Spectrum Surveying for Dynamic Spectrum Access Networks

By

Dinesh Datla
B.E., Electronics and Communication Engineering,
University of Madras, India, 2004

Submitted to the Department of Electrical Engineering and Computer Science and the Faculty of the Graduate School of the University of Kansas in partial fulfillment of the requirements for the degree of Master's of Science

Thesis Committee:

Dr. Gary J. Minden: Chairperson

Dr. Joseph B. Evans

Dr. Alexander M. Wyglinski

Date Defended: 01/17/2007

2007/02/02

The Thesis Committee for Dinesh Datla certifies
That this is the approved version of the following thesis:

Spectrum Surveying for Dynamic Spectrum Access Networks

Committee:

Chairperson

Date Approved

Abstract

The current *command-and-control* regulatory structure for licensing the RF spectrum has been unable to cope with the growing demand for spectrum. This has given rise to an 'artificial scarcity' of usable spectrum. Numerous studies have thus begun to examine how the licensed spectrum is actually used, with the goal of not only reforming the spectrum licensing regime but also opening certain underutilized spectrum to unlicensed secondary usage. The technologies such as frequency agile radios and dynamic spectrum access networks can enable unlicensed users to access the underutilized spectrum in a manner that is transparent to the licensed users.

A thorough understanding of the spectrum utilization and interference issues can be helpful in devising solutions to maximize the spectrum utilization and assist in the design of secondary spectrum access technologies. A spectrum survey can be conducted in order to collect data on spectrum utilization and process the data in order to extract useful information about the spectrum activity.

In this thesis, a generic framework for spectrum surveying has been proposed. This framework outlines the procedures and techniques for the collection, analysis and modeling of spectrum measurements. Techniques have been proposed to perform analysis and extract important parameters of spectrum occupancy. The proposed framework introduces standardization to spectrum surveying as well as automation to the processing of the measurements. The proposed processing techniques have been tested on spectrum measurements collected from the FM band, TV bands, cellular band, and paging band and the results are also presented.

The television spectrum is being underutilized, making it a prime candidate for dynamic spectrum access. Nevertheless, the quality of this spectrum for enabling secondary transmissions has never been assessed. The proposed spectrum survey framework has been implemented and used to examine how non-ideal transmission properties of television broadcasts can potentially impair the performance of secondary transmissions. The framework has also been incorporated into a novel wideband spectrum sensing architecture. The thesis presents the results of the feasibility study as well as simulation results obtained for the proposed spectrum sensing architecture.

Acknowledgments

I am greatly indebted to Dr. Gary Minden for giving me the opportunity to work on future-generation wireless technology in the NSF-sponsored NRNRT project. I have learnt several lessons on conducting research while working with him. I consider it a priviledge to have distinguished professor Dr. Joseph Evans in my committee. Dr. Minden and Dr. Evans have been strong pillars of support for the NRNRT research team. I have interacted several times with Dr. Alexander Wyglinski in my research. I specially thank him for his selfless efforts in editing every word in my thesis and papers before providing me feedback on improving my work. My research work was supported by NSF grants ANI-0230786 and ANI-0335272.

Ted Weidling and Rory Petty are the brains behind the spectrum miner software. Without this software, my work would not have been possible. They have also provided me Matlab tools to retrieve the data from the spectrum repository. We have also collaborated on developing the hypothesis testing algorithm (Chapter 4)and the planning of the feasibility study (see Chapter 5).

I thank Rakesh Rajbanshi for his assistance while working on the spectrum sensing paper. I have used his OFDM simulations to perform the feasibility study. My work on the NC-OFDM simulation is built upon this OFDM simulation. I thank Daniel DePardo for his feedback on the results of the feasibility study. The second phase of the study which is not presented in this thesis was completely performed by him. My discussions with Jordan Guffey on OFDM were also very helpful. Leon Searl, Timothy Newman, Brett Barker, and Qi Chen have provided useful tips to improve my thesis presentation. Michael Hulet, Wesley Mason and the system administrators at ITTC have been very helpful and prompt at crucial times. I thank Annie Francis, Paula Conlin, Robin Hinman, and other administrative staff at ITTC for their help.

I have a wonderful set of roommates (Dileep, Kiran, Madhu) and friends at KU who are always there to help me. I thank my family for making my education in USA possible.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The radio frequency (RF) spectrum is a significant and reusable resource [1] which can be described as a virtual hyperspace, called the *electrospace* [2]. The primary dimensions of the spectrum are: frequency, time, spatial extent, and signal format. The spectrum is an integral part of any wireless system, and a thorough understanding of its behavior is important for effective wireless communications. Fig. 1 shows the transmit and receive spectrum masks that regulate the wireless system's access to the spectrum. The wireless channel acts as a filter that influences the signal that is transmitted through it.

There are two aspects to spectrum management [1]. The technical aspect of spectrum management is concerned with the technology, and the physical world phenomena that affect the spectrum utilization. The policy aspect takes into account the economic and political factors that affect the spectrum market. In the United States, the *Federal Communications Commission* (FCC) regulates the spectrum access for commercial purposes. Under the legacy *command-and-control* regulatory structure for spectrum management, the spectrum is divided into several segments and each segment is allocated for a specific wireless service. A

Figure 1.1. Wireless communication block diagram showing RF spectrum and wireless channel.

licensed user has exclusive rights within a specified geographical area to access a fixed number of frequency channels in a segment [3].

1.1 Research Motivation

The current spectrum allocation policy allows the use of low cost standardized communication equipment and it can ensure that there will be no conflicts in the access to the licensed spectrum [4]. However, it possesses some serious drawbacks that are of concern for future spectrum management.

The entire spectrum has been fully allocated [5], leaving very little space for additional wireless services. In addition, the current spectrum policy makes it difficult for the rapid deployment of new services, which is particularly crucial for emergency services [6].

Moreover, it has been found that the spectrum is underutilized temporally, spatially and spectrally [3, 7]. This has lead to the creation of large portions of underutilized and vacant spectrum, which are termed as *spectrum white spaces* or *spectrum holes*. For instance, the dynamic nature of the spectrum utilization by mobile telephony services has resulted in an inefficient usage of the spectrum temporally [3]. In the television and FM broadcast bands, *buffer spaces* have been created in order to maintain safe distances between broadcasting stations operating on the same frequency channel. In addition, *guard bands* have been assigned between adjacent station frequency channels in order to avoid adjacent

channel interference.

Furthermore, the demand for wireless services has been steadily increasing due to the growing need for wireless broadband connectivity. There has been an increase in the number of users of wireless services, as well as new wireless services that are constantly evolving. Moreover, the requirements of federal agencies and emergency services place high and uncompromising constraints on the spectrum [1]. The current regulatory structure does not possess the flexibility to allow the dynamic reuse of the licensed spectrum even when it is idle as well as fast deployment of new wireless services. This rigid spectrum management system is the main cause for a potential *spectrum scarcity* in the near future.

1.2 Solutions to the Spectrum Scarcity Problem

There has been substantial research efforts aimed at improving the spectrum utilization. New services can be accommodated by redefining the spectrum that has been assigned to existing services that utilize the spectrum sparsely [7]. However, in the current spectrum regime, any potential reallocation of the spectrum can have many political and commercial consequences [8].

In June 2002, the FCC commissioned the Spectrum Policy Task Force (SPTF), which is a body that makes recommendations on reforming the spectrum policy. In its final report [9], the SPTF has suggested various methods to improve the spectrum utilization, such as exploitation of the spectrum along all its dimensions, reuse of the underutilized spectrum, and transformation of the current command-and-control regime into a more flexible market-based system.

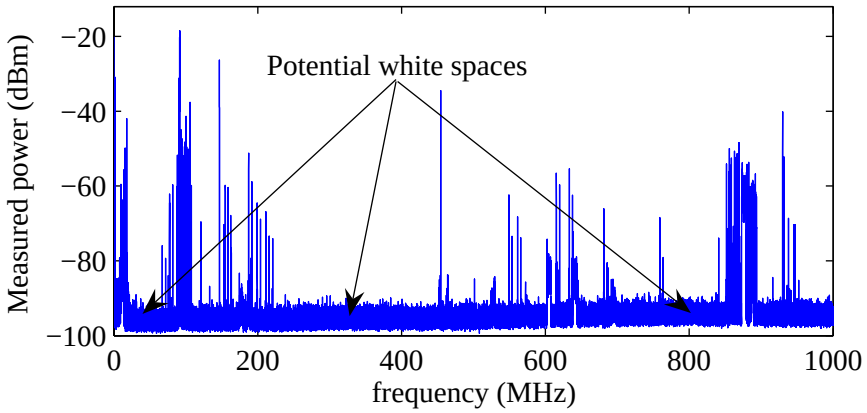

Figure 1.2. Underutilized portions of spectrum (9 kHz - 1 GHz band).

1.2.1 Unlicensed Bands

The FCC has allocated several portions of spectrum for unlicensed usage [1] and has been encouraging the use of these bands for the demonstration of innovative spectrum sharing techniques [8]. This can improve the spectrum utilization as well as allow any FCC-approved wireless service to be deployed without pre-allocation of spectrum. Spectrum sharing in the unlicensed spectrum can be achieved with the ultra wide band (UWB) radio technology that operates in the 3.1-10.6 GHz frequency range [11]. However, since the unlicensed users have to coexist with the licensed users, there are many restrictions on the transmit power [10].

1.2.2 Secondary Usage of Underutilized Spectrum

Past studies have shown that the spectrum utilization can be significantly improved by the reuse of spectrum white spaces [7] (see fig. 1.2). The white spaces comprising of buffer spaces, guard bands, unlicensed bands, and sparsely utilized licensed spectrum are potential candidates for secondary spectrum usage.

[1] For a list of unlicensed bands refer to [10].

A secondary (i.e. unlicensed) user can access the underutilized spectrum in a manner that is transparent to the primary (i.e. licensed) user.

The secondary usage of spectrum white spaces requires a technology that can reliably operate over a large portion of the spectrum including the licensed and unlicensed bands. A frequency agile radio is one such technology that possesses the flexibility to change its transmit parameters, such as its frequency of operation, bandwidth and transmit power, as well as transmit in any frequency channel which has been identified as a potential white space. A network of these radios can form a *dynamic spectrum access* (DSA) *network*, which can optimize the secondary usage of the spectrum. In addition, DSA networks can enable rapid deployment of new wireless services [6].

There has been regulatory and legislative activity that could allow unlicensed devices to access TV band *white space* on a per market basis. In May 2004, the FCC released the Notice of Proposed Rulemaking (NPRM) [12] allowing unlicensed devices to utilize unused spectrum in the TV band. In addition, an IEEE standard for DSA networks has also been proposed [13].

1.2.3 NRNRT Project Overview

At the University of Kansas, the NSF-sponsored National Radio Network Research Testbed (NRNRT) is being developed to support the research and development of agile radios and DSA networks [1]. It is also hoped that the results from the testbed will provide significant input to debates on future spectrum management issues.

The testbed includes a field deployable measurement and evaluation system for long term spectrum data collection, and an experiment facility. The RF data,

which comprises of the RF spectrum utilization measurements and wireless channel propagation measurements, can help in the design of radio technology and medium access protocols for wireless systems. The experiment facility can be used to test prototypes of new radios and wireless networks in a real world situation. The RF data and test measurements taken using this experiment facility can be incorporated into an accurate emulation/simulation model, which will be useful for improved analysis of the new wireless systems.

1.3 Scope and Contributions of Thesis

The main objective of this thesis is: *To conceptualize and implement a generic framework for spectrum surveying that would aid in the research on DSA networks.* Our study of the spectrum will yield important information that will be useful in devising techniques for improving the utilization of the spectrum. In the past, several spectrum surveys [3, 7, 14] have been performed. However, they lacked a formalized structure that can enable collaborations on spectrum surveying and in addition, the processing of the data was not completely automated. This thesis contributes to the research on spectrum surveying by presenting the following:

- A formalized framework for spectrum surveying,
- Statistical methods to process the spectrum survey data,
- Tools to analyze the spectrum data and extract information on spectrum utilization,
- A comprehensive model for the spectrum measurements, and

- Application of the framework to research and implementation of DSA networks.

A novel framework to collect, store, and analyze spectrum utilization data has been conceptualized and implemented. The proposed framework introduces a formal structure to the different stages of a spectrum survey thereby introducing standardization to spectrum surveying.

Techniques for processing and analyzing the spectrum data have been proposed and the efficacy of these techniques has been verified with the help of real world spectrum data. In contrast to the previous spectrum studies, the proposed statistical methods introduce automation to the analysis of the spectrum survey data. A set of parameters of spectrum utilization have been identified, and tools have been developed to extract these features from the processed data. Furthermore, a comprehensive model for the spectrum measurements has been presented.

The thesis also presents the application of this framework to assess the feasibility of secondary usage of spectrum and a novel architecture derived from this framework that can enable efficient spectrum surveying in future communication systems.

1.4 Organization of Thesis

The rest of this thesis is organized as follows:
Chapter 2 begins with an introduction to DSA. The need for the study of the spectrum utilization is stressed along with the possible applications of spectrum surveying to research activities in DSA networks. The chapter also provides a summary of the past research works on spectrum surveying and other related aspects such as processing of the spectrum measurements.

In Chapter 3, the proposed framework for the study of the spectrum utilization is presented along with a description of the various components of the framework including the modeling of the spectrum data. The chapter also describes the preliminary efforts on the processing of spectrum data. Chapter 4 discusses the challenges associated with the processing of spectrum data. The techniques that have been proposed to counter these challenges are presented followed by the results of their performance evaluation.

Chapter 5 describes two applications of the proposed framework. The framework has been applied for assessing the feasibility of secondary usage of the television spectrum. The framework has also been incorporated into a novel architecture for monitoring the spectrum activity in dynamic spectrum access networks.

Chapter 6 presents the conclusions of the thesis and provides directions for future work.

Chapter 2

Background Literature Review

This chapter builds upon the last chapter by providing an overview of DSA networks and spectrum surveying. The secondary access of underutilized spectrum by *Opportunistic Spectrum Access* (OSA) or *Dynamic Spectrum Access* (DSA) is discussed in Section 2.1. This discussion leads on to the importance of spectrum surveys to the research in DSA networks. Section 2.2 provides an introduction to spectrum surveying along with its applications. In Section 2.3, the existing techniques for processing the spectrum data are presented. The different parameters with which the spectrum utilization can be characterized are described in Section 2.4.

2.1 Dynamic Spectrum Access by Frequency Agile Radios

Past studies [3, 7] have assessed the feasibility of secondary usage of spectrum white spaces by a frequency agile operation, and it has been shown that this approach can improve the spectrum utilization. However, there are several challenges to implement this approach. The spectrum operating conditions

including the spectrum utilization and the spectrum regulations can vary as a function of frequency, time of the day, azimuth, polarization, and geographical location. These changes have to be considered for effective frequency agile radio operation [4]. Furthermore, the secondary user might encounter different types of primary users, such as television broadcasters, terrestrial microwave services, and cellular mobile services. Each of these primary users possess different transmission characteristics, such as duty cycle, transmit power, and bandwidth. Moreover, some primary users transmit in bursts, with the transmission characteristics varying with time [10].

To counter these challenges and operate in a manner that is transparent to the existing primary users, the radio must be able to adapt to the varying operating conditions and also dynamically detect the unused spectrum before accessing it. To permit this dynamic spectrum access by secondary users, changes have to be made to the current spectrum management framework in order to allow dynamic and flexible spectrum access management [3]. In addition, we need reliable technology to enable DSA.

These challenges have motivated research on radios that have the ability to learn and adapt to the current spectrum conditions. The DARPA XG program [1] has been pursuing research on a novel approach wherein frequency agile radios can perform dynamic reuse of the underutilized spectrum. This approach basically involves the following stages [6]:(i) The radio dynamically senses its spectrum environment. Spectrum sensing is the process of sampling the channel utilization in order to collect data that would help in identifying potential opportunities for

[1]The Next Generation (XG) program has been initiated by the United States Defense Advanced Research Program Agency (DARPA) and it drives the research and development of next generation communication systems.

secondary usage of the spectrum [15]. Using the spectrum sensing data, the radio detects the presence of primary signals and other secondary signals, and then characterizes the spectrum utilization. (ii) The radio uses the information gathered by spectrum sensing and characterization of spectrum to identify the frequency channels that are suitable for secondary usage. (iii) Using the spectrum sensing and channel sounding data, the radio adapts to the current wireless environment and transmits in these frequencies in a manner that is transparent to the primary users. This OSA approach has several benefits: It has the potential to maximize the spectrum utilization, it allows the secondary user to coexist with the primary spectrum users in a transparent manner, and it enables secondary access in scenarios where there are discontinuities in the available spectrum and the spectrum utilization varies dynamically.

While a frequency agile radio can dynamically detect spectrum opportunities and transmit in spectrum holes located anywhere in the spectrum [6], a *cognitive radio* is a radio platform that is both agile and capable of adapting to the current wireless environment so that its communication does not cause any harmful interference to the primary users of the spectrum [16]. In a dynamic environment, cognitive radios can also make intelligent decisions on the transmit parameters to set in order to optimize the performance of the communication system [7]. Besides cognitive radios, DSA is also being considered for multiband OFDM systems [11] and carrier sense multiple access/collision avoidance (CSMA/CA) networks [17].

Although DSA has several benefits, it is not trivial to implement it. Dynamic sensing requires searching for spectrum holes on a instant-by-instant basis so that the interference to the primary users is limited. The challenges faced by real-time sensing include degraded channel environments [18], detection of weak signals, and

the presence of intermodulation products [10]. Prior knowledge of the spectrum occupancy and the characteristics of the primary signals can greatly help in executing DSA in an effective manner. This knowledge can be gained by surveying the spectrum activity.

2.2 Spectrum Surveying

Precise information on the spectrum behavior cannot be inferred directly from licensing information since it varies with several factors. Hence, there is a need for real world spectrum data in order to characterize its behavior.

Spectrum surveying involves the long term collection of spectrum data by spectrum sensing over a wide range of frequencies. Information about the spectrum activity can be extracted by analyzing the data.

2.2.1 Applications of Spectrum Surveying

Table 2.1 provides a list of previously conducted spectrum studies including the studies conducted at the Radio Spectrum Engineering Lab of Georgia Institute of Technology (GIT) [3, 19], the Institute of Telecommunication Sciences (ITS) in USA [14], the Mobile Portable Radio Group (MPRG) at Virginia Tech [7], and the Shared Spectrum Company (Virginia, USA) [20].

The spectrum data can be used to extract patterns of transmission activity, as well as for interference analysis [19]. A thorough understanding of the spectrum behavior and interference issues can be helpful in devising solutions to maximize the spectrum utilization and assist in the design of secondary spectrum access technologies [21]. Spectrum study 1 in Table 2.1 has been performed [14] to assess the usage of the land mobile radio bands (138-174 MHz) by federal agencies.

Table 2.1. Previously Conducted Spectrum Studies

Name	Frequencies	Purpose of Survey
1. ITS	LMR bands	Improve usage of LMR bands, prediction of usage.
2. MPRG	30-300 MHz	Assess feasibility of DSA.
3. SS Co.	54-3000 MHz	Quantify spectrum occupancy.
4. GIT	0.4-7.2 GHz	Determine spectrum occupancy, interference to radiometric services.

These studies were performed to determine how the spectrum utilization can be improved, and to provide a basis for prediction of the future usage of these bands.

The amount of unused spectrum can be quantified in order to assess the feasibility of broadband communication through cognitive radio operation in spectrum white spaces [7] (spectrum study 2 and 3 of Table 2.1). The spectrum utilization can be quantified using a metric, called *spectrum occupancy*, which is defined as "the probability that a signal is detected above a certain threshold power level" [7].

Resolving the spectrum usage along all the dimensions of the spectrum will provide important information that will help in exploiting the spectrum efficiently [21]. In the real time scenario, the knowledge obtained from spectrum surveying can be input to DSA networks in order to help them in identifying potential opportunities for secondary access. In addition, multi-dimensional spectrum data can be used to identify the characteristics of the primary transmitters such as polarization type and azimuth location [3].

Furthermore, the results of a spectrum survey can also help in interference avoidance (spectrum study 4). The 6.75 - 7.1 GHz band is primarily used to provide fixed microwave services, while radio astronomy studies are also performed in this band. The interference from the primary users can potentially hinder the astronomy studies [19, 22]. In such scenarios, the temporal characteristics of the

primary users can be studied in order to identify periodic spectrum users, and methods can be developed that allow the passive services to dynamically use the spectrum when the primary user is absent.

Besides collecting the spectrum data, a spectrum study mainly involves processing the data in order to detect the presence of signals, and then analyzing the data in order to characterize the spectrum occupancy.

2.3 Processing of Spectrum Survey Data

Improvements in hardware can result in the effective measurement of the spectrum utilization and the detection of signals. However, defects in the measurement data introduced by broadband impulse noise, intermodulation products, and system noise can be conveniently removed by post-measurement data processing [14]. Moreover, processing of the data can improve the sensitivity of the RF front end by providing a processing gain in the detectable signal-to-noise ratio (SNR), and also help in extracting signal features which can be used for the detection of primary signals [23].

2.3.1 Refining of Spectrum Data

Several techniques were proposed in reference [14] to refine the data. The broadband impulse noise affects measurements only at the instant when it occurs. This fact can be used to identify those sweeps of data that are contaminated with it. In the case of a receiver that determines the power at all frequencies simultaneously, the impulse noise affects all the frequencies equally. Sweeps of data with a comparatively higher average noise level can be identified and removed from consideration for further processing [14]. Alternatively, several sweeps of data

can be averaged in order to reduce the average noise power.

In the presence of a strong signal at the input to the RF receiver, when the local oscillator (LO) output is mixed with the strong signal, the LO noise sidebands produce an adjacent-channel response that adds to the signals occurring adjacent to the strong signal. In order to remove the effect of these sidebands, a typical noise sideband response can be measured, and these power levels can be subtracted from the measured power spectrum [14]. In a similar manner, the intermodulation products can be predicted and then subtracted from the measured data. By removing the intermodulation products, the effects of the LO noise sidebands, and the impulse noise, the occurrence of signal-like noise spikes in the measured data is greatly reduced, thereby reducing false alarms.

2.3.2 Detection of Primary Signals

In the literature, the signal detection techniques have been classified into: match filtering, energy detection, and cyclostationary feature detection [23]. Match filtering is achieved by correlating the received unknown signal with a replica of the signal to be detected. While this method performs well even under low SNR conditions, this method requires *a priori* information about the signal characteristics and the type of signal to be expected in the band of interest. This technique has been applied for the feature detection of weak television signals [24].

Cyclostationary feature detection gives the best performance among the three. However, it is a complicated approach that requires both phase and magnitude information [23]. If only power measurements of the spectrum utilization are available, then energy detection is the optimal detection approach [25]. Due to its simplicity and relevance to the processing of power measurements, energy de-

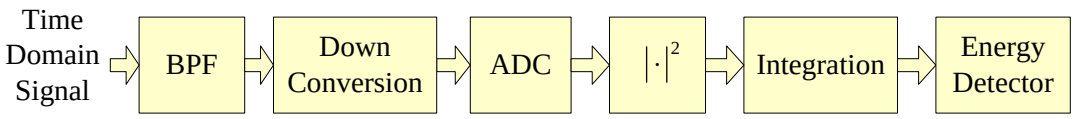

Figure 2.1. Block diagram of energy detection.

tection has been a preferred approach for several past spectrum studies [7, 10, 14]. Experiments have been conducted on the performance of energy detectors [26, 27]. However, it possesses some drawbacks [23]: It cannot distinguish between signal and interference, and it cannot be used to detect signals that occur below the noise level, such as signals from distant transmitters and spread spectrum signals.

2.3.3 Energy Detection

Fig. 2.1 shows the block diagram of energy detection that is implemented with a swept-type measurement system. The received time domain signal is filtered, down-converted, digitized, squared and then integrated over a certain period of time, in order to obtain the average power at the filter's center frequency. The signals can be detected based on a decision threshold, i.e. power measurements above this threshold are identified as signals.

The decision threshold can be estimated in several different ways such as:

1. Empirical analysis of data [10, 20],
2. Computation of threshold from system properties such as noise floor [7, 10],
3. Using *a priori* knowledge of statistics of noise [14, 17], and
4. Estimation of threshold directly from data.

The simplest approach to determining the threshold is via an empirical analysis wherein the collected measurements can be visually inspected. In the second

approach, the decision threshold can be computed as a function of the sensitivity and noise figure of the system.

