, May 2014.

[148] H. Sun, Y. Guo, M. He and Z. Zhong, "A dual-band rectenna using broadband yagi antenna array for ambient RF power harvesting", *IEEE Antennas and Wireless propagation letters*, vol. 12, 2013.

[149] R. Nau, "Forecasting with moving averages", Fuqua school of Business, Duke University, Aug 2014.

[150] L. Mohjazi, M. Dianati, G. K. Karagiannidis, S. Muhaidat, and M. Al-Qutayri, "RF-powered cognitive radio networks: technical challenges and limitations", *IEEE Communications Magazine*, April 2015.

[151] S. Lee, R. Zhang and K. Huang, "Opportunistic wireless energy harvesting in cognitive radio networks", *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, September 2013.

[152] T. Nguyen, I. Koo, "Throughoput of primary user with cognitive radio function", *Proc'. ISCIT*, 2014.

[153] S. Park, H. Kim, D. Hong, "Cognitive radio networks with energy harvesting", *IEEE Transactions on Wireless Communication*, vol. 12, no. 3, March 2009.

[154] Z. Wang, Z. Chen, Y. Yao, B. Xia and H. Liu, "Wireless energy harvesting and information transfer in cognitive two-way relay networks", *Proc. Globecom*, Dec 2014.

[155] G. Zheng, Z. Ho, E. A. Jorswieck, B. Ottersten, "Information and energy cooperation in cognitive radio networks", *IEEE Transactions on Signal Processing*, vol. 62, no. 9, May 2014.

[156] E. Matskani, N. Chatzidiamantis, L. Georgiadis, I. Koutsopoulos and L. Tassiulas, "The mutual benfits of primary-secondary user cooperation in wireless cognitive networks", *Proc'. WiOpt*, 2014.

[157] X. Chen, Z. Zhang, H. Chen and H. Zhang, "Enhancing wireless information and power transfer by exploiting multi-antenna techniques", *IEEE Communications Magazine*, April 2015.

[158] H. Ju, R. Zhang, "Throughput maximization in wireless powered communication networks", *IEEE Transactions on wireless communication*, vol. 13, no. 1, Jan 2014.

[159] F. F. Digham, M. S. Alouini and M. K. Simon, "On the energy detection of unknown signals over fading channels", *IEEE Transaction on Wireless Communications*, vol. 44, no. 1, Jan 2007.

[160] Jeff Hawkins, "Hierarchical temporal memory including cortical learning algorithms", *White paper published by Numenta*, Version 0.2.1, Sept 2011.

[161] Numenta, "Leading the new era of machine intelligence", Available online: http://numenta.com/.

[162] W. Chung, S. Park, S.Lim, D. Hong, "Spectrum sensing optimization for energy-harvesting cognitive radio systems", *IEEE Transaction on Wireless Communications*, vol. 13, no. 5, May 2014.