By measuring a vacant channel that is free from external interference, samples of the system and ambient noise can be collected, and be used to compute an inverse cumulative density function (CDF). If S is a random variable representing the measured noise power, the inverse CDF of S can be defined as $ICDF(S, X) = P(S \geq X)$. If X is the threshold then the corresponding false alarm probability will be $ICDF(S, X)$. From the inverse CDF, the threshold can be chosen for a fixed false alarm probability. For the specific case of Gaussian distribution of noise, the threshold can be determined from the following expressions:

$$P_{FA} = \frac{1}{\sqrt{2\pi}\sigma_N} \int_T^\infty e^{-(x-\mu_N)^2/2\sigma_N^2}\, dx, \tag{2.1}$$

$$P_d = \frac{1}{\sqrt{2\pi}\sigma_{SN}} \int_T^\infty e^{-(x-\mu_{SN})^2/2\sigma_{SN}^2}\, dx, \tag{2.2}$$

where x represents the power levels of the measurements, P_{FA} and P_d are, respectively, the probabilities of occurrence of false alarms and miss detection, σ_N and σ_{SN} are the standard deviations of the noise and signal (signal samples represent sum of signal and noise) samples, and μ_N and μ_{SN} are the mean power levels of the noise and signal samples.

In all these methods, by averaging the data the noise variance is reduced and for the same false alarm rate the threshold can be set lower, resulting in an increase in the probability of detection of weak signals. The drawbacks of these threshold estimation techniques are: (i) The threshold estimated is specific to the receiver, and hence they fail to detect the presence of signals that occur below the receiver's noise floor, (ii) These methods require *a priori* knowledge of the noise

Table 2.2. Characteristics of Spectrum Utilization

Spectrum Parameter	Information Provided
Temporal-spectral statistics on occupancy	Potential bandwidth-time capacity, temporal and spectral agility required, spectrum models
Primary signal features	Algorithms for detection/identification of primary users, to distinguish between primary and secondary users, minimum detection rate.
Ambient noise power characteristics	Interference and noise levels, SNR, sensitivity, detection threshold

statistics, and (iii) They fail to perform well in the presence of noise power that varies throughout the frequency band of interest.

While the above methods require prior knowledge of the noise statistics, in this thesis, we address the problem of estimating the threshold directly from the data itself without requiring any *a priori* knowledge.

2.4 Characterization of Spectrum Behavior

The information gathered by spectrum surveying can be used to characterize the short term as well as the long term occupancy of the spectrum. The activity in the spectrum can be described in terms of certain parameters which can be computed from the spectrum measurements. Broadly, there are two kinds of parameters: the statistics of white space availability, and the parameters describing the other aspects of spectrum activity including the features of the signals occupying the spectrum [28]. Table 2.2 presents a list of these parameters and the information that can be gained from knowledge of these parameters [28].

The unoccupied spectrum can be characterized in terms of temporal statistics of the channel availability and statistics on the available bandwidth. These

statistics indicate the potential capacity (bandwidth-time product) that can be supported by the underutilized spectrum for secondary usage [7]. These statistics can also be used to deduce the minimum agility in frequency and time that is required by the radios in order to access the target spectrum in a DSA manner [28]. If the bandwidth availability is represented by the random variable S_B, the probability distribution of the availability of contiguous bandwidth, $P(S_B = k)$, can be computed. Similarly, if the periods of inactivity for a particular frequency channel is represented by random variable S_t, the inverse CDF of channel availability, $ICDF(S_t, t) = P(S_t \geq t)$, can be computed. Based on these statistics, the future availability of the spectrum white spaces can be predicted. For instance, if the channel has been observed to be inactive for t_n time units, then the probability that the channel will be inactive for the next t_m time units is computed as [17]:

$$P_{n,m} = \frac{ICDF(S_t, t_n + t_m)}{P(S_t = t_n)}. \tag{2.3}$$

The statistics of spectrum occupancy can also be incorporated into spectrum models which have several applications: (i) For long term and short term forecasting of spectrum occupancy, (ii) To provide a summary of the spectrum behavior, and (iii) Aid in the design of protocols for DSA. There are two popular spectrum occupancy models: The Laycock-Gott model and the Markov model. While the Laycock-Gott model requires extensive procedures for fitting the data into a model [29, 30], the markov model is relatively simpler and requires the computation of channel availability probabilities [31].

The features of the primary signal such as signal bandwidth, transmission patterns, duty cycle, and parameters describing the transmitters which utilize the spectrum, can be extracted from the spectrum data. These parameters can be used

in the feature detection of primary signals. The duty cycle and the transceiver mobility specify the minimum rate at which the channel needs to be sensed in order to detect the primary signal. On comparing the mean and maximum power levels among the measurements collected in a certain frequency channel, we can infer the following [7]: If they are nearly equal, we can assume that there is no significant fading of the signal power in that channel. The absence of fading may also mean that the transceiver is not mobile. If the mean is significantly smaller than the maximum, it can be deduced that the channel is being utilized intermittently.

The range of the measurements specifies the minimum dynamic range required by the agile radio's RF front-end [7]. Measurements of the noise power in the underutilized spectrum can be used to determine the transmit power to be maintained for $Z\%$ of the time such that the transmit SNR is fixed at a certain value for the satisfactory operation of a wireless communication system [32]. Since different measuring instruments have different receiver properties, such as sensitivity, noise floor, and spectral resolution, the characterization of the spectrum utilization is a function of the device and the decision threshold used.

2.5 Chapter Summary

This chapter has provided an introduction to OSA. The challenges to OSA have been discussed followed by an introduction to cognitive radios. The applications of spectrum surveying to OSA have been discussed. This was followed by a discussion on the previously used techniques for the processing of the spectrum data. The parameters of spectrum occupancy have been presented along with a brief introduction to the modeling of spectrum occupancy.

Chapter 3

A Framework for Study of Spectrum Utilization

DSA networks and spectrum policy reforms rely on accurate spectrum utilization statistics which can be computed from the data collected by spectrum surveying. The different stages involved in our spectrum survey are: First, the spectrum activity in the target spectrum is captured by collecting spectrum measurements, with the measurement data archived in a suitable format. Second, the spectrum data is processed in order to distinguish the signal and noise measurements. The data is then analyzed in order to extract the characteristics of the spectrum utilization. Third, the occupancy in the target spectrum is modeled for further analysis.

In this chapter, we present a framework that can formalize the spectrum survey by providing a layout for the different stages of the survey and defining the procedures involved in each stage in a convenient mathematical form.

The *Spectrum Survey Framework* (SSF) is expected to aid in the research of cognitive radio networks and provide necessary statistics helpful for debates on

spectrum policy. The SSF enables an automated and efficient approach for performing a spectrum survey. It also facilitates collaborations on collecting and analyzing spectrum measurements. This overcomes the limitations of localized spectrum utilization studies. The proposed framework provides standardized procedures for recording, storing and sharing the measurements. Such standardization will provide a uniform basis for the collaborative study of spectrum utilization.

This chapter also presents the proposed model for the spectrum measurements. This model considers the different random processes that influence the spectrum utilization, including noise, signal power, and the dynamics of the spectrum occupancy. The motivation behind this model is that the different aspects of the spectrum environment, such as signal-to-noise-ratio (SNR), are important in addition to the statistics of channel occupancy.

An introduction to the framework is provided in Section 3.1 followed by a description of each component of the framework. In Section 3.2, we define the metrics to evaluate the efficacy of the spectrum survey results. Section 3.3 presents our implementation of the SSF along with the results for some preliminary data processing algorithms.

3.1 Introduction to Spectrum Survey Framework

Fig. 3.1 shows the high level structure of SSF[1]. The energy (power) in the spectrum is measured by the *measurement subsystem*. The *data management* block formats the measurement data and transfers it to a storage device or to a centralized database where it can be archived for processing and analysis. It enables collaborations on the study of the spectrum by supporting distributed

[1]This framework which was developed by Dinesh Datla has been published in reference [33] and has been reproduced here with the written permission of all the authors.

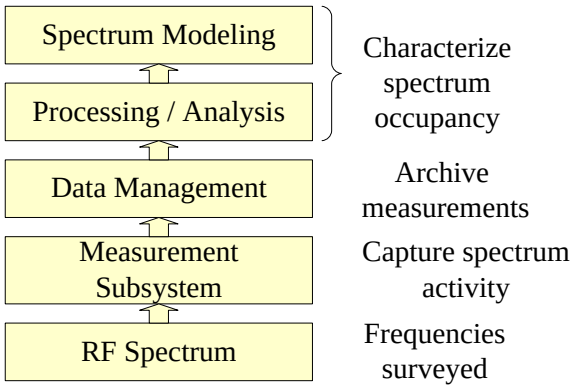

Figure 3.1. High level structure of SSF representing the different phases of spectrum surveying.

and simultaneous data collection, and also providing a mechanism to easily share data among researchers. For instance, the archived data can be published to the research community via the world wide web.

In the *processing and analysis* block, statistical methods are used to process the data. The processed data is analyzed in order to retrieve information about the spectrum utilization, which in turn is used to model the spectrum occupancy in the *modeling* block. Together, the processing and analysis block, as well as the modeling block, perform the characterization of the spectrum.

3.1.1 Measurement Subsystem and Parameters of Spectrum Sensing

A swept-type spectrum analyzer can be used to measure the power distribution in a certain frequency band. The term *measurement test* is used to refer to an experiment designed to collect spectrum measurement data. During a measurement test, the measuring instrument scans across the *sweep bandwidth* B_s, whose limits are specified by F_{start} and F_{stop}. The sweep across B_s does not occur in a continuous manner but in steps of B_r, where B_r is the bandwidth resolution. Every time B_s is scanned, a bandpass filter of bandwidth B_d is stepped in frequency

increments [29] of B_r, where B_d is referred to as the *binwidth*. At each frequency step, the filter bandwidth is centered about channel center frequency f_i, the time domain signal is passed through the filter, and the measurements are collected for a certain period of time. This period of time is referred to as the *dwell time*, T_d. The average of the measurements collected across B_d over the time T_d is stored as the power at that frequency channel. Thus, the sensing mechanism cannot resolve the power of frequencies within the resolution bandwidth. For statistical independence of measurements collected from two adjacent channels, it should be ensured that $B_d \leq B_r$ is satisfied.

The power across all the frequencies in the band B_s is not measured simultaneously, but with a certain delay. The *sweep time* T_s is the total time taken to complete a single sweep over the bandwidth B_s. The sweep time depends on several factors. In addition to the time taken by the measuring instrument to step across the bandwidth, time is consumed by the software that provides control signals to specify the sweep parameters to the measurement subsystem and the data management block also consumes time for transferring the data through buffers to the database. Let this additional time be denoted by T_a. The measurement test is conducted over a certain period of time as specified by the *measurement test duration* T. Within the duration of a measurement test, the measuring instrument can perform one or more sweeps across the sweep bandwidth.

Since T_s is finite, the instantaneous power residing in a certain frequency channel cannot be measured continuously over time, but instead it is measured in time steps as specified by the *sweep time resolution* T_r. Thus, T_r specifies the time elapsed between two consecutive measurements of a certain frequency channel, which implies that T_r is directly proportional to T_s. Fig. 3.2 illustrates spectrum

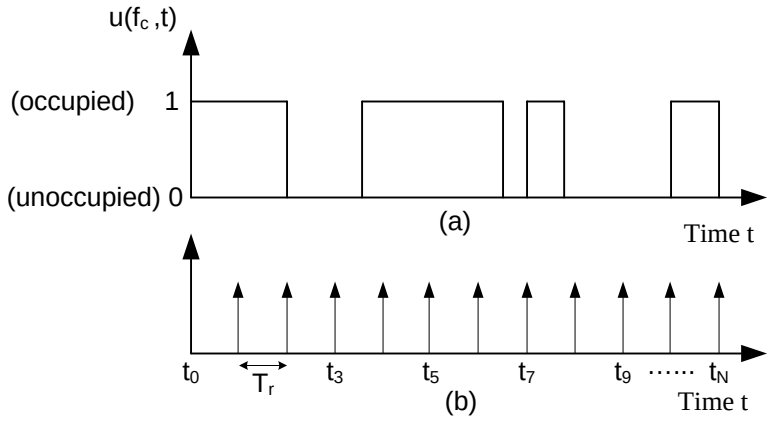

Figure 3.2. (a) Instantaneous channel states where '1' represents the channel being occupied and '0' represents the channel being available, (b) Representation of spectrum sensing as sampling with a pulse train.

sensing to measure the activity in frequency channel f_i. The instantaneous channel occupancy state $u(f_i, t)$ is a continuous function of time t. For time duration T beginning from time instance t_0, the measurement subsystem collects samples of $u(f_i, t)$, and the resulting sequence of N_t samples $\{x(n)\}$ can be represented as:

$$x(n) \quad = \quad u(t_0 + nT_r),\ n \ = \ 1, 2, \ldots N_t. \tag{3.1}$$

Note that T_s and B_r are inversely proportional to each other resulting in a trade off between T_r and B_r. A measurement sweep can be specified by *sweep time*, *sweep bandwidth* (F_{start}, F_{stop}) and the *bandwidth resolution*. Equations (3.2) through (3.6) represent the relationship between all the sensing parameters:

$$T_s \quad = \quad (\ N_f\, T_d\ ) \ + \ T_a\ , \tag{3.2}$$

$$N_f \quad = \quad B_s/B_r, \tag{3.3}$$

$$N_t \quad = \quad T/T_s, \tag{3.4}$$

$$T_r \propto T_s, \tag{3.5}$$

$$M(f_i, t_j) = \frac{1}{B_d\, T_d} \int_{t_j - \frac{T_d}{2}}^{t_j + \frac{T_d}{2}} \int_{f_i - \frac{B_d}{2}}^{f_i + \frac{B_d}{2}} P(f,t)\, \partial f\, \partial t, \tag{3.6}$$

where $P(f,t)$ represents the instantaneous power measured by the measuring instrument, and both f and t are continuous time variables of frequency and time respectively. $M(f_i, t_j)$ represents a measurement sample collected at frequency channel f_i and time instance t_j (see Eq.(3.7)).

The accuracy of the measurements depends on the receiver properties like sensitivity and selectivity, as well as on the resolution parameters of the measurement subsystem.

3.1.2 Representation of Spectrum Measurement Data

The dimensions of the spectrum, namely frequency, time, spatial extent, and signal format, are independent of each other such that any measurement sample can have a unique set of coordinates in the electrospace. The electrospace can be defined with respect to a specific receiver, i.e, it describes the RF operating environment of a receiver in units of radio field strength [2].

The frequency and time attributes refer to the frequency of the RF energy and the time it occurs. Spatial extent is the spatial volume that the receiver senses. It can be specified by the geographical location of the receiver (location type: urban, suburban, rural), the angle-of-arrival (azimuth) of the signal and the beam pattern of the antenna. Theoretically, if an isotropic antenna is used, the spatial extent is a sphere around the receivers antenna.

Signal format can be specified in terms of the type of polarization and modulation. It accounts for the use of orthogonal signal spaces, such as horizontal

polarization, vertical polarization, and code space[2]. In CDMA, two signals can be transmitted at the same time and on the same frequency channel but with different spread codes. In order to resolve between these two signals, knowledge about the specific spreading codes that have been used is required in order to de-spread the signals.

The measurements used for our analysis were taken with an omni-directional antenna, so the spatial extent is not relevant since the angle-of-arrival cannot be resolved. The measurements only include the power of the received signal, so the signal format is also not relevant. This is due to the fact that the power measurements do not convey any information about the polarization of the signal and also we do not attempt to de-spread signals in order to resolve the occupancy at different power levels. From our measurements, we can only resolve the spectrum occupancy along the frequency and time dimensions.

The spectrum measurements were collected along frequency and over a period of time. Accordingly, a set of measurements can be represented as an $N_t \times N_f$ matrix $\mathbf{M}$ defined as [33]:

$$\mathbf{M} = [M(f_i, t_j)] \quad , \tag{3.7}$$

$$\text{where} \qquad F_{start} \leq f_i < F_{stop} \quad , \quad T_{start} \leq t_j < T_{stop},$$

$$i = 1, \ldots N_f \quad , \quad j = 1, \ldots N_t,$$

given that $M(f_i, t_j)$ is a sample of the RF power (expressed in dBm) residing in frequency channel f_i at sweep time instance t_j, F_{start} and F_{stop} specify the start and stop frequencies for the measurement sweep, T_{start} and T_{stop} specify the start and stop time instances for the sweep, and N_t and N_f are the number of time

[2]Achieved using spreading codes.

instances and the number of frequency channels for which the measurements are collected. The sweep time parameters can be set such that the measurements can be collected over a few minutes or hours or days.

Across the sweep bandwidth (along frequency), the first measurement sample is collected at frequency F_{start}. Thereafter, the measurements are taken in steps of F_{step} at frequencies specified by:

$$\begin{aligned} f_i &= F_{start} + (i - 1 \times F_{step}) \ , \ i = 1...N_f \\ \text{where} \quad N_f &= \frac{F_{stop} - F_{start}}{F_{step}}. \end{aligned} \tag{3.8}$$

Note that the power at F_{stop} is not measured, instead the sweep ends at $F_{stop} - F_{step}$. Likewise, the measurement test duration is quantized into steps of T_{step}. The time instances at which the measurements are collected are:

$$\begin{aligned} t_j &= T_{start} + (j - 1 \times T_{step}) \ , \ j = 1...N_t \\ \text{where} \quad N_t &= \frac{T_{stop} - T_{start}}{T_{step}}. \end{aligned} \tag{3.9}$$

While T_{step} can be set by the user, the minimum time step is directly proportional to the sweep time. The parameters F_{step} and T_{step} correspond to the bandwidth resolution and the sweep time resolution of the measurements.

For analysis purposes, we can represent a sub-matrix of $\mathbf{M}$ by $\mathbf{M_{F,T}}$ which is defined over a range of frequencies and time instances specified by F and T, where F represents a sub-range of the frequency range $[F_{start} \ldots F_{stop})$ and T represents a sub-range of the time range $[T_{start} \ldots T_{stop})$.

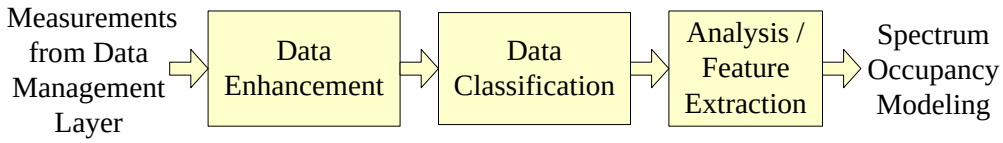

Figure 3.3. Stages involved in the processing and analysis of spectrum measurements.

3.1.3 Pre-processing and Classification of Spectrum Data

The data can be retrieved from the database for processing and analysis. Fig. 3.3 presents the block diagram of the different stages of processing involved in the processing and analysis block.

Initially, pre-processing operations involving data enhancement are performed on the data, followed by classification of the data. Data enhancement is done to condition the data and make it more suitable for classification. The pre-processed data can be represented as $\mathbf{M_p} = [M_p(f_i, t_j)]$.

In a given matrix of enhanced data $\mathbf{M_p}$, each matrix element $M_p(f_i, t_j)$ is classified into two classes, namely signal and noise, based on a decision threshold. The decision threshold η can be an estimate of the average noise power above which the signals occur. The threshold η is used for bilevel classification as:

$$M_c(f_i, t_j) \quad = \quad \left\{ \begin{array}{l} 1, M_p(f_i, t_j) \geq \eta \\ 0, M_p(f_i, t_j) < \eta \end{array} \right\}, \tag{3.10}$$

where $\mathbf{M_c} = [M_c(f_i, t_j)]$ represents the matrix of classified data which is called the *spectrum availability function*.

The threshold needs to be optimum in order to achieve classification with minimum errors. If the threshold is set too high, weak signals may not be identified. If the threshold is set too low, then even some noise samples can get classified

as signals. Either of the two cases result in erroneous classification. For critical applications such as spectrum sensing for cognitive radios, highly accurate classification techniques are required.

3.1.4 Characterization of Spectrum Utilization

The processed spectrum data is analyzed in order to determine the characteristics of the spectrum occupancy from the data. Since the frequency channel is treated as the fundamental unit for the purposes of dynamic spectrum management [15], the characterization of the spectrum can be done on a channel by channel basis. The following *spectrum utilization parameters* (SUPs) can be determined from the spectrum measurements:

1. Statistics on channel availability,
2. Received signal power,
3. Ambient noise power, and
4. Signal characteristics such as duty cycle, 'on' and 'off' times of a bursty signal, and signal bandwidth.

The matrix $\mathbf{M}$, in conjunction with $\mathbf{M_c}$, can be used to extract samples of the parameters, compute the probability distributions of the parameters from these samples, and estimate statistics, such as mean and variance, from these distributions.

Signal and Noise power Statistics: We can define mathematical functions to convert a measurement sample x expressed in dBm to milliwatts and vice versa:

$$\text{linear}(x) = 10^{x/10}\ (\text{mW}) = y \tag{3.11}$$

$$\text{decibelm}(y) = 10\log_{10} y\ (\text{dBm}) \tag{3.12}$$

Given a column vector of measurements collected from frequency channel f_i, along with their classification $\{M(f_i, t_j), M_c(f_i, t_j)\}$, samples of signal power and noise power can be extracted. The measurements $M(f_i, t_j)$ and the decision threshold η can be converted to the linear scale as:

$$M_l(f_i, t_j) = \text{linear}(M(f_i, t_j)), \tag{3.13}$$

$$\eta_l = \text{linear}(\eta). \tag{3.14}$$

Note that, η can be the local or global decision threshold which has been used to classify the measurement $M(f_i, t_j)$, or it can be the average noise level in the channels that are adjacent to the signal.

From the spectrum availability function, $\mathbf{M_c}$, the number of time instances when the signal has been present in the channel can be determined as:

$$K = \sum_{j=1}^{N_t} M_c(f_i, t_j). \tag{3.15}$$

A measurement which has been classified as signal has both signal and noise power components. A set of K samples of the signal power can be extracted from the

spectrum data $M(f_i, t_j)$ as:

$$\begin{aligned} S(f_i) &= \{\, s(k) \,:\, s(k) = M_l(f_i, t_j) - \eta_l,\ \forall\ t_j \text{ where } M_c(f_i, t_j) = 1\} \\ \text{where} \quad k &= 1, 2, \ldots, K. \end{aligned} \tag{3.16}$$

In a similar manner, a set of $N_t - K$ samples of the noise power $N(f_i)$ can be extracted as:

$$\begin{aligned} N(f_i) &= \{\, n(r) \,:\, n(r) = M_l(f_i, t_j),\ \forall\ t_j \text{ where } M_c(f_i, t_j) = 0\}\,, \\ \text{where} \quad r &= 1, 2, \ldots, N_t - K. \end{aligned} \tag{3.17}$$

The extracted samples can be used to compute histograms of the signal and noise power in the channel and statistics such as mean, standard deviation, dynamic range $= Max\,\{S(f_i)\} - Min\,\{S(f_i)\}$, and threshold crossing rate $P(S(f_i) \geq s)$. For instance, the signal-to-noise ratio can be determined as:

$$\overline{S}(f_i) = Mean\,\{S(f_i)\}\,, \tag{3.18}$$

$$\overline{N}(f_i) = Mean\,\{N(f_i)\}\,, \quad \text{and} \tag{3.19}$$

$$SNR(f_i) = 10\,log_{10}\left(\frac{\overline{S}(f_i)}{\overline{N}(f_i)}\right). \tag{3.20}$$

Signal Characteristics: Given a row vector of classified spectrum data from $\mathbf{M_c}$, the signal edges can be identified in order to determine the signal bandwidth. In a similar manner, given a column vector from $\mathbf{M_c}$, the start and stop times of signal transmissions can be identified in order to determine the 'on' and 'off' times of the signals.

3.1.5 Modeling of Spectrum Measurements

The SUPs extracted from the measurements can be incorporated into the proposed comprehensive model of the spectrum measurements that characterizes both the dynamics of (a) spectrum utilization, and (b) primary signal characteristics. In our spectrum model, a measurement can be represented as a function of various components, such as signal, noise, and channel occupancy, using:

$$\begin{aligned} M(f_i, t_j) &= (M_c(f_i, t_j) \times S(f_i, t_j)) + N(f_i, t_j)\,, \qquad (3.21) \\ \text{where} \quad M(f_i, t_j) &\in \mathbf{M}. \end{aligned}$$

$S(f_i, t_j)$ and $N(f_i, t_j)$ are the signal and additive noise power components of the measurement sample $M(f_i, t_j)$. Each component of the model can be modeled as a random variable. For instance, the channel occupancy is represented by a random variable X_{f_i} and $M_c(f_i, t_j)$ represents a sample of the random variable. To model the spectrum occupancy in a channel, we propose a modified Markov model that incorporates the time varying nature (non-stationarity) of the channel occupancy. While the Laycock-Gott occupancy model is complicated and not suited for real time applications, the previously proposed Markov models [34] assumed the channel to be stationary. However, in practice the model parameters may be time varying. For instance, the usage of cellular mobile channel varies drastically with the time of the day with the peak usage being during the business hours.

The dynamics of X_{f_i} can be modeled by a two-state Markov chain, which is shown in Fig. 3.4. The parameters of the markov model are the state probabilities denoted by P_0 and P_1 corresponding to the states '0' and '1' and the state tran-

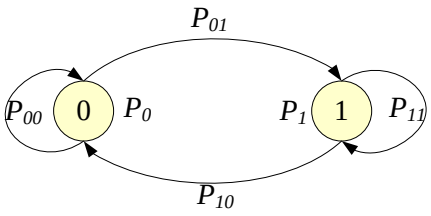

Figure 3.4. A simple markov model of a channel's occupancy states

sition probabilities P_{00}, P_{01}, P_{10}, and P_{11}. A weighted approach has to be taken to compute the model parameters where, more weight is given to the current and most recent instantaneous channel states. For a particular frequency channel f_i and at a particular time instance t_k, the weighted count of past occurrences of state '1' is determined from $\mathbf{M_c}$ and is denoted by $N_w(M_c(f_i, t_j) = 1)$ and the corresponding weighted state probability is defined by:

$$P_1^i(t_k) = \frac{N_w(M_c(f_i, t_j) = 1)}{N_{t_k}} \tag{3.22}$$

$$P_0^i(t_k) = 1 - P_1^i(t_k) \tag{3.23}$$

$$\text{where} \quad j = 1, \ldots N_{t_k}.$$

The weights are given based on the time of occurrence of the channel state relative to t_k. The variable t_k is a discrete time variable, where the total number of measurement samples collected from time t_j= 1 to $(t_k - 1)$ is denoted by N_{t_k}. The estimate of the true probability $P_1^i(t_k)$ is computed from the finite set of measurement samples using:

$$\bar{P_1^i}(t_k) = \frac{1}{N_{t_k}} \sum_{j=1}^{N_{t_k}} M_c(f_i, t_j) \cdot \exp\left(\gamma(t_k - j)\right) \tag{3.24}$$

where γ is the *forgetting factor* for $0 \leq \gamma \leq 1$. The forgetting factor has been in-

troduced into 3.24 in order to give more weightage to the most recent samples and less weightage to past samples. Similarly, the weighted probability of transitions from state '0' to state '1' in channel f_i is computed as:

$$\begin{aligned} P_{01}^{i} &= \frac{N_w(M_c(f_i, t_{j+1}) = 1, M_c(f_i, t_j) = 0)}{N_{t_k} - 1}, \\ \text{where} \quad j &= 1, ...N_{t_k} - 2. \end{aligned} \tag{3.25}$$

Note that the term in the numerator represents the weighted number of occurrences of transitions from state '0' to state '1' in channel f_i. In the same manner, the other transition probabilities and model parameters can be computed.

3.1.6 Efficient Characterization of Spectrum Utilization

The time taken t_k to estimate a statistically accurate markov model has to be kept to a minimum so that the characterization period is small. In the case of a stationary random process, as the number of measurement samples used for computing the estimates increases, the accuracy of the estimates increases, assuming that a consistent estimator is used. There arises a question on what is the minimum number of measurement samples that is required in order to estimate the parameters of the model, given finite measurement data and the sampling limitations of the spectrum sensing mechanism. In this discussion, the model parameters represent statistics of the random variable X_{fi}. We make an assumption that the channel occupancy is piecewise stationary. We base our discussion over a segment of time when the channel occupancy is stationary. We also assume that the random variable X_{fi} is ergodic and that the measurement samples are observed independently and under similar experimental conditions. With these assumptions, the N_{t_k} measurement samples that are collected from

the frequency channel f_i can be viewed as constituting the sample space of X_{fi}. The time spacing, T_r, between the samples is the sweep time resolution of the sensing mechanism. If T is the total time duration for which the measurement samples are collected then $T = N_{t_k} \cdot T_r$. If the resolution is improved then the number of samples N_{t_k} is increased. The relationship between the measurement test parameters and the statistical error [3] of the estimated parameters can be derived.

Consider the following expression for estimation of the true probability P_1^i for the case of the stationary channel occupancy:

$$\bar{P_1^i} = \frac{1}{N_{t_k}} \sum_{j=1}^{N_{t_k}} M_c(f_i, t_j), \tag{3.26}$$

$$\overline{\sigma}_s^2 = \frac{1}{N_{t_k} - 1} \sum_{j=1}^{N_{t_k}} (M_c(f_i, t_j) - \bar{P_1^i}(t_k))^2. \tag{3.27}$$

In Eq. (3.26), $M_c(f_i, t_j)$ represents a sample of the two state random variable X_{fi} and the summation of the samples, which is the number of instances when the channel has been occupied, follows a binomial distribution $B(N_{t_k}, P_1^i)$. $\bar{P_1^i}$ is a consistent estimator of P_1^i and it can be treated as a random variable with a certain *sampling distribution*. For large N_{t_k} ($N_{t_k} > 10$), the binomial distribution can be approximated with a normal distribution[4] and $\bar{P_1^i}$ can be treated as a Gaussian random variable with normalized form denoted as z $\sim \mathcal{N}$(0,1). If α is the confidence coefficient, we can denote the 100α percentage point by z_α [35]:

$$P(z_\alpha) = \int_{-\infty}^{z_\alpha} p(z)\, dz = Prob[z \leq z_\alpha] = 1 - \alpha. \tag{3.28}$$

[3] Statistical error is the error between the estimate of a parameter and the actual parameter itself and it arises due to estimation from a finite number of samples.

[4] The normal distribution is $\mathcal{N}(N_{t_k} P_1^i, N_{t_k} P_1^i (1 - P_1^i))$.

Let t_n be the corresponding random variable with a student t distribution of n ($n = N-1$) degrees of freedom and its 100α percentage point is denoted as $t_{n,\alpha}$. χ^2_n represents the corresponding random variable with a chi-square distribution of n ($n = N-1$) degrees of freedom with 100α percentage point denoted as $\chi^2_{n;\alpha}$. In the derivation of the sampling distribution, we make an assumption that the samples of the channel occupancy are not correlated with each other and this assumption can be met by sampling the channel occupancy at random intervals as described in [36]. From the sampling distribution of $\bar{P_1^i}$, the confidence interval for P_1^i can be defined as [35]:

$$P\left[\bar{P_1^i} - d \leq P_1^i < \bar{P_1^i} + d\right] = 1 - \alpha, \tag{3.29}$$

$$d = \frac{\overline{\sigma}_s\, t_{n,\alpha}}{\sqrt{N_{t_k}}}. \tag{3.30}$$

It is seen from Eq. (3.29) that for a given degree of uncertainty, as the number of samples N_{t_k} increases, the confidence interval narrows and the accuracy of the estimate increases.

This discussion is applicable to the estimation of statistics of the other SUPs and parameters of the markov model where we can treat the SUP as a random variable X (or a random process) with true mean μ_x and true standard deviation σ_x, represent a statistical parameter of X by Φ_x, and its estimate by $\overline{\Phi}_x$. $\overline{\Phi}_x$ can be treated as a random variable with a certain sampling distribution [35]. The estimators of mean and variance of X, are shown below:

$$\overline{x} = \frac{1}{N_{t_k}} \sum_{j=1}^{N_{t_k}} x_j, \tag{3.31}$$

$$s^2 \quad = \quad \frac{1}{N_{t_k} - 1} \sum_{j=1}^{N_{t_k}} (\overline{x} - x_j)^2. \tag{3.32}$$

The corresponding confidence intervals are [35]:

$$P\,[\,\overline{x} - d \;\leq\; \mu_x \;<\; \overline{x} + d\,] \quad = \quad 1 - \alpha, \tag{3.33}$$

$$\text{where} \quad d \quad = \quad \frac{s\, t_{n;\alpha/2}}{\sqrt{N_{t_k}}} \quad = \quad \frac{\sigma_x\, z_{\alpha/2}}{\sqrt{N_{t_k}}},$$

$$P\left[\frac{n\, s^2}{\chi^2_{n;\alpha/2}} \;\leq\; \sigma_x^2 \;<\; \frac{n\, s^2}{\chi^2_{n;1-(\alpha/2)}}\right] \quad = \quad 1 - \alpha, \tag{3.34}$$

$$\text{where} \quad n \quad = \quad N_{t_k} - 1.$$

3.2 Evaluation of Spectrum Survey Results

To evaluate the efficacy of the SSF and compare the performance among different implementations of this framework, a uniform performance evaluation procedure is required.

3.2.1 Quantitative Performance Evaluation

The processed data that appears at the output of the classification algorithm is compared with the *ground truth*. The ground truth which is obtained from a reliable source represents knowledge about the true presence or absence of a signal at a particular frequency and time instance. The decision made by the classification algorithm may or may not conform with the ground truth, which results in correct or incorrect classification. A *false alarm* occurs when a noise sample has been incorrectly classified as signal, and a *miss detection* occurs when a signal sample has been incorrectly classified as noise.

The ground truth can be obtained from the licensing database maintained

by the spectrum regulatory agencies. For instance, the FCC maintains licensing information on the channel allocations of the FM broadcast spectrum (88-108 MHz). The FCC website provides a database of all the radio stations within a certain radius around the measurement site which is located in Lawrence, Kansas, USA, along with the height and GPS coordinates of the towers[5]. Using this data, along with FM propagation curves provided by the FCC, we can determine the approximate ERP (effective radiated power) that the measurement subsystem should be recording [33]. This data yields a list of FM stations (center frequencies of the FM stations) that can be captured above a specified power level (this has been assumed to be -100 dBm). This list constitutes the ground truth. All the frequency bands of 150 kHz each with the center frequencies obtained from this list are tagged as 'occupied' (represented as '1'). In this way, a matrix $\mathbf{T} = [T(f_i, t_j)]$, of the same dimensions as the spectrum utilization matrix $\mathbf{M_c}$, can be constructed from the ground truth. By comparing every element of $\mathbf{M_c}$ with the corresponding element of $\mathbf{T}$, we can construct a model contingency table as shown in Table 3.1. From this table, we can compute the false alarm rate and the miss rate [37].

Table 3.1. Model contingency table showing relationship between ground truth and measurement classifications

		Decision by classification algorithm		
		Signal	Noise	Total
Ground Truth	Signal	a	b	$a + b$
	Noise	c	d	$c + d$
	Total	$a + c$	$b + d$	$N = a + b + c + d$

In Table 3.1, the quantities a, b, c, and d represent the count of the joint

[5]http://www.fcc.gov/mb/audio/fmq.html, coordinates of Lawrence: 38^o 57' 36"N, 95^o 15' 12"W

occurrences of the respective variables. The false alarm rate, the miss rate, and the weighted error can be computed as:

$$FA\,(\,\%\,) \;=\; \frac{c}{c+d} \;\times\; 100, \tag{3.35}$$

$$Miss\;Rate\,(\,\%\,) \;=\; \frac{b}{a+b} \;\times\; 100, \tag{3.36}$$

$$WE\,(\,\%\,) \;=\; \frac{w_1\,b + w_2\,c}{N} \;\times\; 100. \tag{3.37}$$

In the case of dynamic spectrum access systems, the highest priority must be given to minimizing the miss rate. This is to prevent the radio from inadvertently identifying transmissions from other services as noise and subsequently transmitting to cause interference to them. On the other hand, the false alarm rate should also be kept to a minimum so that the largest available area of spectrum can be identified and used [33]. Thus, for our evaluation a miss detection has been penalized more than a false alarm by considering $w_1 = 0.85$ and $w_2 = 0.15$.

3.2.2 Qualitative Performance Evaluation

A linear transformation from a matrix of measurements $\mathbf{M}$, into a matrix of gray scale values $\mathbf{I}$, is given by:

$$I(f_i, t_j) \;=\; \frac{1.0 - 0.0}{Max\,\{\mathbf{M}\} - Min\,\{\mathbf{M}\}} \;\times\; (\; M(f_i, t_j) \;-\; Min\,\{\mathbf{M}\}\;)\,. \tag{3.38}$$

The matrix $\mathbf{I} = [I(f_i, t_j)]$, can be displayed as a discrete image or a gray scale intensity plot. The range of the pixel values is $0.0 \;\leq\; I(f_i, t_j) \;\leq\; 1.0$, where 0.0 represents the luminance of the darkest pixel and 1.0 represents the luminance of the brightest pixel. This image is termed as a *spectrum image*. In the absence of the ground truth, the performance of the classification algorithms can be evaluated

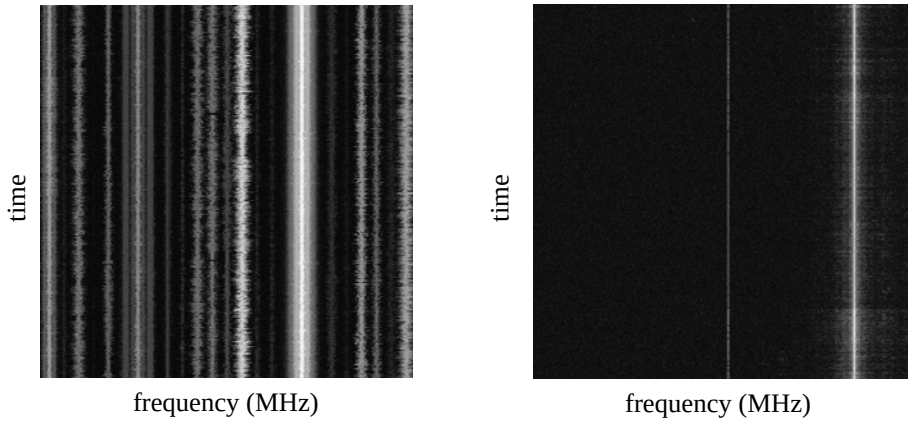

Figure 3.5. Spectrum images generated from FM band measurements (88-92.99 MHz) (left) and upper TV band measurements (72.99-77.99 MHz)(right).

qualitatively by viewing the spectrum image which is generated from the matrix of classifications that occurs at the output of the classification algorithm. It is called so since the signal and noise data in the measurement data set have been segmented. This method may not be as accurate as the quantitative evaluation.

Fig. 3.5 displays the spectrum images that were generated from FM band measurements and the upper TV band (54-87 MHz) measurements. In these images, bright patterns (straight lines) can be observed against a dark background. The bright patterns are either due to signals or intermodulations, while the dark background is mostly due to noise.

3.3 Implementation of Spectrum Survey Framework

Fig. 3.6 shows the block diagram of our implementation of the SSF. A HP 8594E spectrum analyzer has been used as the measurement subsystem. The *Spectrum Miner*[6] software implements the measurement subsystem's user interface, the software module that controls the measurement subsystem, and the data

[6]The spectrum miner software has been developed in-house at the University of Kansas

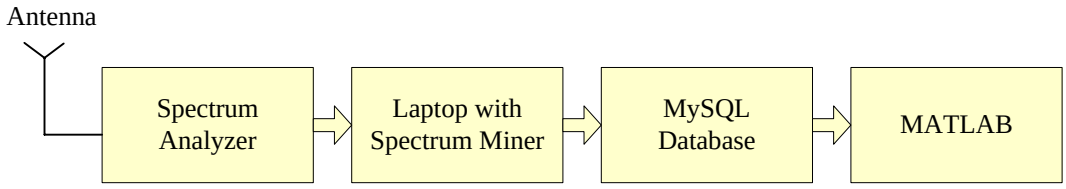

Figure 3.6. Implementation of SSF

management layer. The Spectrum Miner is a software tool designed to automate the collection and storage of spectrum measurements. These measurements can be exported from the database into analysis programs in Matlab or Mathematica or can be archived in a web-based spectrum data repository. Reference [33] gives more details on the architecture of the spectrum miner software and the database. In this implementation, the main focus of this thesis is the processing and analysis of spectrum measurements[7].

Simple techniques have been used for performing preliminary investigation on the processing of the spectrum measurements which are discussed next.

3.3.1 Classification Based on Analysis of Cumulative Density Function

Let S be a random variable, representing the measured power. A cumulative distribution function (CDF) of S, computed from the measurement data $\mathbf{M}$, can be defined as:

$$CDF(S, \mathbf{M}) = P(S \leq M(f_i, t_j)) \ , \ M(f_i, t_j) \ \in \ \mathbf{M} \tag{3.39}$$

Based on this CDF, simple techniques can be used to determine the decision threshold. The next two examples will illustrate this approach.

[7]In this implementation, the MySQL database, the spectrum miner software, and the Matlab codes to import data from MySQL into Matlab have been developed by Ted Weidling and Rory Petty at the University of Kansas.

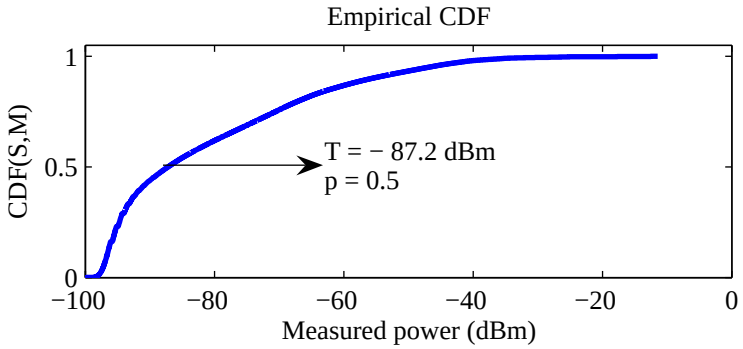

Figure 3.7. CDF plot of the FM band measurement data. Threshold T for p = 0.5 is determined from CDF.

Example 1: p-tile thresholding -

The threshold is chosen based on *a priori* knowledge about the spectrum utilization [38]. For instance, if it is known that the fraction of the spectrum being utilized is p, then choose a threshold T such that p fraction of the measurements have values greater than T, i.e.

$$1 - p \;=\; CDF(S = T, \mathbf{M}) \tag{3.40}$$

Fig. 3.7 shows the CDF computed from the FM band measurements.

Example 2: Compute a marginal CDF, denoted by CDF_{f_i}, for each frequency in the sweep bandwidth and classify a frequency channel f_i based on a statistic of CDF_{f_i}. For example, determine the maximum measurement for frequency channel f_i and if the maximum is above a certain threshold, classify f_i as an occupied frequency channel. This threshold can be manually set based on some factors such as the noise floor of the measurement subsystem.

These techniques possess the following limitations: First, *a priori* knowledge about the spectrum occupancy is required for the classification, and second, the

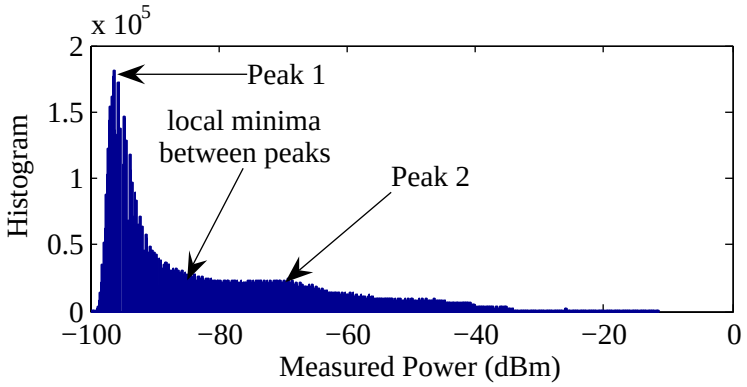

Figure 3.8. Histogram of the FM band (88-108 MHz) measurements.

classification is dependent upon manually selecting a threshold value. Due to these drawbacks, these classification techniques cannot be implemented in an automated fashion in a cognitive radio.

3.3.2 Classification Based on Histogram Analysis: Mode Method

The global histogram of the measurements can be expected to be bimodal[8]. In the bimodal histogram, the two peaks belong to the signal and noise samples respectively. The measurement values at the edges[9] of the signals occur less frequently in the measurement data, as compared to the signal and noise values. Thus, the valley between the peaks in the histogram may belong to the measurement values at the edges of the signals. The value of the local minimum between the two peaks or the center point (mean) between the peaks can be chosen as the decision threshold [38] (see Fig. 3.8). Table 3.2 presents the performance evaluation of this method.

If we consider a sparsely occupied band, such as the upper TV band (54-87 MHz) where there is presence of a large percentage of background noise, the

[8]A bimodal histogram has two peaks.
[9]Boundary separating a signal pattern from the noise in the vacant adjacent channels.

Table 3.2. Performance evaluation of mode method applied on FM band (88-108 MHz) data.

Threshold (dBm)	Miss (%)	FA (%)	Error (%)	Weighted error (%)
Local Minima = -81	17.3194	23.6210	21.9196	6.5613
Mean of peaks = -82.5	15.3942	26.0085	23.1426	6.3809

histogram may not be bimodal as seen from Fig. 3.9. In such cases, additional processing has to be done to make the histogram bimodal [38]. The Laplacian operator is applied on $\mathbf{M}$. The Laplacian forms the spatial second partial derivative of a function $F(x, y)$ (i.e., the rate of change in slope) and has the mathematical form [38]:

$$\begin{aligned} G(x,y) &= -\nabla^2 \{ F(x,y) \} \\ \text{where} \quad \nabla^2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}. \end{aligned} \tag{3.41}$$

Consider the case when a signal is surrounded by uniform noise[10] present in the adjacent channels. Ideally, at the signal edges the gray values increase from a low plateau level (belonging to the uniform background noise) to the peak power level of the signal waveform in a smooth ramp-like manner. In the plateau and along the ramp where the slope is constant, the Laplacian is zero. However, in the regions where there is a transition from the low plateau to the ramp or from the ramp to the signal peak, the Laplacian has a large magnitude. A histogram formed only from measurement samples that lie at coordinates corresponding to a high magnitude of Laplacian is expected to be bimodal [38].

A threshold[11] $L(n) = \mu_L + (n\,\sigma_L)$ is used to identify the high magnitude

[10]Uniform noise refers to noise which does not vary much across the target frequency band.
[11]This threshold is different from the decision threshold.

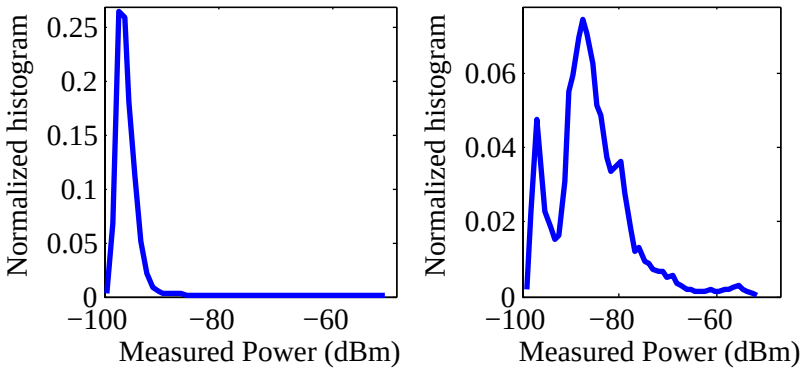

Figure 3.9. Normalized histogram of the Upper TV band (54-87 MHz) measurement data (left), and normalized histogram of selected samples in the TV band data (right).

Laplacian values, where μ_L and σ_L are the mean and standard deviation of the Laplacian values, and n is a positive integer which can be specified. Fig. 3.9 shows the normalized histogram of the TV band (54-87 MHz) measurements with Laplacian greater than $L(4)$. Fig. 3.10 shows the results of classification of the TV band (54-87 MHz) data for different values of $L(n)$. As seen from Fig. 3.10, the best classification has been obtained for $n = 4$.

The drawback in the mode method is that, the local minima has to be selected manually. This approach can be automated by analytically representing the shape of the histogram and then performing an optimization of this analytical expression. However, such methods are not always accurate [38].

3.4 Chapter Summary

The high level structure for the Spectrum Survey Framework has been presented followed by discussion of each component of the framework. In this framework, the procedures involved with the collection, processing, and analysis of spectrum measurements have been described in terms of mathematical expres-

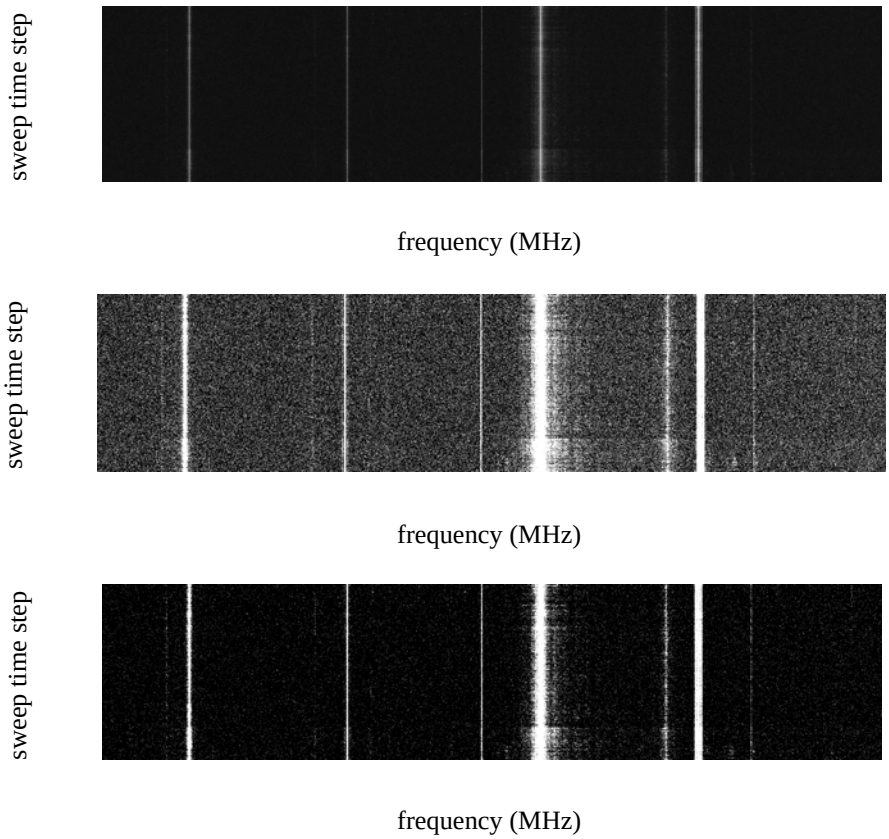

Figure 3.10. Results of classification of TV band (54-87 MHz) measurements using various values of $L(n)$ (from top to bottom):(a) Original spectrum image, (b) Image of data classified with n = 2, (c) n = 4.

sions and suitable notations. A modified markov model that accounts for the time varying nature of the channel occupancy was also presented. The issue of efficient characterization of the channel occupancy was addressed by presenting an expression to determine the minimum time required to estimate the channel model parameters. Our implementation of the SSF has been presented along with preliminary processing methods.

Chapter 4

Threshold-Based Classification of Spectrum Measurements

In this chapter, we discuss the various challenges for classifying the spectrum measurements (Section 4.1). In Sections 4.2-4.5, techniques have been proposed to over come these challenges and process the data. The proposed classification algorithms can estimate the threshold based on the statistical properties of the data and do not require any *a priori* knowledge about the signals present in the spectrum. The proposed techniques have been applied to the spectrum data and the results are shown in Section 4.6.

4.1 Challenges for Threshold-Based Classification of Spectrum Data

There are several factors that affect the performance of the classification algorithms:

Range of ambiguous classification - Due to the presence of weak signals that

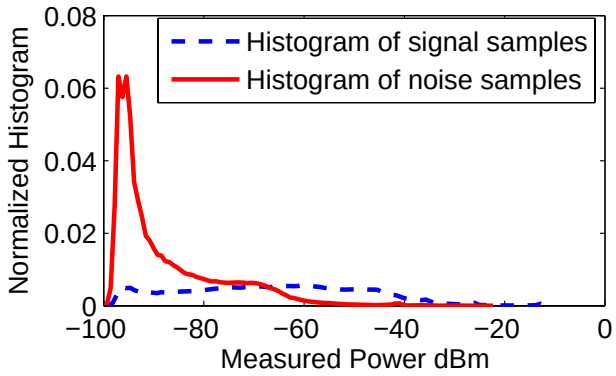

Figure 4.1. Histograms of noise and signal measurement samples taken from FM band (88-108 MHz) measurement data. The histograms are seen to overlap over a large area.

occur below the noise level, there can be some overlap between the histograms of the signal[1] and noise samples present in the measurement data (see Fig. 4.1 [2]). Across the range of measurement values over which the histograms overlap, there can be ambiguity in the classification of the measurements. Due to this overlap in the histograms, by setting a certain threshold there is always a tradeoff between the miss detection rate and the false alarm rate.

Variance in signal and noise power levels- The noise and signal samples in the data can have a wide range of values. In addition, the signal-to-noise ratio can vary across the data.

Biasing effect- The spectrum measurements may capture signals received from distant transmitters as well as transmissions from nearby sources. As a result, weak signals need to be detected in the presence of strong signals. When the statistical classification techniques are applied on this data which contains a high dynamic range of signals, the stronger signals in the data set may bias the thresh-

[1]The signal being referred here is not the pure signal but contains noise added to it.
[2]The signal and noise samples were sorted from the data with the help of the ground truth.

old estimation process. The resulting threshold may be so high that it may not detect weak signals, thereby leading to miss detection.

Ghost signals and ghost noise- An inherent property of threshold-based techniques is that they bifurcate a set of data into two classes, irrespective of whether samples from both the classes are present or not in the data. Thus, when applying thresholding techniques on a data set containing only noise samples, the stronger noise samples get wrongly classified as signals. Such signals are termed as *ghost signals* and they results in false alarms. Similarly, *ghost noise* samples can appear when classifying data that contains only signals resulting in miss detection.

To counter these problems, we propose the following techniques for the classification of measurement data:

1. Optimum thresholding,
2. Data enhancement and noise suppression,
3. Recursive thresholding, and
4. Adaptive thresholding.

Ideally, optimum thresholding can achieve the best classification results in scenarios where there is overlap in the signal and noise histograms. Furthermore, by suppressing the noise without affecting the signals, the noise histogram is shifted to lower power levels such that the overlap in the histograms is reduced. Recursive thresholding and adaptive thresholding can be used to counter the *biasing effect* and the problem of non-uniform power levels. Otsu's algorithm and the data enhancement techniques have been presented from existing concepts [38–40]. The algorithms for recursive thresholding are being proposed in this thesis. The

adaptive thresholding algorithm presented in this thesis has been built upon the original sliding window approach which has been proposed in [10].

4.2 Optimum Thresholding using Otsu's Algorithm

Otsu's algorithm [39] selects an optimum threshold based on the properties of the histogram of the data and it does not assume any model for the histogram. The optimum threshold results in the maximum separation between the two classes of data, namely the signal and the noise classes. The algorithm also returns a metric that indicates the separability of the two classes, which is useful to quantify the *goodness* of the threshold. Before applying Otsu's algorithm, the measurement data in $\mathbf{M}$ is converted to the gray scale image $\mathbf{I}$.

The data is quantized into L levels with values $s \times [1, 2, ..L]$, where s is a scaling factor. Let the i^{th} gray level value be denoted by g_i and its probability of occurrence is denoted as p_i. The mean of the distribution is defined as:

$$\mu_T \;=\; \sum_{i=1}^{L} g_i \;\cdot\; p_i.$$

A threshold, $T \;=\; g_k$, can be used to bifurcate the probability distribution into the noise class C_0 and the signal class C_1, with the levels $[1, 2, .. k] \in C_0$ and levels $[k+1, ..L] \in C_1$. For a certain threshold set at the k^{th} gray level, the *between-class variance* (see Appendix A) is defined as:

$$\sigma_B^2\,(k) \;\;=\;\; \frac{[\,\mu_T\,\omega_k - \mu(k)\,]^2}{\omega_k\,(1-\omega_k)}, \tag{4.1}$$

$$\text{where} \quad \omega_k \;=\; \sum_{i=1}^{k} p_i, \tag{4.2}$$

$$\text{and} \quad \mu\,(k) \;=\; \sum_{i=1}^{k} g_i\; p_i. \tag{4.3}$$

A measure of class separability can be defined as [39]:

$$\alpha \;=\; \sigma_B^2/\sigma_T^2. \tag{4.4}$$

Otsu's algorithm involves determining the gray level of the optimum threshold, k^*, that maximizes the measure of class separability:

$$\alpha\,(k^*) \;=\; max_{1\leq k<L}\,\sigma_B^2(k)/\sigma_T^2. \tag{4.5}$$

4.3 Enhancement of Spectrum Measurement Data

The measurement data is refined by performing the following operations:

1. Suppress the high spatial frequency components by low pass filtering and reduce the noise variance by time averaging, and

2. Improve the contrast between signals and noise i.e. enhance the weak signals and suppress the noise, which leads to an improvement in the SNR of the data.

Eventually, these operations can result in increasing the separation between the noise and signal histograms and thus improve the class separability.

4.3.1 Clipping and Contrast Manipulation

Clipping is done to reduce the dynamic range of the measurements and thus reduce the *biasing effect*. Assuming that the left most and the right most portions of the histogram [3] contain only noise and signal samples respectively, the histogram can be clipped at power levels η_l and η_r. The clipping operation can be combined with contrast manipulation where the contrast between the signals and noise can be manipulated by appropriate amplitude scaling, as represented by:

$$
\begin{aligned}
I_p(f_i,t_j) &= \left\{ \begin{array}{l} (\,\eta_r\,)^p\,,\ I(f_i,t_j) \geq \eta_r \\ [\,k\ \cdot\ (\,I(f_i,t_j)\ -\ \eta_l\,)\,]^p\,, \eta_l < I(f_i,t_j) < \eta_r \\ (\,\eta_l\,)^p\,,\ I(f_i,t_j) \leq \eta_l \end{array} \right\}, \qquad (4.6) \\
\text{where}\quad k &= \frac{1.0-0.0}{\eta_r-\eta_l},\ 0.0\ <\eta_l,\eta_r<\ 1.0, \\
M(f_i,t_j) &\in \mathbf{M}, \quad \text{and} \\
p &\geq 1.
\end{aligned}
$$

$I(f_i,t_j)$ is the gray scale equivalent of $M(f_i,t_j)$, and $I_p(f_i,t_j)$ is the processed gray scale measurement sample. The measurements can be raised to the power p. For $p<1$, the right most portion of the histogram, consisting mainly of the signal samples, gets suppressed while the left portion, consisting mainly of the noise samples, gets scaled up.

[3] Note that this is an histogram of power measurements.

4.3.2 Low Pass Filtering for Noise Suppression

The low pass filtering operation is performed using a spatial-domain averaging filter, of dimensions $r \times c$, with the impulse response given by:

$$\mathbf{H} = \frac{1}{r \times c} \mathbf{1_{r \times c}}, \tag{4.7}$$

where $\mathbf{1_{r \times c}}$ is a $r \times c$ matrix of ones.

The Gaussian low pass filter can also be used and its normalized impulse response is [40]:

$$\begin{aligned} \mathbf{G_N} &= [\, G_N(r,c) \,], \\ \text{where} \quad G_N(r,c) &= \frac{G(r,c)}{\sum_r \sum_c G(r,c)},\ G(r,c) = exp\left(-\frac{r^2 + c^2}{2\,\sigma^2}\right), \end{aligned} \tag{4.8}$$

where r, c are the dimensions of the filter with the impulse response matrix $\mathbf{G_N}$ and σ is the standard deviation of the Gaussian impulse response[4].

The filtering involves convolution between input matrix $\mathbf{I}$ and the impulse response $\mathbf{H}$, to give the filtered output as [38]: $\mathbf{I_p} = \mathbf{I} \otimes \mathbf{H}$. Multiple stages of filters can also be used. The enhancement operations described so far are generally applied on gray scale values.

4.3.3 Time Averaging of Spectrum Measurement Data

In certain bands with fixed channelization, such as the FM and TV broadcast bands, most active licensees transmit continuously for 24 hours. In such cases, all the measurement sweeps of data collected across a band represent redundant

[4]The Matlab built-in functions for Otsu's algorithm, Gaussian low pass filter, and the median filter have been used for the processing.

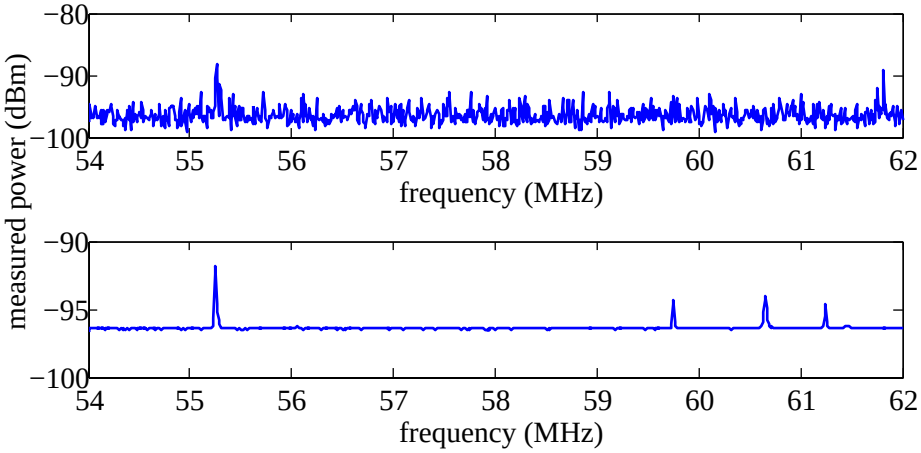

Figure 4.2. Power spectrum of the 54-62 MHz band in the upper TV spectrum before time averaging (top) and after time averaging (bottom).

data. By averaging over such redundant sweeps of data, which is affected by independent random noise, the noise variance is reduced, as shown in Figs. 4.2 and 4.3. However, this method may not be effective when applied to bands occupied by bursty signals. The average power at frequency f_i, $M_a(f_i)$, can be computed from the measurements as:

$$M_a(f_i) = \frac{1}{N_t} \sum_{j=1}^{N_t} M(f_i, t_j),\ F_{start} \leq f_i < F_{stop}. \tag{4.9}$$

4.3.4 Post-classification Processing: Median Filtering

The classified output may contain ghost signal samples that appear as grains in the spectrum image of the classified data. Median filtering can be done to remove these grains and reduce the false alarms. The median filter consists of a sliding window that extracts subsets of the data and replaces the center data sample in the subset by the median of the subset.

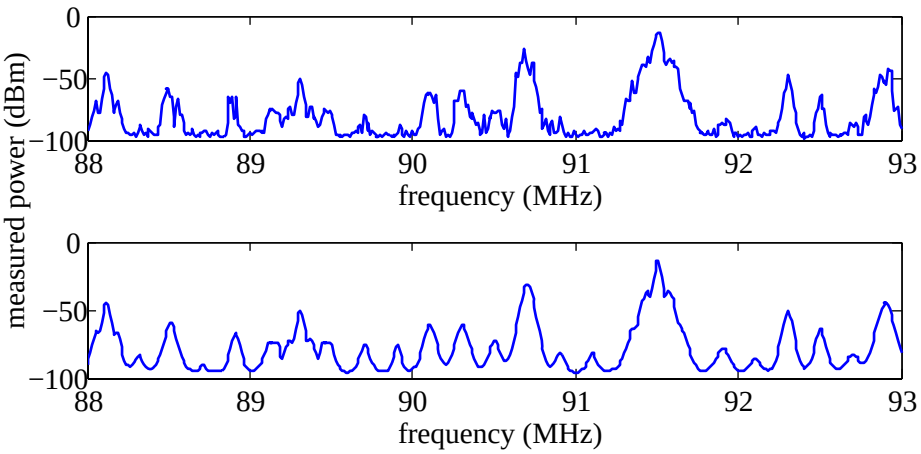

Figure 4.3. Power spectrum of the 88-93 MHz band in the FM broadcast spectrum before time averaging (top) and after time averaging (bottom).

4.4 Classification by Recursive Thresholding

Recursive thresholding classification is useful for detecting signals with a wide range of power levels.

4.4.1 Recursive One-Sided Hypothesis Testing algorithm

We[5] have proposed the *Recursive One-Sided Hypothesis Testing* (ROHT) algorithm that performs classification of the spectrum measurements based on the concept of one-sided hypothesis testing. The algorithm works for various levels of statistical significance. This algorithm makes the assumptions that the measurement data follows a Gaussian distribution and that there are sufficient number of measurement samples so that the estimates of mean and standard deviation obtained from the data are accurate.

The algorithm begins by assuming that the set of measurements contains

[5]The ROHT algorithm has been conceptualized and developed by Dinesh Datla, Ted Weidling, and Rory Petty at the University of Kansas, and has been published in [33]. The algorithm and some of the results have been reproduced here with the written permission of all the authors.

mostly noise samples. A percentage of the measurement data (specified by the z-value) on the far right of the Gaussian distribution is identified as signal samples and the rest as noise. The signal portion is discarded and this process is repeated iteratively on the remaining unclassified measurements. As a result, at every iteration, the standard deviation is reduced as shown by Fig. 4.4. The algorithm stops iterating when the change in the standard deviation between two consecutive iterations becomes less than or equal to ϵ, where ϵ is an arbitrary positive value that can be specified. In every iteration, the stronger signals are discarded, thereby reducing their biasing effect on the weaker signals. As the number of iterations increases, there can be an improvement in the miss rate but at the cost of an increase in false alarm rate.

The arguments that are passed to the algorithm are the z-value corresponding to the confidence level and the ϵ value. The algorithm is mathematically represented as:

Let

- **M** be the set of measurement samples,
- S be the set of signals within **M**,
- S_k be a subset of S for the k^{th} iteration of the algorithm,
- Q be the set of noise samples within **M**,
- Q_k be a superset of Q for the k^{th} iteration of the algorithm, Q_k may contain signals,
- μ_k , σ_k = mean and standard deviation of the elements of Q_k, and
- θ_k= decision threshold to identify the signal portion for the k^{th} iteration.

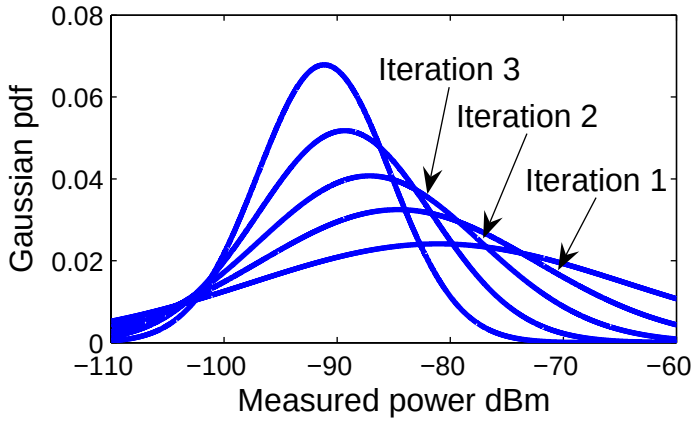

Figure 4.4. Normal distribution of measurement samples to illustrate the first four iterations of ROHT algorithm for a 99 % confidence interval.

Initialize S = $\oslash$, $S_o = \oslash$, $Q_o = M$, k = 0

do

- $\theta_{k+1} = zvalue * \sigma_k + \mu_k$
- $S_{k+1} = \{q_k \mid q_k \in Q_k\ ,\ q_k \geq \theta_k\}$
- $Q_{k+1} = Q_k - S_{k+1}$ (set subtraction)
- $S = S \cup S_{k+1}$
- $k = k + 1$

Until $(\sigma_{k-1} - \sigma_k) \leq \epsilon$

Fig. 4.4 shows multiple iterations of the algorithm[6], demonstrating how it progressively extracts signals from the noise in a band. Adjusting the confidence interval changes the amount of the distribution's tail that is considered signal.

[6]The Gaussian curves have been generated from the statistics of the FM band (88-108 MHz) measurements after every iteration of the ROHT algorithm and not generated directly from the data itself. These curves illustrate the working of the ROHT algorithm.

One drawback with this algorithm is that not all distributions are Gaussian. Central limit theorem is applicable only when there are a large number of samples available such that the actual distribution converges to the Gaussian distribution. Similarly, the sample mean and standard deviation may not be equal to the actual statistics if sufficiently large number of samples are not available (Eqn. 3.31). The resulting uncertainties impose limitations on the detection of signals in low SNR conditions [25].

4.4.2 Modified Recursive Otsu's Algorithm

We have modified the recursive Otsu's algorithm [41] such that its working principle is similar to that of the ROHT algorithm. This algorithm differs from the ROHT algorithm in that at every iteration the threshold is now estimated using Otsu's algorithm instead of the one-sided hypothesis testing used in the ROHT algorithm.

4.5 Adaptive Thresholding

A global threshold may not be optimum for the entire set of measurements, especially when the noise and signal statistics vary across the data. In this case, classification can be done using local threshold values [40] that vary over the measurement set as a function of the sample values $M(f_i, t_j)$ and local statistics of the data. While a global threshold is represented as $T = C\{\mathbf{M}\}$, a local threshold is $T(f_i, t_j) = C\{\mathbf{M_{F,T}}\ ,\ S(\mathbf{M_{F,T}})\}$ where C represents an algorithm that estimates the local threshold $T(f_i, t_j)$ as a function of the measurement samples in $\mathbf{M_{F,T}}$ and some local statistic $S(\mathbf{M_{F,T}})$. In the case of a partitioned window approach, the measurement matrix $\mathbf{M}$ is divided into disjoint submatrices, each denoted by

$\mathbf{M_{F,T}}$, and the threshold is estimated independently for each subset.

4.5.1 Sliding Window Approach

In the sliding window approach, $\mathbf{M}$ is divided into overlapping submatrices. This approach offers the benefit of a partitioned window approach, as well as improves over it by:(i) Helping in reducing the biasing effect, and (ii) Reducing the appearance of ghost signals and ghost noise. Every iteration, the sliding window is moved by a certain step[7] along the data and it extracts a submatrix $\mathbf{M_{F,T}}$ from the data $\mathbf{M}$. A local threshold is computed and used to classify the measurement samples in $\mathbf{M_{F,T}}$. Since the classification is performed in overlapping submatrices, every element of $\mathbf{M}$ is classified more than once. However, each time it is classified in the presence of a different set of neighboring matrix elements. Suppose, there is a strong signal in the vicinity of a weak signal, there will be at least one submatrix that will contain the weak signal and not the strong signal and in that particular submatrix, the weak signal may be identified correctly. In this manner, both biasing effect and the appearance of ghost samples among the classified data can be reduced.

If the sliding window is of dimensions $l \times m$, then every measurement sample in $\mathbf{M}$ will be classified Y times, where $Y = l \cdot m$. The independent votes $v(k)$ (or classifications) of a measurement sample $M(f_i, t_j)$ can be combined to obtain the final vote as:

$$v_{net} = \frac{1}{Y} \sum_{k=1}^{Y} v(k) \tag{4.10}$$

Then a final decision on the classification of the sample can be made based on

[7]In our processing, we have taken this step to be 1 unit.

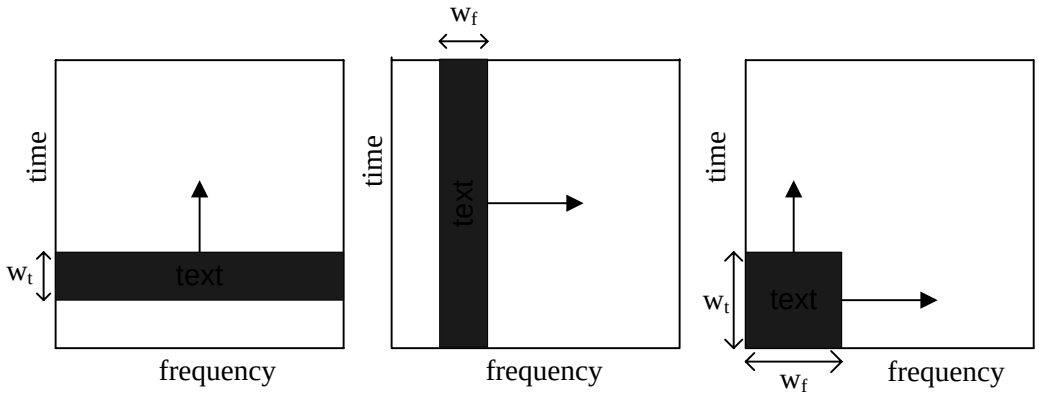

Figure 4.5. Various types of sliding windows moved along: (from left)(a) time, (b) frequency, and (c) both time and frequency.

a criterion specified on the combined vote. The different criteria are defined as:

- Simple combining criterion - According to this criterion, $M(f_i, t_j)$ is classified as signal only if $v_{net} > 0$.
- $n\%$ majority criterion - $M(f_i, t_j)$ is signal only if $v_{net} \geq n/100$. For instance, in the case of 50% majority, $v_{net} \geq 0.5$ must be satisfied.
- Super majority criterion - $M(f_i, t_j)$ is signal only if $v_{net} = 1$.

The different types of sliding windows that can be used are (see Fig. 4.5):

1. Strip of width w_t slided along time: The submatrix $\mathbf{M_{F,T}}$ at each iteration is defined as $\mathbf{M_{F,T}} = [M(f_i, t_j)]$; $f_i \in F$, $t_j \in T$, where F is the range of frequency channels in the sweep bandwidth $[F_{start} \ldots F_{stop})$. T is a subrange of the range $[T_{start} \ldots T_{stop})$ and it spans w_t sweep time instants every iteration. The total number of iterations will be $N_t - w_t + 1$ where N_t is the number of time instants within the measurement test duration.
2. Strip of width w_f slided along frequency range: The submatrix $\mathbf{M_{F,T}}$ at each iteration is defined as $\mathbf{M_{F,T}} = [M(f_i, t_j)]$; $f_i \in F$, $t_j \in T$, where F

Table 4.1. Spectrum measurement data sets.

Spectrum	Band	BW Resolution	Binwidth	Time Resolution
FM Band	88-108 MHz	10 kHz	10 kHz	19.1 sec
Upper TV Band	54-87 MHz	10 kHz	10 kHz	34.4 sec
Paging band	929-931 MHz	10 kHz	10 kHz	-
DTV band	638-668 MHz	10 kHz	10 kHz	36 sec
Analog TV band	198-128 MHz	10 kHz	10 kHz	36 sec
Cellular band	824-849 MHz	10 kHz	10 kHz	27.47 sec

is a sub range of $[F_{start} \ldots F_{stop})$ and it spans w_f frequency channels every iteration. T is the entire measurement time range $[T_{start} \ldots T_{stop})$. The total number of iterations will be $N_f - w_f + 1$ where N_f is the number of frequency channels within the sweep bandwidth.

3. A sliding window of dimensions $w_f \times w_t$: Here F and T are subranges of the full frequency range and time range. The total number of iterations will be $(N_f - w_f + 1) \times (N_t - w_t + 1)$.

4.6 Performance Evaluation of Processing Techniques

The processing techniques have been applied on real world spectrum measurements. The spectrum measurement data (see Table 4.1) was collected at the Information and Telecommunications Technology Center at the University of Kansas, Lawrence (rural environment), Kansas (USA), over a period of 24 hours. The data has been calibrated. While all the measurements have been collected with a preset attenuation of 10 dB, the cellular band measurements were made with 0 dB attenuation.

Three types of spectrum bands have been targeted: Bands with fixed channelization such as the FM and TV broadcast bands, and bands occupied by digital 'on/off' signals such as the 929-931 MHz paging band.

4.6.1 FM Radio Spectrum: 88-108 MHz

Data enhancement and Otsu's classification - All the data enhancement operations were performed on the FM spectrum measurements in a cascade manner, followed by classification using the Otsu's algorithm. Table 4.2 presents the results for various values of the data enhancement parameters. In this table, the following key has been used: *sc* and *nc* refer to signal and noise clipping, $L = x$ refers to a cascade of the spatial averaging filter and the Gaussian filter where the length of each of the filters is x, CS refers to contrast stretching, and p refers to the power law of the amplitude scaling. In this table, the parameters are listed in the order of their occurrence in the cascade of operations.

From Table 4.2, it is evident that the data enhancement operations have improved the performance of Otsu's classification algorithm. Case 2 shows improvement in the results as compared to Case 1, where the classification has been done without any data enhancement. It has been observed that, contrast stretching and the use of multiple stages of the filters did not improve the results. Case 3 is worth noting for the improvement in the miss rate although the false alarm rate is degraded by a small percentage as compared to Case 2. Fig. 4.6 shows a single time sweep of the FM band measurement data after data enhancement (Case 3). The measurements were time averaged before applying Otsu's algorithm and the results are shown as Cases 6-8. It is observed that time averaging improves the false alarm rate since the noise power is reduced. Overall, the best results have

Table 4.2. Results of data enhancement operations on FM broadcast spectrum (88-108 MHz) measurement data.

Case No.	Parameters	Miss (%)	FA (%)	Error (%)	Weighted error (%)
1	Otsu's algorithm, no enhancement	24.9279	16.5625	18.8212	7.5345
2	L = 4	21.7692	14.4307	16.4121	6.5762
3	sc = -55 dBm, nc = -98 dBm, L = 4	17.4581	18.0527	17.8921	5.9834
4	sc = -55 dBm, nc = -98 dBm, CS, p = 1, L = 4	17.4959	17.9333	17.8152	5.9790
5	sc = -64 dBm, nc = -98 dBm, p = 1, L = 4	13.8231	21.9614	19.7641	5.5772
6	Time averaging and classification	22.7778	11.7123	14.7000	6.5100
7	sc = -55 dBm, nc = -98 dBm, L = 4, time averaging	17.4074	15.4795	16.0000	5.6900
8	Time averaging, sc = -55 dBm, nc = -98 dBm, L = 4	18.1481	14.6575	15.6000	5.7700
9	sc = -55 dBm, nc = -98 dBm, L = 4, median filter L = 5	17.2852	17.9336	17.7585	5.9307

been obtained by applying the data enhancement techniques along with the time averaging technique before Otsu's classification. From Case 9 in Table 4.2, it is observed that there is little improvement in the results after applying median filtering to the classified FM band data.

Recursive thresholding - The ROHT algorithm was applied on the FM band data. The results of the ROHT algorithm with $\epsilon = 0.5$ for various values of the confidence intervals are illustrated in Figs. 4.7 and 4.8, and also tabulated in Table 4.3(a). From these figures, the tradeoff between miss rate and false alarm rate is clearly pronounced. From the plots we can also infer that good results

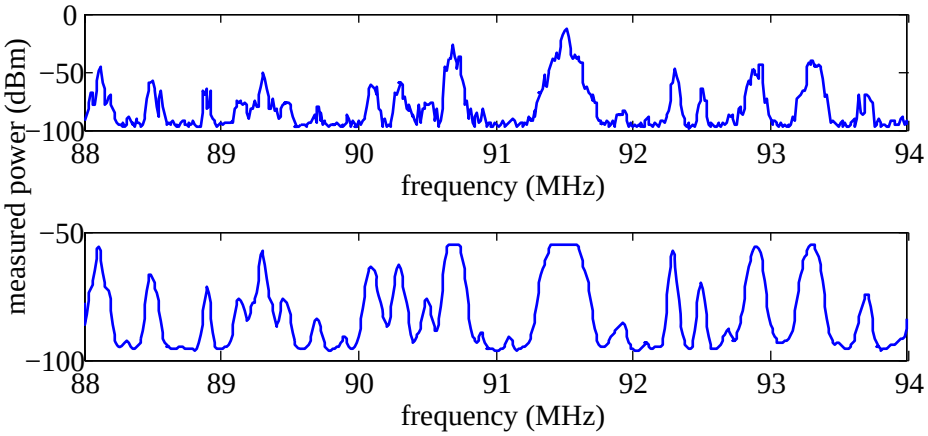

Figure 4.6. Power spectrum of FM band (88-94 MHz) before (top) and after data enhancement (bottom).

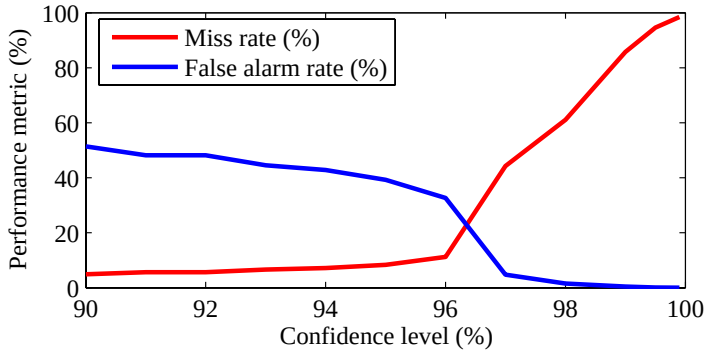

Figure 4.7. Results of ROHT classification with $\epsilon = 0.5$ applied on FM band data (without data enhancement):Tradeoff in miss rate and false alarm rate for various confidence levels.

can be obtained for the FM band by operating the algorithm at around 96 % confidence level beyond which the miss rate drastically increases. A similar trend has been observed for the case when $\epsilon = 0.05$ as shown in Table 4.3(b) and also when the ROHT algorithm was applied on enhanced data (with parameters as in Case 3 of Table 4.2). Table 4.3(c) shows the results of applying the modified recursive Otsu's algorithm on FM measurement data.

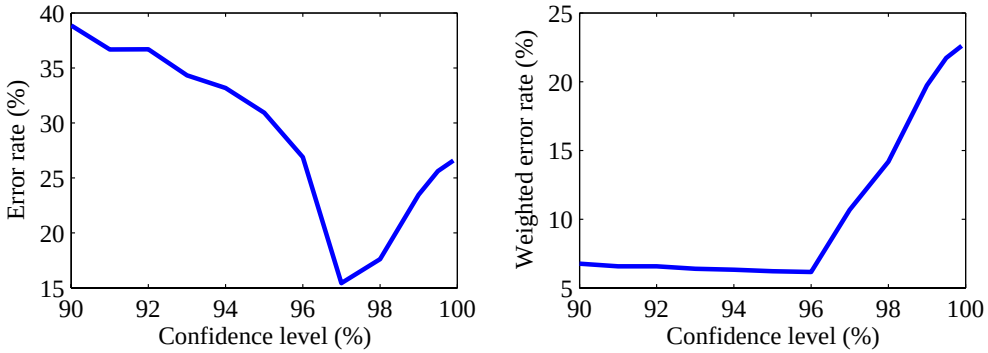

Figure 4.8. Results of ROHT algorithm with $\epsilon = 0.5$ applied on FM band data (without data enhancement) for various confidence levels: error rate (left), and weighted error rate (right).

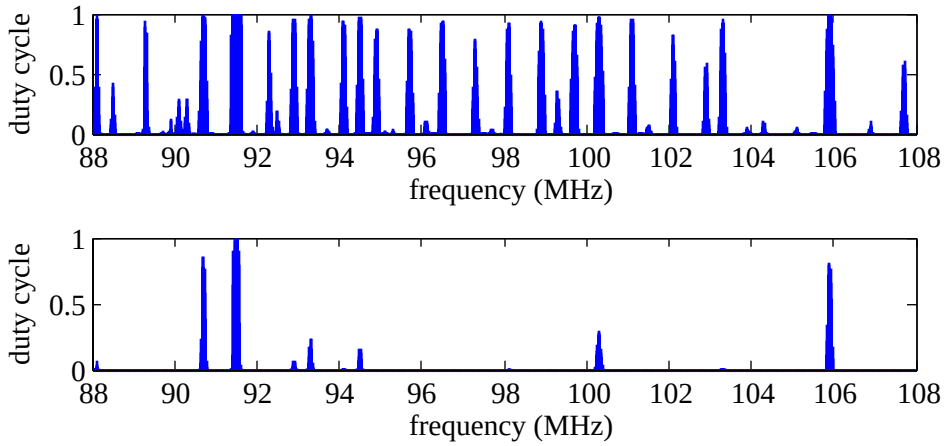

Figure 4.9. Results of ROHT algorithm with $\epsilon = 0.5$ applied on FM band data (without data enhancement): Duty cycle plots for the confidence intervals of 98 % (top) and 99.5 % (bottom).

Adaptive thresholding- The sliding window approach was used along with the Otsu's algorithm, as well as the ROHT algorithm. Table 4.4 presents the classification results for various dimensions of the sliding window while using Otsu's algorithm. The simple combining operator criterion has been.

Interesting trends are observed from the results in Table 4.4. There was little improvement in the results by using the sliding window moved across time, as compared to the case where sliding window has not been used. The reason behind

Table 4.3. Performance evaluation of recursive thresholding applied on FM band (88-108 MHz) data: (a) ROHT algorithm with $\epsilon = 0.5$, (b) ROHT algorithm with $\epsilon = 0.05$, and (c) Recursive Otsu's algorithm . Data enhancement with sc = -55 dBm, nc = -98 dBm, and L =4 has been performed to the data before thresholding.

Confidence Level (%)/ϵ	Miss (%)	FA (%)	Error (%)	Weighted error %)
(a) $\epsilon = 0.5$				
90	1.6985	50.5119	37.3323	5.9209
94	5.9117	34.5668	26.8299	5.1418
95	8.3757	29.6694	23.9201	5.1710
99.9	99.9702	0.0909	27.0583	22.9531
(b) $\epsilon = 0.05$				
90	0.0993	74.6542	54.5244	8.1974
92	0.1535	71.8168	52.4677	7.8992
95	0.5766	61.4434	45.0094	6.8604
99.9	99.9702	0.0909	27.0583	22.9531
(c) Recursive Otsu				
$\epsilon = 0.5$	0.0348	80.9202	59.0812	8.8688
$\epsilon = 1.5$	0.4211	64.1711	46.9586	7.1234

this trend can be explained as follows: All the sweeps of measurement data possess similar statistical properties since the occupancy in the FM band (transmit power, and noise) does not vary significantly over time. Thus, by increasing the window size, the number of sweeps of data available for the local threshold estimation increases, however it results in increased redundancy in the data which does not improve the classification results much. As before, there was no change in results with increasing window size, when using the window moved along time along with the ROHT algorithm.

In the second case, where a window is moved along the frequency range, as the window width is increased to 25 there is an improvement in the false alarm rate and miss rate. As seen in Fig. 4.10, a similar trend has been observed when the

Table 4.4. Results of applying Otsu's algorithm on enhanced ($sc = -55dBm$, $nc = -98dBm$, $L = 4$) FM data (88-108 MHz) along with sliding window approach

Type of Sliding Window	Window size w	Miss rate (%)	FA rate (%)	Error rate (%)	Weighted error rate (%)
No sliding window	-	17.4581	18.0527	17.8921	5.9834
Strip of width	5	17.3731	17.6063	17.5433	5.9150
w_t slided	10	17.3659	17.6113	17.5450	5.9139
across time	15	17.3593	17.6193	17.5491	5.9133
	25	17.3431	17.6393	17.5593	5.9117
Strip of width	5	11.9233	41.9454	33.8394	7.3294
w_f slided	10	5.8415	42.3225	32.4726	5.9749
across frequency	15	6.0272	41.6021	31.9969	5.9387
	25	5.7628	38.2955	29.5117	5.5159
$w_t \times w_f$	5	0.2350	90.2441	65.9416	9.9357
window	10	1.1385	74.3384	54.5744	8.4013
	15	3.0132	66.3498	49.2489	7.9568
	25	3.4771	50.7768	38.0058	6.3580

sliding window approach was used along with the ROHT algorithm on enhanced data ($sc = -55dBm$, $nc = -98dBm$, $L = 4$). There are two reasons for this trend to occur. First, an increase in the window size would lead to an increase in the number of samples available for estimating the local threshold and thus the classification results improve. It should be noted that the measurement samples along frequency do not represent redundant data since the signal and noise powers can vary drastically from station to station. Second, by increasing the window size beyond 20, the incidence of ghost signals and ghost noise is decreased. Every FM station band occupies 150 kHz. Thus, in the FM band data, the measurements

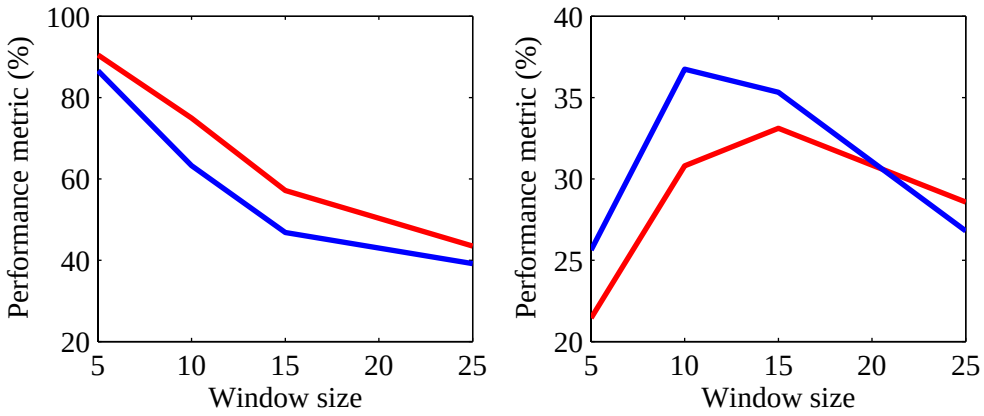

Figure 4.10. Results of ROHT algorithm used with sliding window approach on enhanced FM band data: Miss rates (left), and false alarm rates (right). Red curve represents the case of a strip slided along frequency and blue curve represents the case when square window has been used.

collected at 15 consecutive frequency steps which coincide with a station band, represent samples of signal power. The measurements from the adjacent guard bands and buffer spaces represent samples of noise power. If a window of width less than 15 is used, then as the window is slided along frequency, there will be instances when the window overlaps over data consisting of only signal measurements (measurements from occupied station band) and no noise measurements. In this case, the classification of the windowed measurements results in the occurrence of ghost noise. In a similar manner, the classification of measurements from the vacant portion of the FM spectrum results in the appearance of ghost signals.

In the third case, where a square window has been used, the miss rate increases and the false alarm rate decreases with increase in window size.

Figs. 4.11 and 4.12 shows the image of the measurement data before and after classification. The measurements have been enhanced before classification with $sc = -55dBm$, $nc = -98dBm$, and $L = 4$.

It is evident from Fig. 4.12 that almost all the signals have been correctly

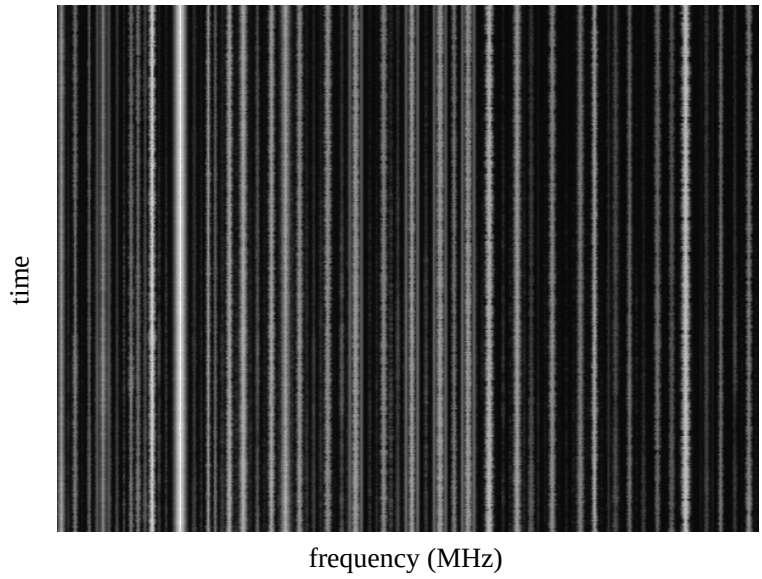

Figure 4.11. Spectrum image of FM band (88-108 MHz) measurement data.

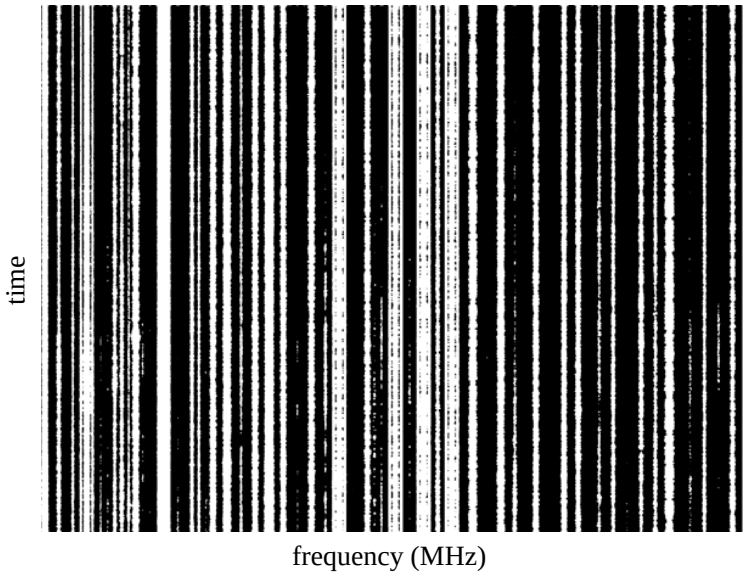

Figure 4.12. Spectrum image of FM band (88-108 MHz) measurement data after classification using sliding window along time of width 25 and Otsu classification.

classified excluding some weak signals. In addition, in the case of the strong signals the sidelobes have been wrongly identified as signal. As a result the strong signals appear wider in Fig. 4.12 as compared to Fig. 4.11.

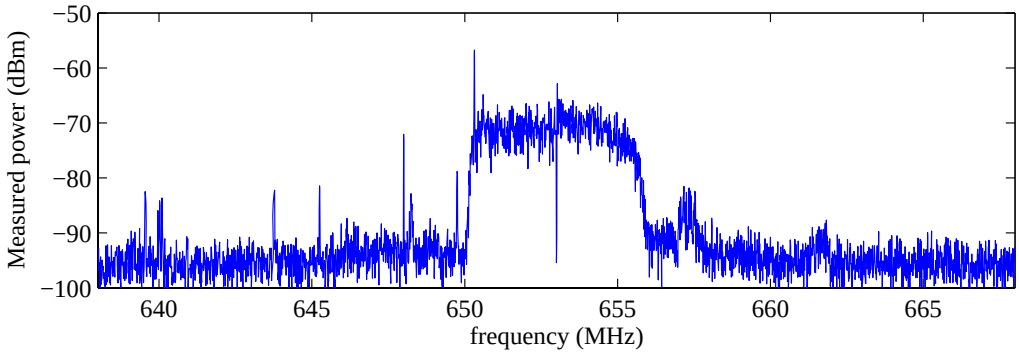

Figure 4.13. Instantaneous power spectrum of digital TV band (638-668 MHz).

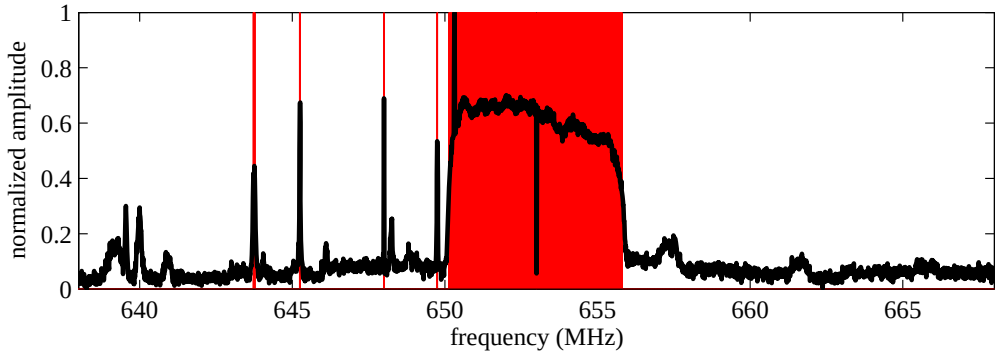

Figure 4.14. Digital TV band (638-668 MHz): Otsu's classification (red) of averaged measurements (black).

4.6.2 Digital Television Band: 638-668 MHz

A single time snapshot (instantaneous power spectrum) of the digital TV band (638-668 MHz) is shown in Fig. 4.13. In this figure, the channel 44 signal can be observed at 650-656 MHz. It is also observed that there is a lot of noise power variation in the instantaneous power spectrum, which can degrade the accuracy of classification algorithms. A comparison of Fig. 4.13 and Fig. 4.14 clearly shows that there is significant reduction in the noise variations by time averaging the data over 25 time snapshots. The time averaged data was then classified using Otsu's algorithm.

Fig. 4.15 shows the result of applying ROHT (95% confidence level and $\epsilon = 1.5$)

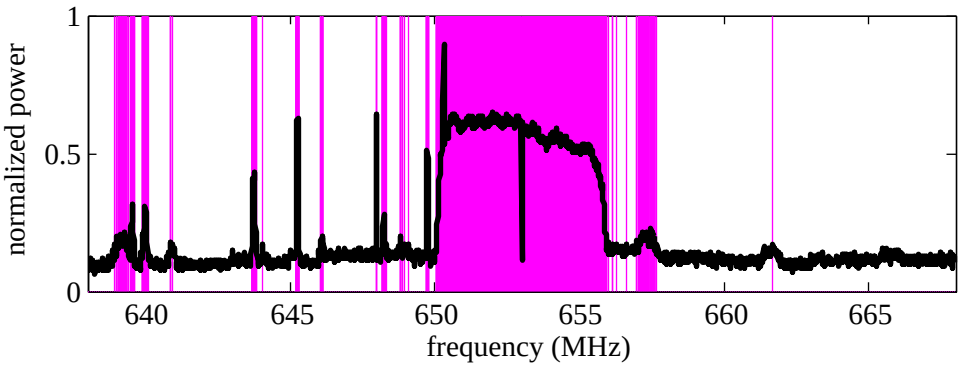

Figure 4.15. Digital television band (638-668 MHz): mean power spectrum (black) and its ROHT (95% confidence level and $\epsilon = 1.5$) classification (magenta).

algorithm on time averaged data. Similar classification result has been obtained by using the recursive Otsu's algorithm on the mean power spectrum. It has also been observed that the sliding window approach (width = 1000, moved along frequency) improved the results of Otsu's algorithm.

4.6.3 Analog Television Band: 198-228 MHz

Measurements were collected from the analog television spectrum (198-228 MHz) at a distance of 200 ft from the channel 13 (210-216 MHz) television tower[8]. The presence of spurious signals, intermodulation products, and the raised noise level is clearly seen from Fig. 4.16. As a result, the noise level varies with frequency and a sliding window approach to classify the data can produce good results.

The recursive Otsu's algorithm gave very poor false alarm rate even after time averaging. As compared to Otsu's algorithm, better results were obtained by time averaging the data followed by Otsu's classification (see fig. 4.17). Time averaging followed by ROHT algorithm with 96% confidence level and $\epsilon = 0.5$ (see fig. 4.18) gave good results.

[8]More details on the measurement campaign are provided in Section 5.1.1

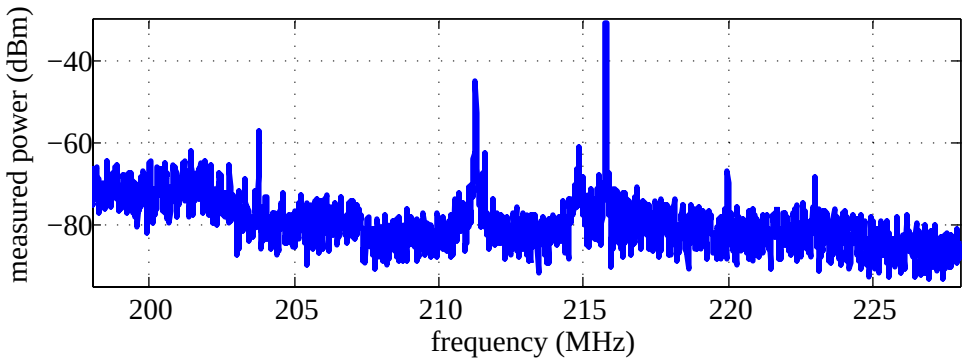

Figure 4.16. Instantaneous power spectrum of analog television band (198-228 MHz).

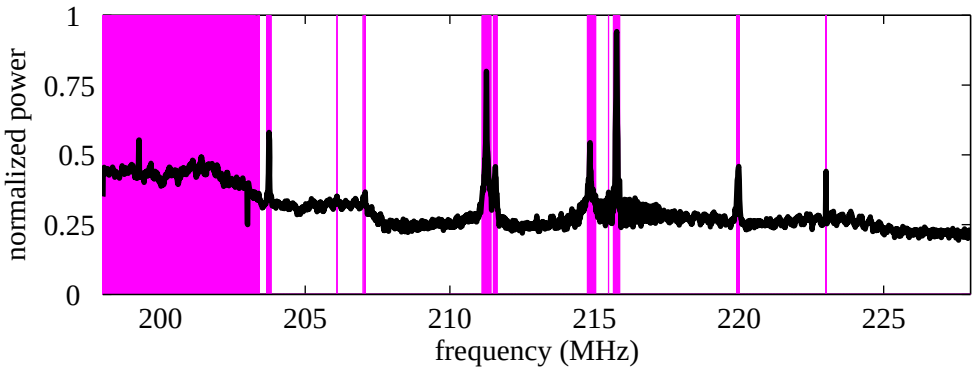

Figure 4.17. Analog television band (198-228 MHz): mean power spectrum (black) and Otsu classification of time averaged data (magenta)

Otsu's algorithm was used with a sliding window moved along frequency (window size = 2000) on time averaged analog TV band data. By using a 20% combining criterion decent results were obtained (see fig. 4.19). By using a sliding window moved along time and the square sliding window with window sizes of 5 - 25 on refined data there were several false alarms.

Compared to Fig. 4.18 and Fig. 4.19, the best results have been obtained by using the ROHT algorithm (96% confidence, $\epsilon = 0.5$) along with a sliding window of size 2000 that is moved along frequency on time averaged data (see fig. 4.20). In this case, 20% combining criterion has been used. However, the same sliding

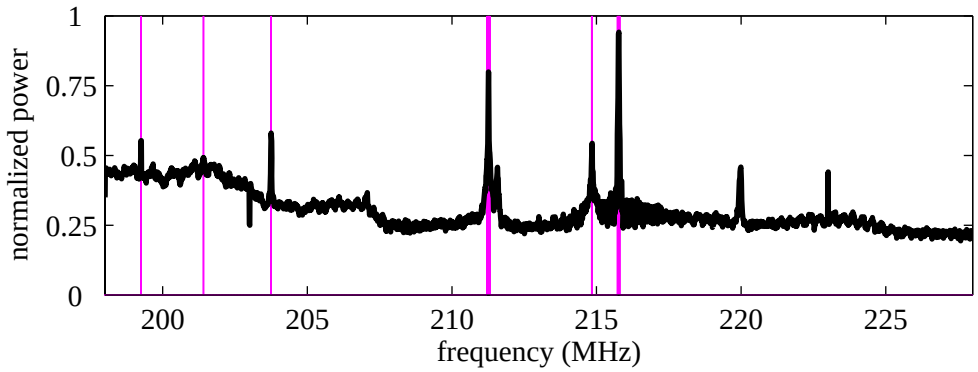

Figure 4.18. Analog television band (198-228 MHz): mean power spectrum (black) and its classification (magenta) using ROHT with 96% confidence and $\epsilon = 0.5$.

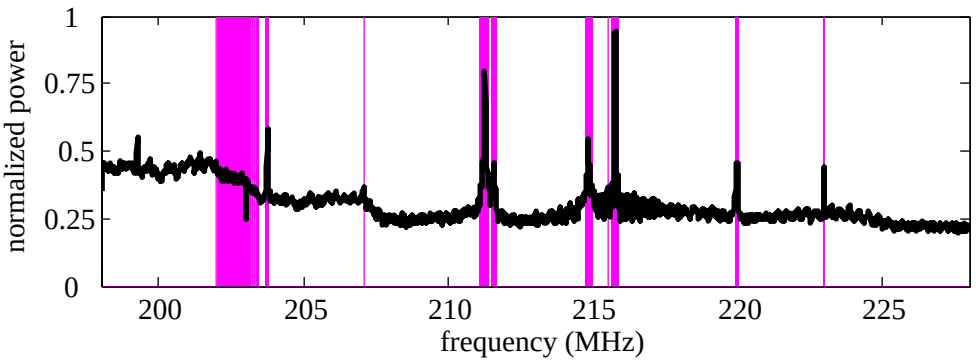

Figure 4.19. Analog television band (198-228 MHz): mean power spectrum (black) and its Otsu's classification (magenta) using sliding window width 2000 moved along frequency (20% criterion used).

window when used with the recursive Otsu's algorithm yielded several false alarms.

4.6.4 Paging Band: 929-931 MHz

Otsu's algorithm was applied to the measurements collected through a single sweep across the 929.4-930 MHz band. Otsu's algorithm using a non-overlapping window size of 30 was also applied on the data and this produced slightly improved results. As seen from Fig. 4.21, the signal detection at the edges of the signal has been improved. This is due to the localization of the noise level estimation and moreover, by using a partitioned window approach the influence of the stronger

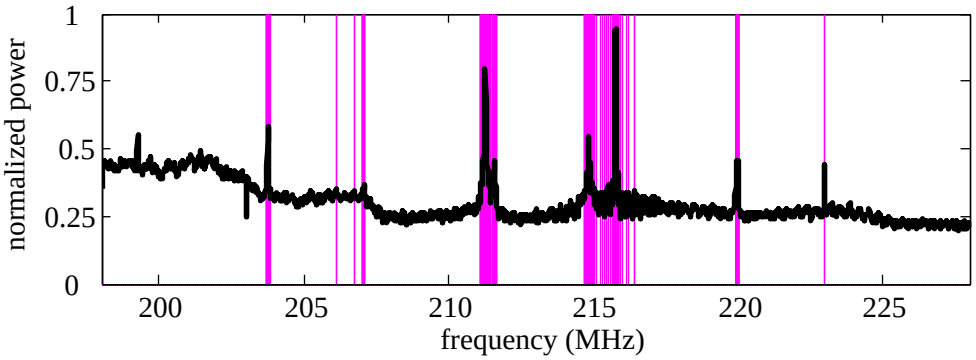

Figure 4.20. Analog television band (198-228 MHz): mean power spectrum (black) and its ROHT (96% confidence, $\epsilon = 0.5$) classification (magenta) using sliding window width 2000 moved along frequency (20% criterion used).

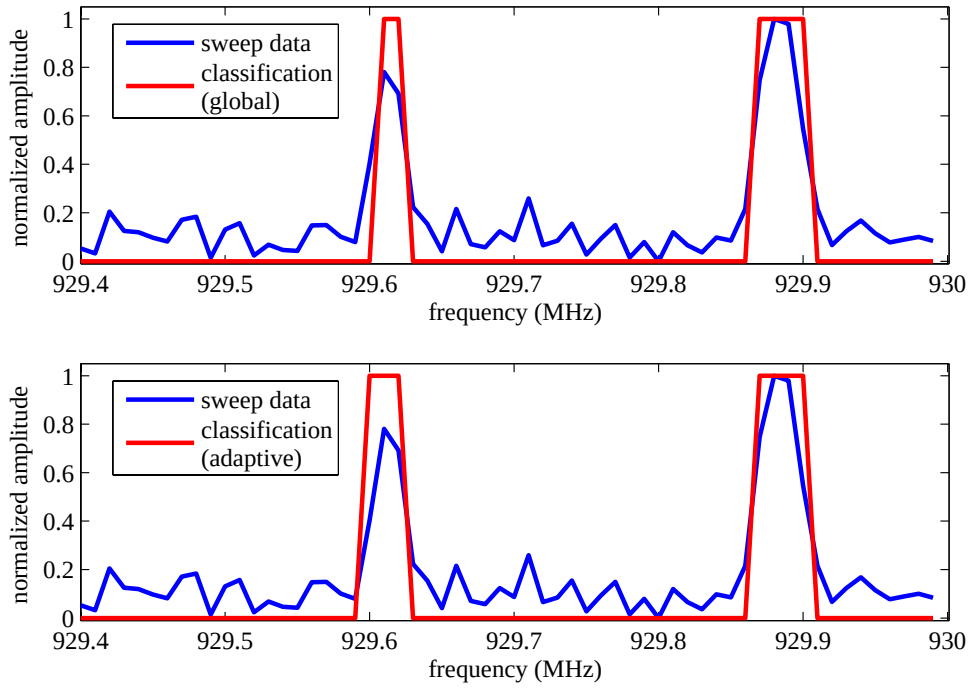

Figure 4.21. Otsu's algorithm applied to paging band (929.4-930 MHz) data using global threshold (top) and using local threshold (bottom).

signal on the weaker signal is eliminated thereby enabling improved threshold estimation.

Measurements have been collected from the 929-931 MHz paging band. Fig. 4.22

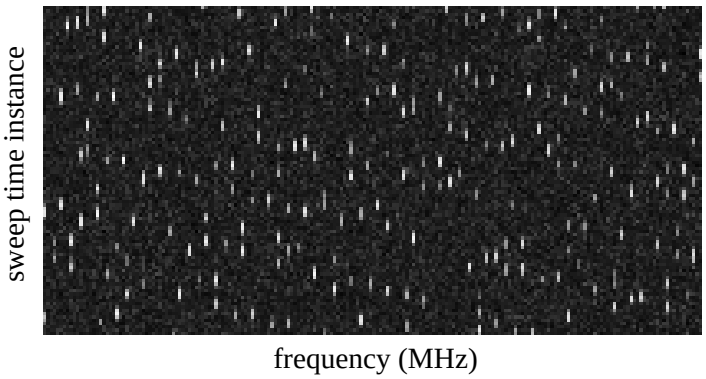

Figure 4.22. Spectrum image of paging band (929-931 MHz) data.

shows the intensity plot of the paging band data. In the figure, the dots represent the signals that go 'on' and 'off' over time. When applied on this data, the ROHT algorithm with 99.5% confidence level and $\epsilon = 1.5$ gave the same performance as the Otsu's algorithm. In this data, the dynamic range of the signals is small thus, a higher confidence level and higher epsilon has been used in order to avoid false alarms. This shows that the confidence level and epsilon are dependent on the dynamic range of the signals present in the data. The recursive Otsu's algorithm generated a lot of false alarms.

In order to demonstrate the feature extraction tools, the measurements along channel 930.04 MHz were first classified using Otsu's algorithm (see fig. 4.23). By analyzing the classified data, the mean 'on' time of the signals occupying the channel was computed as $3.8T_s$ where T_s is the time resolution of the measurements.

4.6.5 Cellular Band: 824-849 MHz

The measurements collected from the cellular band (824-849 MHz) were classified using the ROHT algorithm with 96% confidence and $\epsilon = 0.5$. By applying

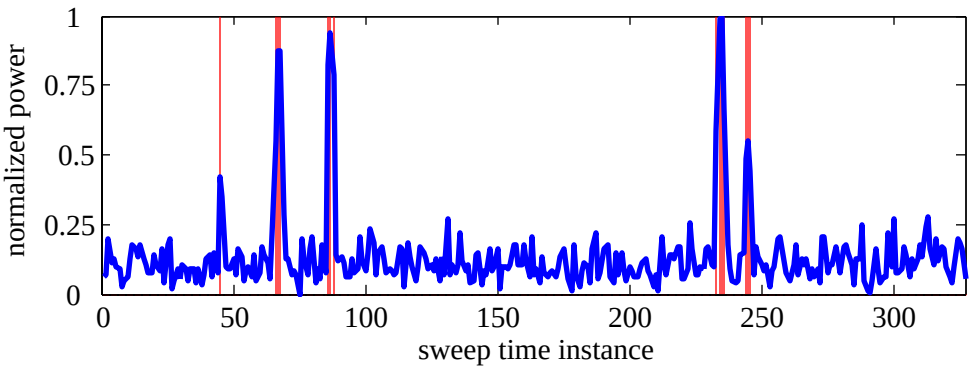

Figure 4.23. Measurements from channel 930.04 MHz classified by Otsu algorithm.

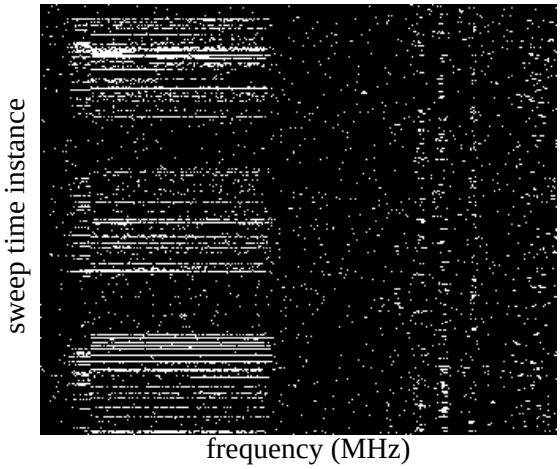

Figure 4.24. Spectrum image of cellular band (835.71-838.96 MHz) after ROHT (96% confidence and $\epsilon = 0.5$) classification.

median filtering to the classifications, most of the ghost signals that appear as grains in Fig. 4.24 were removed (see Fig. 4.25).

Figs. 4.26 and 4.27 show a sweep of the data before and after the classification. Otsu' algorithm gave nearly 100% miss detection even when used with noise filtering. However, when applied to the averaged spectrum measurements it gave good results (see Fig. 4.28). The recursive Otsu's algorithm gave very high false alarms even after averaging the measurements.

While using a sliding window along with Otsu's algorithm on time averaged

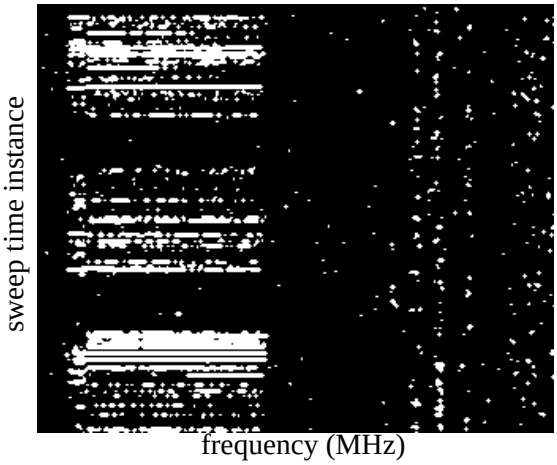

Figure 4.25. Spectrum image of cellular band (835.71-838.96 MHz) after ROHT (96% confidence and $\epsilon = 0.5$) classification and median filtering.

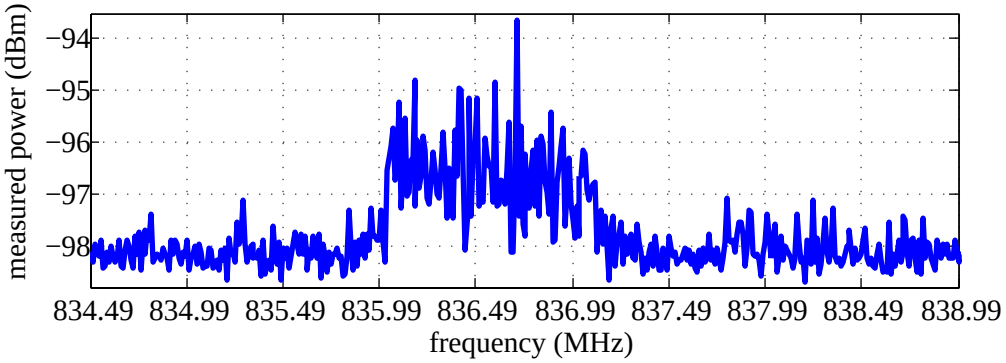

Figure 4.26. Instantaneous power spectrum of cellular band (834.49-838.99 MHz).

data, a window size of 500 gave very high false alarm rate while a window size of 1000 gave decent results. This indicates that the window size depends on the occupancy of the band under consideration. A band with low occupancy is prone false alarms when classification is done. However, this can be reduced by using a sliding window approach with a suitable window size such that every subset of the data that is extracted by the sliding window contains signals.

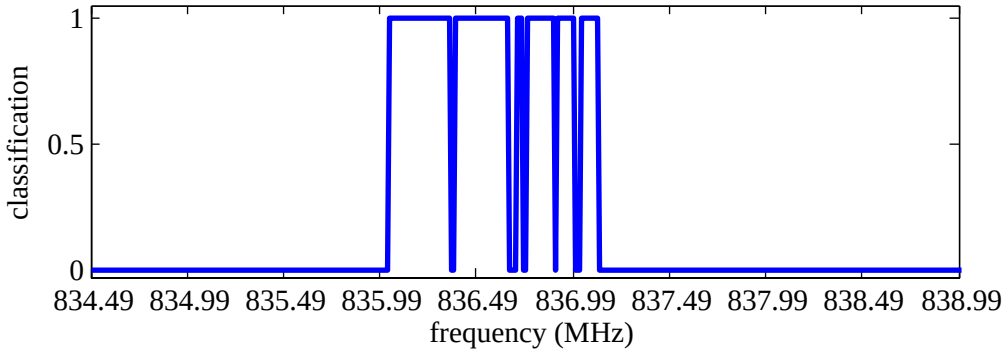

Figure 4.27. Cellular band (834.49-838.99 MHz): Classification of data after ROHT (96% confidence and $\epsilon = 0.5$) classification and median filtering.

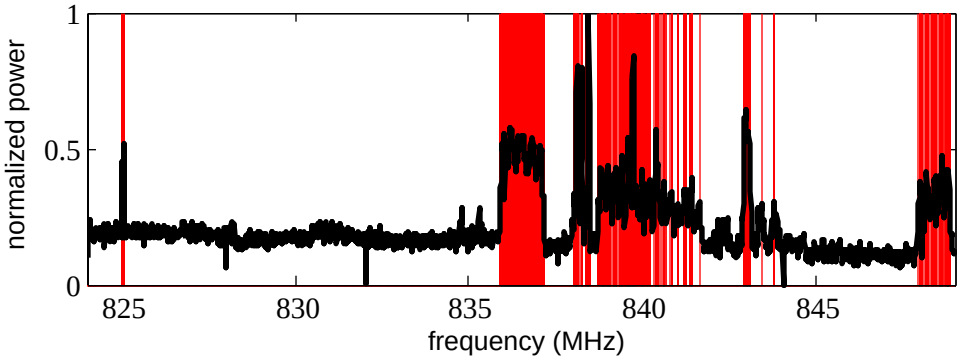

Figure 4.28. Cellular band (824-849 MHz): (black) mean power spectrum and its classification (magenta) using Otsu algorithm.

4.6.6 Summary of Results

The results achieved on the FM band measurements were not as good as the results obtained with the other data sets. In fact, the FM band data represents the worst case scenario due to several reasons: The FM band data was collected at a close proximity from the transmission tower for the radio station KJHK 90.7 MHz. Hence the presence of intermodulations and LO sideband noise is expected. In fig. 4.29, sidelobes for FM signals at 91.5 MHz and 96.5 MHz are clearly seen. The sidelobes get classified as signals thereby resulting in false alarms. In addition, there are weak signals at frequencies 101.75 MHz and 106.5 MHz which occur close

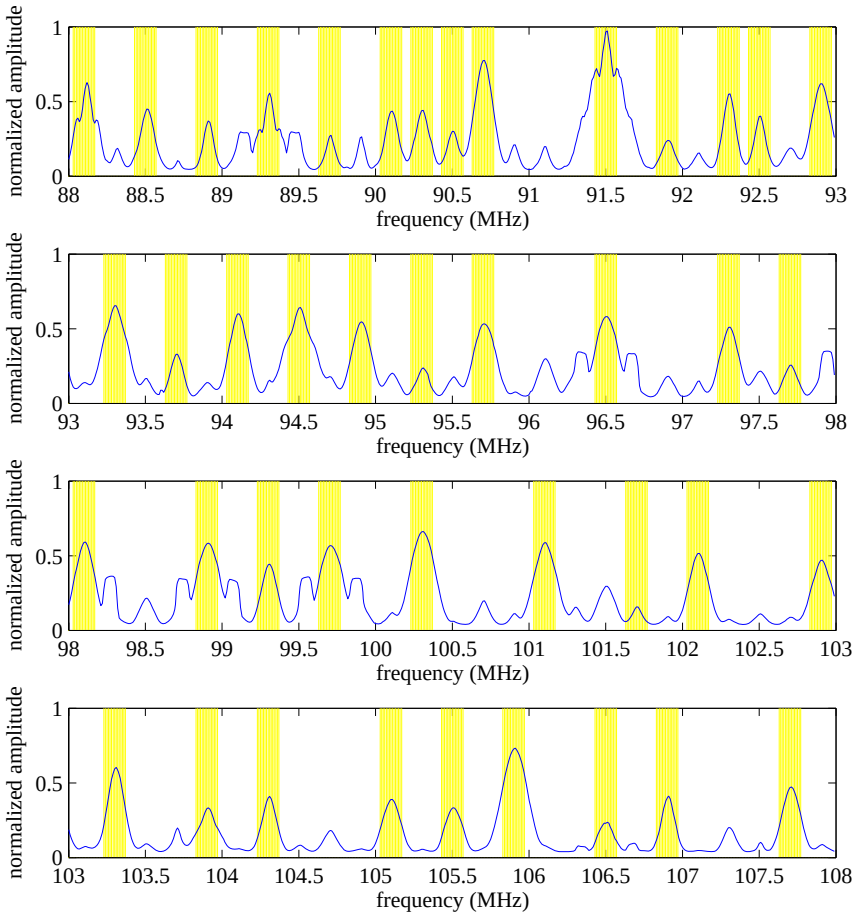

Figure 4.29. The time averaged FM band (88-108 MHz) measurements (blue plot) shown along with the ground truth (yellow plot).

to the noise level and these may not be identified by the classification algorithms. Furthermore, the dynamic range of the measurements is high.

Table 4.5 presents the comparison among the various processing techniques used on the measurements. Time averaging improved the performance of the classification algorithms when applied on the bands with fixed channelization such as the FM band and the television bands. However, time averaging did not perform well when applied to the paging band due to the random nature of the channel occupancy in this band. The ROHT and Otsu's algorithms performed well on all the bands with the help of time averaging. The recursive Otsu's algorithm gave

Table 4.5. Summary of results.

Bands	Time averaging	Otsu's algorithm	ROHT algorithm	Rec. Otsu algorithm	Sliding window
FM band	√	√	√×	×	√
Analog TV	√	√	√	×	√
Digital TV	√	√	√	√	√
Paging band	×	√	√	×	NA
Cellular band	√	√	√	×	√×

a high false alarm rate with most of the bands. The sliding window approach improved the results of the classification algorithms. However, in the case of the cellular band which is sparsely occupied, smaller window sizes of less than 1000 gave high false alarms. This exposed a drawback that the window size has to be set according to the occupancy of the target band. The sliding window was not required for the paging band since the noise level in the band is uniform.

4.7 Chapter Summary

In this chapter, we have identified the challenges to the energy detection of signals from spectrum measurement. The proposed techniques have been applied on spectrum bands with different types of spectrum activity. Using suitable techniques such as data enhancement, optimum thresholding, recursive thresholding, and adaptive thresholding, we have shown through results that the performance of the classification algorithms can be improved.

Chapter 5

Applications of Spectrum Survey Framework

In this chapter, we demonstrate two proposed applications of the SSF that are pertaining to DSA networks. In Section 5.1, we present a feasibility study of unlicensed DSA in the presence of television signals in the underutilized television bands. The study has been divided into two phases with the overall objective of determining the conditions by which both television and unlicensed transmission can co-exist with each other in the same spectral bandwidth.

The spectrum might contain different types of signals, individually requiring scans with different resolutions. Conventional non-adaptive wideband spectrum sensing approaches could potentially be inefficient since they employ the same scanning resolution. In Section 5.2, we present a novel spectrum sensing framework that can adapt its parameters across the spectrum of interest according to the characteristics of its occupancy. The SSF can be incorporated into this framework in order to support the radio in learning its wireless environment in an efficient manner. We also propose a dynamic scheduling algorithm for spectrum

sensing. This algorithm allocates different time resolutions to different portions of the spectrum.

5.1 Feasibility Study of Unlicensed Cognitive Radio Operation in TV bands

Although several methods have been proposed for avoiding harmful interference to the TV receivers [24, 42–44] there is still a debate on the non-interfering operation of unlicensed devices. There are many who claim that the unlicensed devices will cause harmful interference to the primary users [45], while others debate that DSA can be performed in a transparent manner [46]. Proponents of the DSA approach favor the TV bands for DSA due to several reasons: There is substantial amount of unused spectrum available for DSA and, in addition, the propagation properties in these frequency ranges, such as low propagation attenuation, are beneficial for long range mobile and line-of-sight (LOS) communications [44]. Moreover, the fixed channel allocations resulting in deterministic usage patterns in these bands are favorable for accurate spectrum sensing [23].

However, there are several challenges for the unlicensed usage of these bands. The presence of strong TV signals near the secondary user can lead to the generation of intermodulation products, and saturation effects in the vacant bands [47]. The unoccupied portions of the spectrum might also be licensed for other purposes, such as public safety [48].

In order to provide input to these debates and assess the challenges to DSA, the feasibility of unlicensed device operation in the TV spectrum needs to be studied. In addition, there is a need for a standardized procedure to measure the effects of interference from unlicensed users on the TV signals [49]. Our feasibility

study has been split into two phases:

- Phase 1 - Test viability of cognitive radio operation near TV transmitters by determining the interference from TV transmissions to the unlicensed user (cognitive radio), and
- Phase 2 - Viability of TV co-existing with cognitive radios by studying the impact of cognitive radio transmissions on TV reception.

In the first phase, we measure spurious signals in the vacant TV bands and evaluate the error robustness of OFDM-based transmissions (secondary signals) in the unoccupied portions of the spectrum at varying distances from the TV station. In the second phase, we determine the impact of increased interference from unlicensed devices directly to the TV received video quality. While the two scenarios have been examined in reference [50], only the first scenario is being discussed in this thesis[1].

An investigation studying the operation of public safety transmissions in television bands, when both digital and analog television were present, was conducted as described in reference [51]. Although several insights were obtained regarding the interaction between licensed and unlicensed transmissions, the investigation did not quantify the impact on the video quality of the television signal nor the effects of operating unlicensed devices at close distances to television transmitters. Reference [52] describes an experiment conducted on a discontiguous OFDM testbed to show that unlicensed device operation in discontiguous bands is possible. Unlike this study, our feasibility study is achieved by conducting a combination of field measurements and computer analysis.

[1]Dr. Alexander Wyglinski, Ted Weidling, and Rory Petty of the University of Kansas were also involved in planning this study.

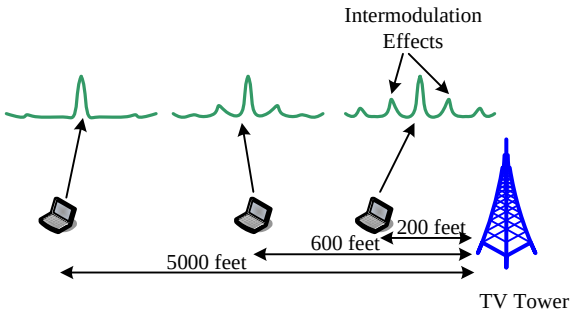

Figure 5.1. Measurements collected at distances 200 feet, 600 feet, and 5000 feet from the WIBW television station (210-216 MHz, 650-656 MHz) tower.

5.1.1 Measurement Campaign

The station broadcast tower for the WIBW television channel located west of Topeka, Kansas, USA, was selected because it was in a spectrally quiet location, and its analog and digital stations were separated by over 400 MHz. This allowed for the measurements taken to include unused surrounding channels such that any intermodulation or saturation effects can be clearly identified.

Fig. 5.2 displays the measurement equipment. An omni-directional discone antenna is connected to the input port of a IFR-2398 spectrum analyzer. The Spectrum Miner software was installed on a laptop computer and it controls the spectrum analyzer over either an RS-232 or a general purpose interface bus (GPIB) connection.

The measurements were collected in the TV channels 13 (analog television, 210-216 MHz, ERP [2]: 316 kW) and 44 (digital television, 650-656 MHz, ERP: 193 kW). As shown in Fig. 5.1, the measurements were collected[3] at increasing line-of-sight distances of 200, 600, and 5000 feet from the WIBW TV tower. The

[2]ERP is an abbreviation for effective radiated power.

[3]Dr. Alexander Wyglinski, Ted Weidling, and Rory Petty of the University of Kansas helped me by collecting these measurements.

Table 5.1. Measurement site GPS coordinates.

Site	Coordinates	Elevation
A	39° 00.408 N, 96° 02.956 W	1298 ft.
B	39° 01.565 N, 96° 02.914 W	1090 ft.
C	39° 05.261 N, 96° 03.169 W	1000 ft.

GPS coordinates of the measurement locations in Topeka are listed in Table 5.1.

In order to study the impact of intermodulation and saturation effects on secondary transmissions, 12 MHz of bandwidth on either side of the TV channels were also recorded. Thus, the total bandwidth spanned by each measurement set is 30 MHz with a spectral resolution of 10 kHz. At each measurement site, 25 sweeps were recorded over the 30 MHz bandwidth for both the analog and digital TV channels. These measurements can be used in an evaluation of the cognitive radio performance as well as help set guidelines on the effective operation of cognitive radios in the vicinity of TV stations.

5.1.2 Viability of Unlicensed Device Operation Near TV Transmitters

In order to analyze the impact of primary user transmissions over the secondary transmissions, we considered the vacant bands in the vicinity of the TV band and used it to transmit OFDM symbols. The measured power spectrum was averaged over 25 sweeps in order to obtain the average noise power in the frequency channels adjacent to the occupied TV bands. The plots of the average power spectrum measured at increasing distances from both the analog TV and the DTV towers are shown by Figs. 5.3 and 5.4. This computed average noise power was used for the OFDM simulation.

The analog television measurements collected at 200 ft from the tower were averaged and classified using Otsu's algorithm with sliding window moved along

Figure 5.2. Picture of the field measurement setup showing the discone antenna mounted on a tripod stand, laptop and spectrum analyzer (placed inside vehicle).

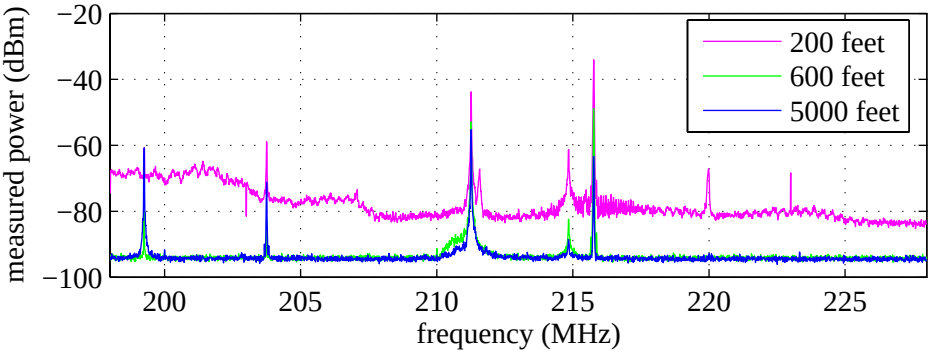

Figure 5.3. Mean power in the analog TV Spectrum with channel 13 (210-216 MHz), at various distances from TV tower.

frequency (window size = 1300). The feature extraction tools were applied to the classified data and the mean noise level was found to be -75.06 dBm with a standard deviation of 4.6311. The maximum and minimum noise power levels

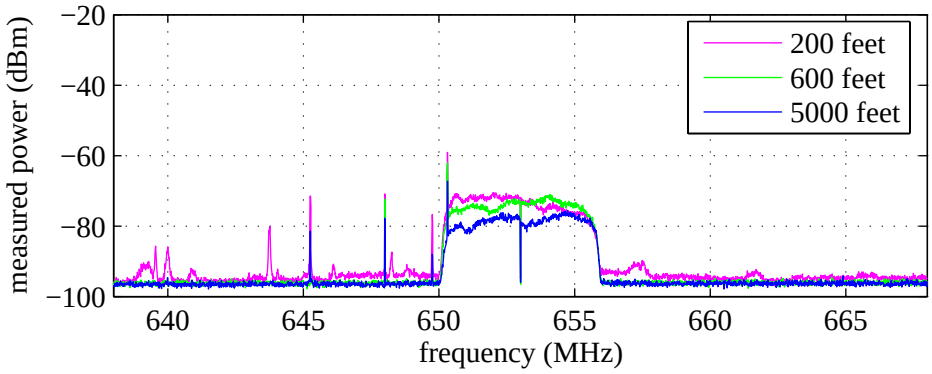

Figure 5.4. Mean power in the digital TV Spectrum with channel 44 (650-656 MHz), at various distances from TV tower.

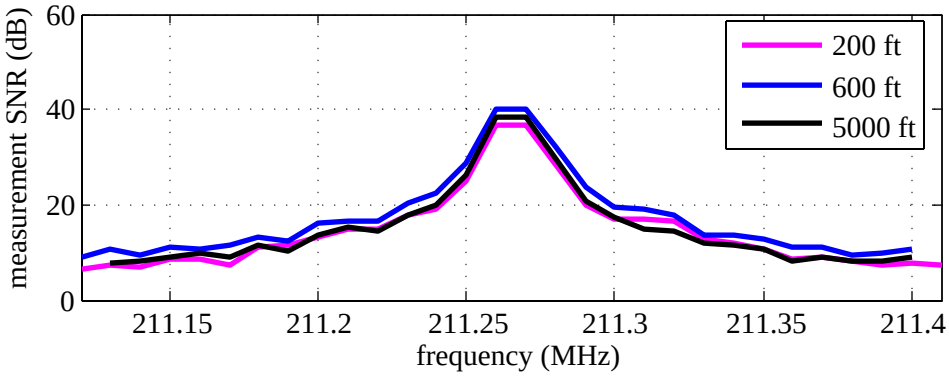

Figure 5.5. SNR of analog TV Signal at 211.13 MHz, at various distances from TV tower.

were -60.65 dBm and -85.29 dBm. Analysis was also done in order to plot the SNR of the television signal at 211.13 MHz at various distances from tower (see Fig. 5.5). The bandwidth of this signal was found to be 0.31 MHz.

5.1.3 Simulation Setup

For the simulation[4], we considered a QPSK modulated OFDM system having 512 subcarriers each with a bandwidth of 10 kHz. The OFDM transmission utilizes an overall bandwidth of 5.12 MHz adjacent to the occupied TV band. Since

[4]The simulation has been performed using the OFDM model developed by Rakesh Rajbanshi at the University of Kansas.

the measurement sites were located in an open area with less obstructions in the outdoor channel, we assumed a negligible effect of multipath signals on the system performance. A quasi-analytical error rate estimation approach was used in the monte carlo simulation in order to determine the bit error rate (BER) performance of the OFDM system. In our simulation, the probability of bit error was computed as [53]:

$$P_{be}(n_i) = \mathcal{Q}\left(\sqrt{\frac{4\,S}{N(n_i)}}\right), \tag{5.1}$$

$$P_{avgbe} = \frac{1}{512}\sum_{i=1}^{512} P_{be}(n_i), \tag{5.2}$$

where S is the transmit power at each of the subcarriers, $N(n_i)$ is the average noise power at frequency n_i, $P_{be}(n_i)$ is the probability of bit error for the i^{th} subcarrier at frequency n_i, and P_{avgbe} is the probability of bit error for the OFDM transmission computed by averaging the error probabilities over all the subcarriers. For a certain signal-to-noise ratio SNR, S is computed as:

$$S = Minimum\,\{N(n_i)\} \times SNR,\ 1 \leq i \leq 512. \tag{5.3}$$

5.1.4 BER Performance Results

In Figs. 5.3 and 5.4, the presence of spurious signals and the average noise level in the bands adjacent to the occupied TV bands 210-216 MHz and 650-656 MHz is of particular interest to us. The spikes of narrow bandwidth in the plots might represent the channel activity of other licensed users such as public safety. The spurious signals, which might be other licensed users such as public safety radios or intermodulation products of the TV transmissions, can potentially interfere

with the secondary OFDM transmissions and degrade the BER performance.

The measurements collected at a close proximity of 200 feet from the analog TV tower contain strong spurious signals. Moreover, the noise floor is quite high due to the saturation effects of strong TV signals at the receiver. At distances of 600 and 5000 feet, the spurious signals are weaker, and the average noise level is lower than -90 dBm. In Fig. 5.5, it is observed that the SNR of the analog TV signal is the least at 200 ft. from the TV tower. In spite of a strong received signal power at this distance, due to the presence of strong intermodulation products and noise, the SNR is low when measured close to the tower. However, at distances further away from the tower the SNR is improved. In the vicinity of the digital TV transmissions, the average noise level does not vary much with the distance from the TV tower and it is below -90 dBm at all the three distances. However, there are several spurious signal spikes in the average power spectrum when measured at 200 feet from the DTV station.

The BER results for an OFDM transceiver operating in an AWGN channel as well in the TV bands, are shown in Figs. 5.6 and 5.7[5]. The BER performance of the OFDM transceiver can be explained in terms of the varying levels of spurious signals and noise levels at the three distances from the TV tower.

Due to the presence of spurious signals, the performance in both the DTV and the analog TV bands is poor when compared with the AWGN channel case. In the vicinity of the analog TV band, the performance is worse at 200 feet from the tower. At distances of 600 feet and 5000 feet, the performance is comparable to the AWGN channel case at low SNR conditions of up to 3 dB. However, as the SNR increases there is no remarkable improvement in the performance as the

[5]These results have been published in [50] and have been reproduced here with the written permission of all the authors.

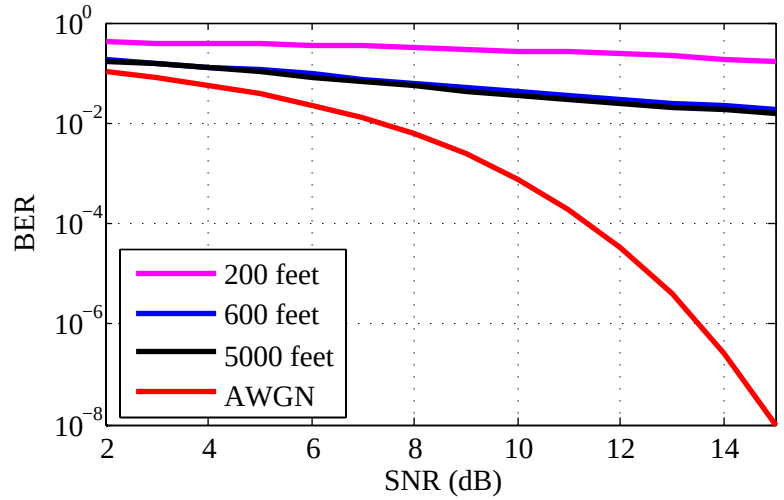

Figure 5.6. OFDM transceiver error performance in AWGN channel and vacant Analog TV bands.

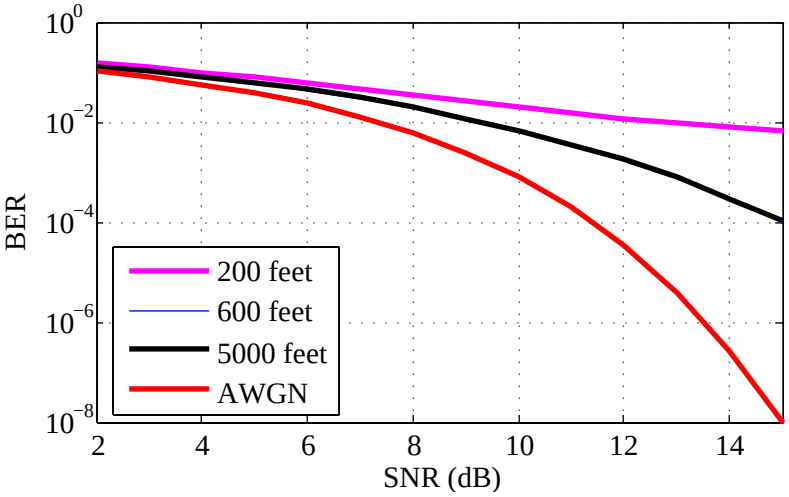

Figure 5.7. OFDM transceiver error performance in AWGN channel and vacant Digital TV bands.

BER never drops below 10^{-2}.

Better performance has been obtained in the case of the DTV band. The performance is poor at 200 feet distance with not much improvement at high SNRs. However, the performance at distances farther away from the tower improves with increasing SNR and it drops to 10^{-4} at an SNR of 14 dB. Moreover, at low SNR conditions, the performance at 600 feet and beyond is comparable to that in the

AWGN channel. It should also be noted that there is no change in the performance for distances beyond 600 feet from the tower, due to the presence of the spurious spikes which could be other licensed users.

The classifications of the DTV measurements performed using the Otsu's algorithm on time averaged DTV measurements were used along with an OFDM simulation that uses 2048 subcarriers. Among the classifications, the frequencies that have been identified as occupied by the DTV signal as well as the spurious signals were avoided while the frequencies that were found to be vacant were used for the transmission of the OFDM subcarriers. It was found that there is a small improvement in the BER performance as compared to the case where no knowledge of the presence of spurious signals was used.

5.2 PASS Framework for Wideband Spectrum Sensing in DSA networks

Various challenges are faced by DSA networks and real-time spectrum sensing such as varying spectrum behavior, co-existence with different primary users, and wireless channel fading. The radio should also be able to sense the spectrum over a wide bandwidth and detect the presence of signals with a low probability of interception [23]. Conventional approaches to wideband spectrum sensing possess several limitations. First, sensing the entire target spectrum continuously for white spaces may be inefficient in terms of time and power consumed by the sensing mechanism. Frequency channels with a high occupancy rate need not be sensed frequently, while channels with a high channel availability need to be monitored more frequently. Second, the channel occupancy state[6] can change rapidly

[6]The channel may be 'occupied' or 'available'.

such that the sensing mechanism may fail to keep track of the instantaneous states due to limitations on the sampling time resolution. Such a situation can result in an inaccurate characterization of the spectrum occupancy that can eventually lead to inefficient secondary usage of the spectrum, as well as allowing for transmissions that can interfere with the primary signals. Third, the characterization and learning period of the radio needs to be short in order to allow for rapid radio deployment. Finally, low cost cognitive radio might have a single RF front end. The RF circuit resources need to be shared efficiently between the sensing mechanism and the radio communication subsystem of the radio[7].

In this section, we address these issues with the proposed *Parametric Adaptive Spectrum Sensing* (PASS) architecture for dynamic spectrum sensing. In our approach, the radio learns the statistics of the channel availability by building an updated model of the channel occupancy, while the spectrum sensing mechanism adapts its parameters based on this model. Furthermore, we propose an algorithm for efficient scheduling of the sensing mechanism's time-frequency assignments, which enables efficient utilization of the radio's resources (time, power, and RF front end). While there has been preliminary work done on channel-aware spectrum sensing techniques [17], there has not been a clear solution to these wideband sensing issues. Nevertheless, the sensing in [17] is performed at fixed intervals, while our algorithm varies the sensing schedule dynamically according to the channel occupancy statistics.

[7]The radio communication subsystem handles the primary tasks of the radio namely the transmission and reception of signals

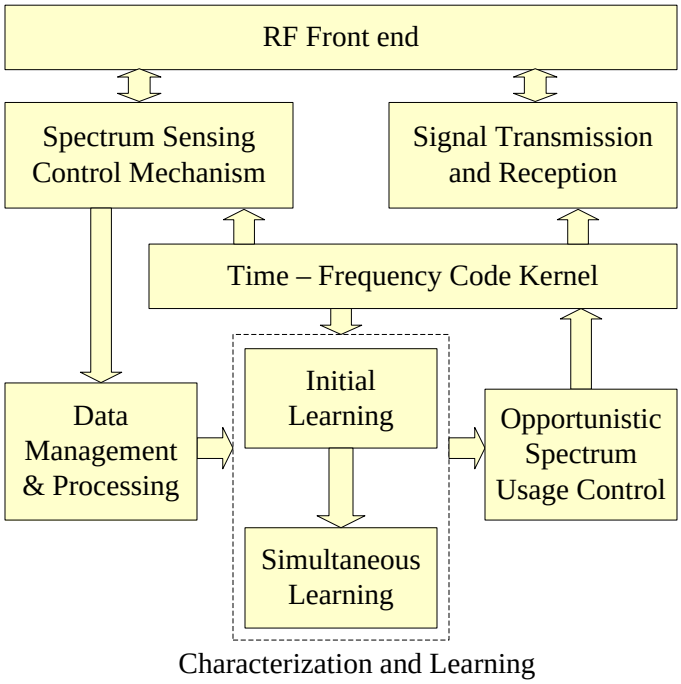

Figure 5.8. Block diagram of the proposed Parametric Adaptive Spectrum Sensing architecture.

5.2.1 Parametric Adaptive Spectrum Sensing Architecture

The proposed PASS architecture enables the radio to adapt its spectrum sensing mechanism to variations in the spectrum occupancy. The block diagram for the proposed PASS architecture is shown in Fig. 5.8.

The *RF front end* consists of the RF components of the radio, such as the antenna, local oscillator, mixer, and bandpass filters. During the sensing operation, the *RF front end* functions as a passive receiver that captures the spectrum measurements. The *spectrum sensing control mechanism* controls the parameters of the sensing operation. It translates the requirements of the sensing mechanism into the parameters of the RF components, such as the local oscillator's center frequency and the intermediate frequency filter bandwidth. The *RF front end* and the *spectrum sensing control block* constitute the cognitive radio's spectrum

sensing mechanism.

The *data management and processing* (DMP) block maintains a database of the spectrum measurements that are either collected by the radio itself or obtained from other radios. This database of measurements can be shared with other radios in the DSA network, thereby enabling collaborative spectrum sensing [18]. The data from the database is retrieved for processing to detect the presence of signals.

The *characterization and learning* block performs the analysis on the processed measurement data and characterizes the radio's operating spectrum by building a spectrum occupancy model, as well as update the parameters of the model. The radio's learning process occurs in two stages: *initial learning* and *simultaneous learning*. Initial learning refers to the stage of learning when the radio is deployed and it learns its wireless environment before becoming fully-functional for radio operations. In this stage of learning, the model parameters are estimated from the measurement data. When the estimates of the model parameters reach a certain confidence level, initial learning stops and the radio becomes fully-functional. At this point, the stage of simultaneous learning begins. In this stage of learning, the *RF front end* is shared by the sensing mechanism and the radio communication subsystem, and the learning process occurs simultaneously along with the primary operations of the radio[8]. As more data is collected and analyzed, the parameters of the model are updated, thereby improving the confidence levels of the parameter estimates. The statistics on channel availability are made available to the *opportunistic spectrum usage* (OSU) *control* block.

The OSU control block may be implemented in the radio resource management

[8]Sensing operations involve scanning for spectral white spaces and collecting data for characterizing spectrum occupancy, and radio communication operations include transmission and reception of information-bearing signals.

layer of a DSA network that regulates the access to the available spectrum. The *OSU control* block chooses a set of frequency channels among a pool of unoccupied channels based on the wireless channel characteristics such as signal-to-noise ratio, multipath fading, channel attenuation, and regulations imposed by a cooperative spectrum access system or a spectrum broker. Accordingly, it decides on which channels to sense and at what time intervals to sense them. Thus, the OSU block decides the time-frequency assignments for the sensing operations, and radio communication operations, as well as how the *RF front end* resources are shared among them. The *time-frequency code kernel* block stores these assignments. After every transmission, the *learning* block is notified by the kernel so that the model parameters are updated.

The spectrum sensing mechanism adapts its sweep parameters and scanning assignments according to the estimated model of the spectrum occupancy, hence the name *parametric adaptive spectrum sensing*. The spectrum sensing control mechanism, the DMP block, and the learning mechanism have all been derived from the spectrum survey framework.

5.2.2 Fine Tuning of Spectrum Sensing Parameters

In the initial learning stage, the spectrum sensing mechanism must be able to capture a signal in a frequency channel, track its instantaneous channel occupancy and characterize its occupancy. The characterization of the channel occupancy involves the estimation of various channel occupancy properties, such as the mean occupancy in the channel, and the 'on' time of the signal. This will aid in identifying potential opportunities for reuse. For instance, in the case of a channel occupied by a digital signal, the sensing mechanism can be switched off during

the 'on' times of the signal when the channel is occupied and a secondary signal can be time multiplexed during the 'off' time.

If the channel occupancy can be monitored continuously, then the channel state and channel properties can be specified with 100% reliability. However, sensing the wideband channel continuously is highly inefficient. In practical systems, the sampling of the continuous-time channel occupancy process can either be done in a periodic manner or with random sampling times [36]. While periodic sampling with short sample times is suitable for reliable estimation of 'on' time, the random sampling approach proposed in [36] overcomes the drawbacks of periodic sampling, such as biasing and correlation between samples. In this work, we choose periodic sampling for estimating the temporal properties of the channel occupancy such as the signal 'on' time. We present an iterative version of the periodic sampling method wherein the sample time recursively decreases until an accurate estimate of the signal 'on' time can be made. In our approach, for improved reliability, the parameters of the sensing mechanism are set adaptively in order to capture the particular type of signal that is being targeted. In the case of signals that occupy the spectrum continuously over time, such as FM radio and TV broadcast signals, the channel occupancy is nearly fixed and this makes it easier for the sensing mechanism to capture and track the channel occupancy with less stringent requirements on the time resolution [23]. However, in the case of digital and bursty signals, the temporal characteristics of the channel occupancy vary dynamically, making it difficult for spectrum monitoring and also making it complex for secondary spectrum access due to the need to avoid interference to the primary signals [16]. For instance, the time resolution may not be fine tuned to track a digital signal that has a short 'on' time. However, such signals can be

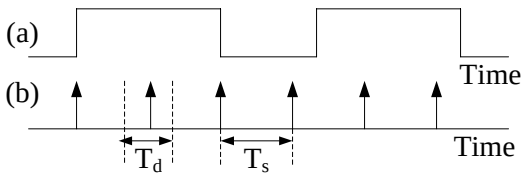

Figure 5.9. (a) Digital signal, (b) Pulse train representing the time sampling of digital signal by sensing mechanism

captured by suitable fine tuning of the sensing parameters in an iterative manner.

Consider a digital signal with 'on' time T_{on} occupying a certain frequency channel as shown in Fig. 5.9(a). Fig. 5.9(b) shows a pulse train with period T_s illustrating the time instances when that signal's channel occupancy is sampled. It is desired that the dwell time T_d is long enough to average out the random amplitude variations in the instantaneous power due to the effects of modulation [14], noise (such as the impulsive noise), multipath fading, and channel attenuation. The reduction in the amplitude variations will result in a uniform background noise which in turn helps in accurate estimation of the decision threshold for data classification. However, it is important that the dwell time does not extend over an entire cycle of 'on' and 'off' times, i.e., $T_d \leq T_{on}$.

The 'on' time of a digital signal can be estimated from the samples of the channel occupancy as follows:

$$\hat{T}_{\text{on}} = (N_{\text{on}} + 2) \cdot T_s, \tag{5.4}$$

where N_{on} is the number of samples collected during the 'on' time of the signal. This number can be determined by analyzing the spectrum measurement data collected over a channel. A factor of $2T_s$ has been added to the estimate to account for uncertainties in the channel occupancy due to the finite sampling time. It accounts for the channel occupancy during the time elapsed between the

last sample in the signal's 'off' time and the first sample in the succeeding 'on' time, and between the last sample in the signal's 'on' time and the first sample in the succeeding 'off' time. Due to this uncertainty, we safely assume that the channel is occupied during these time instances and hence $2T_s$ has been added to the 'on' time estimate. This factor introduces an error in the estimate which is expressed as follows:

$$\text{error} \;\leq\; (\,2\,T_s/T_{\text{on}})\;\times\;100\,\%, \tag{5.5}$$

$$\hat{T}_{\text{on}} \;=\; T_{\text{on}} + (\,\text{error}\;\times\;T_{\text{on}}\,/\,100\,). \tag{5.6}$$

Eq. (5.6) shows the relationship between the estimated 'on' time and the error. From Eq. (5.6), it is clear that by reducing T_s, the error is reduced and, in turn, the estimate of 'on' time is more accurate. T_s can be iteratively decreased such that the error reaches a tolerable limit and the signal is reliably captured.

5.2.3 Proposed Scheduling of Spectrum Sensing Assignments

An efficient scheduling of the time-frequency assignments of the spectrum sensing mechanism would improve the time resolution of the sweeps besides increasing the efficiency of the spectrum sensing. We present a novel algorithm that schedules the sweeps both along time and frequency, based on the channel occupancy statistics. The algorithm performs a backoff in the time resolution every time the channel is detected to be occupied. The backoff can be linear, or exponential. Fig. 5.10 illustrates the linear backoff in time resolution for the case when the channel is always occupied and for the case when it is vacant throughout the observation time. The objective is to minimize the number of scanning assignments and avoid wasteful sensing over presumably occupied segments of the spectrum. In

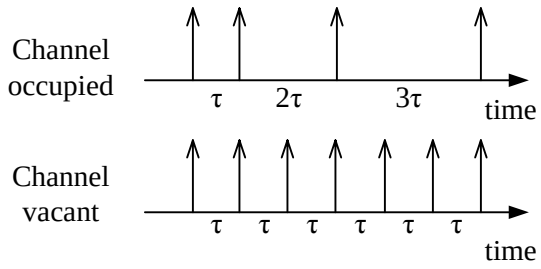

Figure 5.10. Backoff in sweep time resolution.

addition, by minimizing the scanning assignments the sweep time can be reduced. In this way, better time resolution can be provided to channels that are available with high probability. The details of the proposed algorithm are presented next.

Let N_c be the number of channels required by the radio for transmission. Depending on the time-bandwidth requirements of the radio and the availability of the same in the spectrum, N_c can be varied over time. For our discussion, we assume N_c is fixed, with no specific requirement on the time duration of the channel availability. N_c number of channels with the highest probability of availability P_0^i are selected and we can represent them as a set of channels $\{f_i\}, i = 1, 2...N_c$. These channels may not be contiguous in frequency. For the i^{th} channel, the sweep time resolution is expressed as $n_i\tau$ which is a multiple of the fundamental sample time τ and n_i is a variable scaling factor. The time-frequency assignments for the spectrum sensing mechanism are stored in matrix $\mathbf{A} = [A(f_i, t_j)]$ where $A(f_i, t_j)$ specifies if the radio has to scan channel f_i at time instance t_j. Fig. 5.11 shows the flowchart for the proposed algorithm for a linear backoff case.

We assume a total time of T for our simulation. Initially, the sensing mechanism is assigned to scan all the frequency channels $\{f_i\}$ at all time instants $t_j = 1, 2, \ldots, T$ and all the values in $\mathbf{A}$ are set to 'scan'. During each time instance t_j, before scanning the frequency channels, the sensing mechanism first

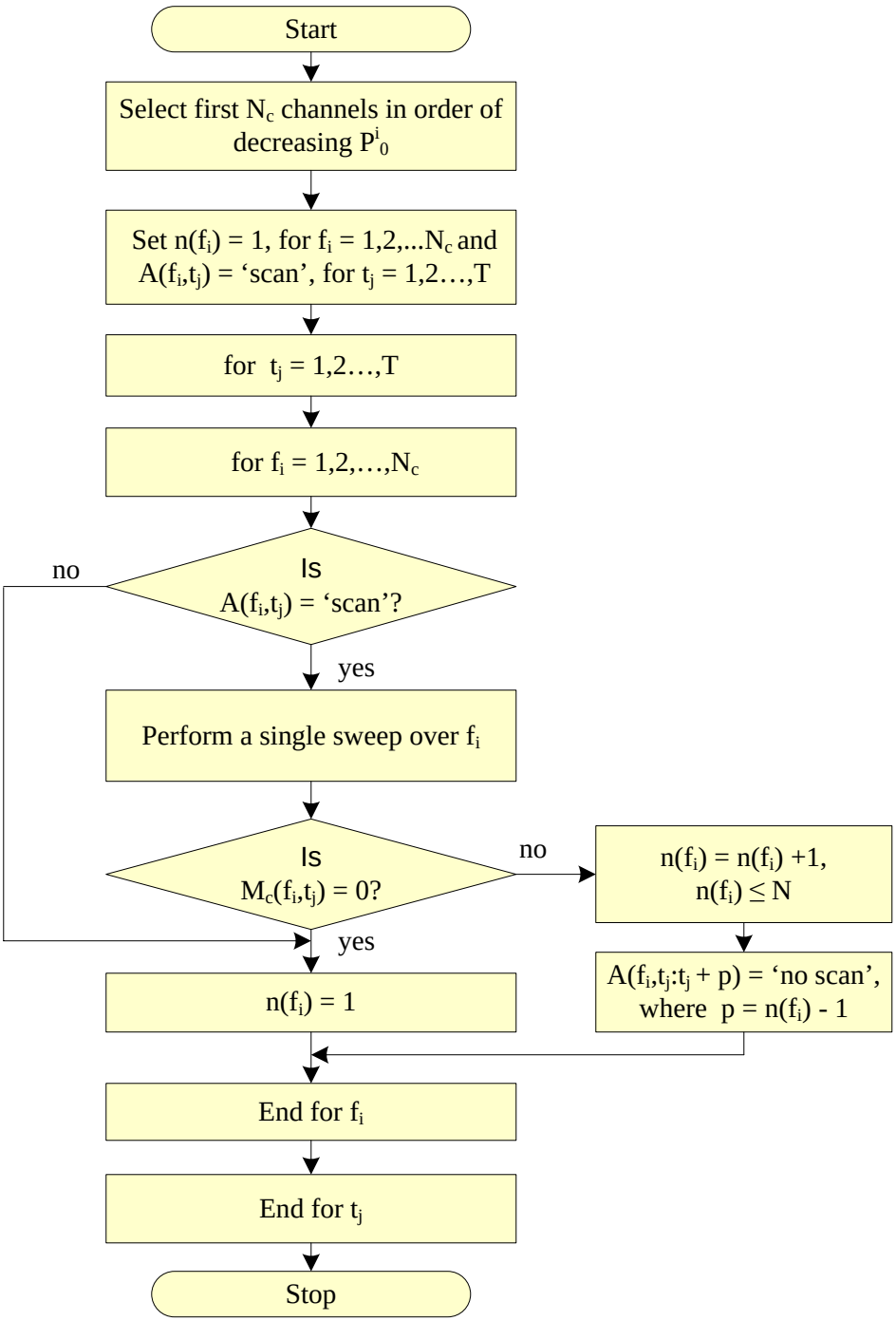

Figure 5.11. Flowchart of proposed scheduling algorithm.

reads from the corresponding row in **A** and scans only those channels that have been scheduled for sensing. After sensing, based on the perceived channel occupancy state $M_c(f_i, t_j)$, n_i is either incremented or set to the default value of one. Incrementing n_i results in backoff in the time resolution for sensing channel f_i. Accordingly the sensing schedule for the frequency channels at future time instances is updated in **A**. The maximum backoff N can be specified.

5.2.4 Simulation Setup

In our computer simulation, we randomly generated the spectrum occupancy in frequency channels with a bandwidth of 33 MHz and a 10 kHz spacing between adjacent spectral samples. In the simulated spectrum, some of the channels are assumed licensed and the primary users remain 'on' throughout the observation time. Each primary user occupies 6 MHz of the spectrum which also includes the guard bands. The guard bands between adjacent licensed bands are occupied by secondary users that transmit in bursts. The bursty signals have exponentially distributed 'on' times, where the mean 'on' time is randomly selected from a uniform distribution with the interval (0.0, 1.0). We generated 1000 instances of the spectrum occupancy. Each channel has different mean 'on' times and different channel occupancy statistics and this allows us to exploit the adaptive nature of the proposed algorithm. It is further assumed that the channel occupancy state does not vary over the sampling intervals. Fig. 5.12 shows the intensity plot for the simulated spectrum occupancy. In the figure, the darker areas represent the occupied spectrum while the unoccupied spectrum is represented by the bright regions.

The proposed scheduling algorithm using a linear backoff was run over this

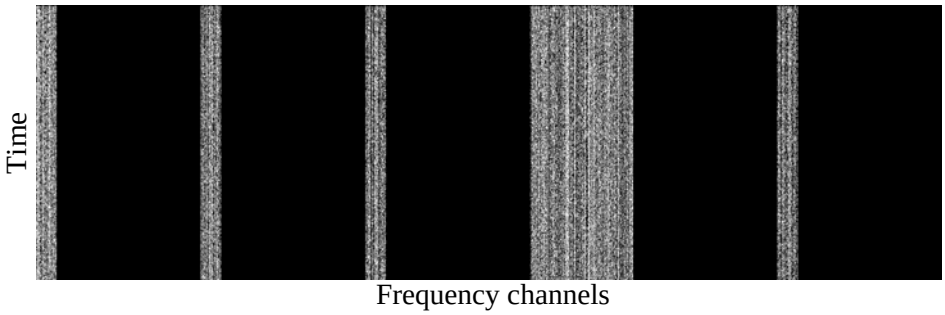

Figure 5.12. Intensity plot of simulated spectrum occupancy. The darker areas represent the occupied spectrum.

randomly generated spectrum. To test the algorithm in the worst case scenario, it was assumed that $N_c = 33$ MHz, which means that all the channels are considered for scanning initially including the licensed channels with 100 % occupancy.

5.2.5 Simulation Results

Fig. 5.13 shows the proportion of the total number of channels that were sensed during each time instance and Fig. 5.14 shows the proportion of the total time for which each channel was sensed. If the conventional fixed scheduling approach was followed, ideally all the channels should have been sensed 100% of the time. However, by applying the proposed algorithm, it is seen from the results that there is a drastic decrease in the sweep time at each time instance (see Fig. 5.13) and over the occupied bands the sensing mechanism scans for only a very small percentage of the total observation time, as seen from Fig. 5.14. The portions of the spectrum occupied by the primary signals are scanned with a low time resolution with large spacing between consecutive time samples. The sweep time resolution is high for the frequency channels that are sparsely occupied and the resolution varies according to the spectrum occupancy in each of those channels. The simulation was repeated 100 times and it was determined that, on an average, there is a

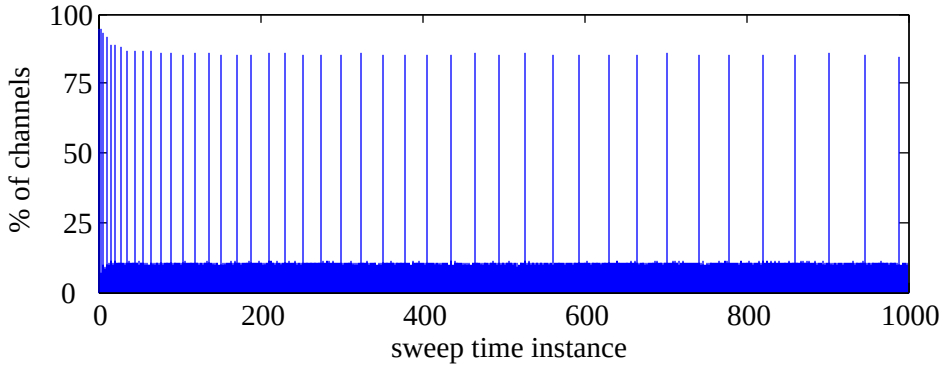

Figure 5.13. Simulation results for linear backoff when applied to simulated spectrum shared with secondary signals: Number of channels sensed at each time instance.

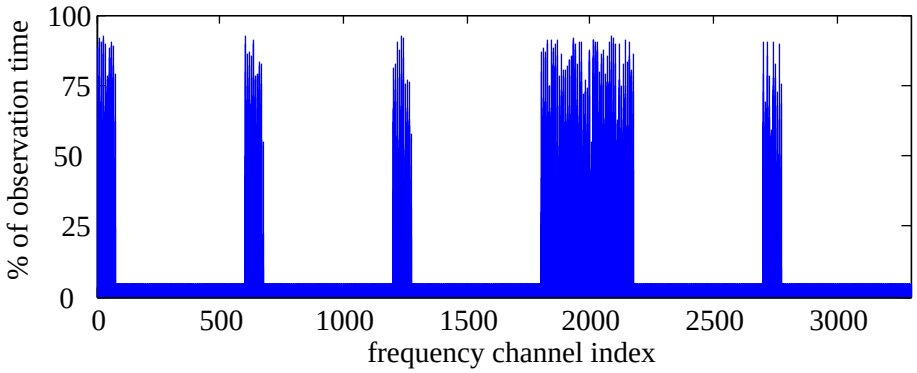

Figure 5.14. Simulation results for linear backoff: Total sensing time for each channel

84.23% improvement in the efficiency over the conventional approach. However, on an average, 1.84 % of the potential channel vacancies were not scanned due to the backoff, which is insignificant when compared to the drastic improvement in the sensing efficiency. This improvement in efficiency reflects as savings in time and power consumption, assuming that the sensing mechanism is switched-off when it is not assigned to scan.

5.3 Chapter Summary

In this chapter, we have presented a feasibility study of secondary usage of the TV spectrum. The spectrum measurements collected were used in an OFDM simulation to determine the error performance of OFDM symbols transmitted through the vacant TV spectrum. The PASS framework for efficient spectrum sensing has also been presented. A novel algorithm for scheduling the assignments for spectrum sensing both over frequency and time has been presented along with simulation results.

Chapter 6

Conclusion

This chapter provides the concluding remarks for the thesis along with a summary of the novel contributions of this thesis. Section 6.1 provides the future works.

Spectrum Survey Framework - The proposed framework introduces standardization to spectrum surveying which can enable collaborations on studying the spectrum behavior. The proposed spectrum survey framework has been implemented as shown in Fig 3.6. Currently, it is being used to collect measurements along time and frequency, and process the data offline (non-real time). Till date, several sets of measurements have been collected from different bands ranging from 9 kHz - 1 GHz including measurements from the FM broadcast spectrum (88-108 MHz), television spectrum (54-87 MHz, 638-668 MHz, 198-228 MHz), and paging band (929-931 MHz).

Processing and analysis of spectrum measurements - Techniques have been presented to counter various challenges to the processing of spectrum mea-

surements. The processing techniques have been tested on the collected spectrum data. The FM band measurement data which is affected by spurious sidelobes and high dynamic range of signals represented the worst case data. Although the results obtained in the case of FM band data were not excellent, these results provided us an insight into how various parameters of the processing techniques such as sliding window size affect the results and also allowed the comparison between the proposed classification algorithms. Good results were obtained when the techniques were applied on the TV spectrum data, cellular band, and the paging band data.

Otsu's algorithm and the ROHT algorithm performed well with most of the data sets except the FM band data. Time averaging the measurements improved their performance in all the bands except the paging band. However, when applied on the cellular band data, the ROHT algorithm could perform well on the data without the need for time averaging while the Otsu's algorithm performed well only on the time averaged data.

The ROHT algorithm and the modified recursive Otsu's algorithm have been proposed for classification of data which has signals with a high dynamic range. The recursive Otsu's algorithm gave a lot of false alarms. The ROHT algorithm was found to possess the following drawbacks: (i) The parameters of the ROHT algorithm had to be set depending on the dynamic range of the signals (or the standard deviation of the signal portion of the data histogram). While ROHT algorithm with 96% confidence level and $\epsilon = 0.5$ performed well with the television bands, cellular band, and FM band data, the parameters had to be increased to 99.5% confidence level and $\epsilon = 1.5$ when the ROHT algorithm was being applied to the paging band. (ii) It requires a large number of samples for estimating

the threshold. This was observed when the performance on the FM band data improved by increasing the window size while using a sliding window approach.

Adaptive thresholding using the sliding window approach improved the classification results in bands such as the analog television band (198-228 MHz) where the noise level varies across the band. However, the drawback of this approach is that the window size depends on the occupancy of the band. For instance, when used with the cellular band which is sparsely occupied, a window size of 1000 gave much lesser false alarms as compared to a window size of 500.

The energy detection techniques that we have presented rely entirely on the ability to distinguish between the power levels of signals and noise. The main drawback of the proposed techniques is that they cannot distinguish between signal, noise, and interference. Strong signals that appear against a relatively uniform background noise are easily detected while signals that are close to the noise level may not be easily distinguished from the noise.

A set of parameters of spectrum utilization were identified and tools have been developed in Matlab in order to extract these features from the processed data.

Modeling of spectrum measurements and spectrum occupancy - A comprehensive model for the spectrum measurements has been presented. This model can, in turn, incorporate the models for channel occupancy, noise, and signals in the target channel. A modified markov model that accounts for the time varying nature of the channel occupancy was also presented. The issue of efficient characterization of the channel occupancy was addressed by presenting an expression to determine the minimum time required to estimate the channel model parameters.

Study of feasibility of secondary usage in television spectrum - We have presented a feasibility study of secondary transmissions into the TV spectrum. It has been observed that usage in the vicinity of a TV transmitter can result in poor secondary user performance results, but more distant operation is relatively free of intermodulation or spurious signal effects, yielding better conditions for DSA. Our hope is that this study, and future results from the continuing work at the University of Kansas in this subject area, will be of value in regulatory discussions concerning spectrum policy decisions that will ultimately define access to a valuable national asset, the TV band spectrum.

Application of SSF to spectrum sensing - In this thesis, we have introduced the PASS framework for efficient and fast DSA by spectrum sensing, channel occupancy modeling and adapting the spectrum sensing parameters based on the estimated channel occupancy model. By following a channel-aware adaptive approach, the PASS framework is able to overcome the limitations of the conventional non-adaptive sensing mechanisms. We analyzed fine tuning of the channel sampling times and we have shown that there is a decrease in the error in estimating a signal's temporal characteristics with improvement in the time resolution of the sampling process. We have also presented a novel algorithm for scheduling the assignments for spectrum sensing both over frequency and time. The simulation results showed a drastic improvement in the sensing efficiency as compared to a conventional non-adaptive approach.

6.1 Future Research Directions

Spectrum surveying along spatial dimension of spectrum - SSF has been implemented to conduct studies along frequency and time dimensions of the spectrum. However, by using a network of measurement nodes which can capture spatial spectrum data, the full potential of the SSF's collaborative spectrum surveying can be tested. The processing techniques that have been presented in this thesis can be extended for processing multi-dimensional data. For instance, using the data collected simultaneously at different azimuth angles, analysis algorithms can be developed in order to resolve the spectrum utilization spatially.

Implementation of SSF in KUAR - A novel spectrum sensing architecture that uses the SSF has been proposed and tested in simulation. The SSF as well as the sensing architecture can be implemented in the KUAR[1] radio in order to improve its spectrum sensing capabilities. The processing and analysis techniques can be modified in order to enable real-time spectrum surveying.

Incorporate spectrum measurements into NRNRT - The spectrum measurements that have been collected using the SSF can be incorporated into the National Radio Networks Research Testbed. The data collected can be utilized in an emulation of the radio spectrum which in turn can be used to test innovative radio technologies for spectrum sensing, primary signal detection, and interference avoidance.

Advanced processing and signal detection algorithms - A major limitation of energy detection techniques is that they fail to perform in low SNR

[1]The KUAR refers to the software radio platform that is being developed for agile radio operation at the University of Kansas.

conditions or in the presence of interference. Advanced techniques such as cyclostationary signal detection [23], which require *a priori* knowledge about the primary signal characteristics can be developed in order to distinguish between noise, interference, and intermodulation products.

Appendix A

Derivation of Optimum Threshold in Otsu's Algorithm

This derivation has been adapted from [39]. Otsu's algorithm [39] selects an optimum threshold based on the properties of the histogram of the data and it does not assume any model for the histogram. The optimum threshold results in maximum separation between the two classes in the data, namely the signal and the noise classes. The algorithm also returns a metric that indicates the separability of the 2 classes, which is useful to quantify the 'goodness of the threshold'. Before applying Otsu's algorithm, the measurement data in **M** is converted to the gray scale image **I**.

The data is quantized into L levels with values $s \times [1, 2, ..L]$, where s is a scaling factor. Let the i^{th} gray level value be denoted by g_i and its histogram is denoted by h_i. The probability of occurence of the i^{th} level in the data can be estimated as:

$$p_i = \frac{h_i}{N}, \; where \, N = \sum_{i=1}^{L} h_i$$

The mean of the distribution is defined as:

$$\mu_T \;=\; \sum_{i=1}^{L} g_i \;\cdot\; p_i$$

A threshold, $T \;=\; g_k$, can be used to bifurcate the probability distribution into the noise class C_0 and the signal class C_1, with the levels $[1, 2, .. k] \in C_0$ and levels $[k+1, .. L] \in C_1$. The probability of occurence of the classes can be defined as a function of the threshold:

$$\omega_0 \;=\; \sum_{i=1}^{k} p_i \;=\; \omega_k$$

$$\omega_1 \;=\; \sum_{i=k+1}^{L} p_i \;=\; 1 - \omega_k$$

The probability distribution of the gray levels within each of the classes are:

$$P(g_i \,|\, g_i \in C_0) \;=\; \frac{h_i}{\sum_{i=1}^{k} h_i} \;=\; \frac{p_i}{\omega_0}\,,\; i \;=\; 1 \;\ldots\; k$$

$$P(g_i \,|\, g_i \in C_1) \;=\; \frac{h_i}{\sum_{i=k+1}^{L} h_i} \;=\; \frac{p_i}{\omega_1}\,,\; i \;=\; k+1 \;\ldots\; L$$

The corresponding means of the probability distributions are defined as:

$$\mu_0 \;=\; \sum_{i=1}^{k} g_i \; P(g_i \,|\, g_i \in C_0) \;=\; \sum_{i=1}^{k} g_i \; \frac{p_i}{\omega_0} \;=\; \frac{\mu\,(k)}{\omega_k}$$

$$\mu_1 \;=\; \sum_{i=k+1}^{L} g_i \; P(g_i \,|\, g_i \in C_1) \;=\; \sum_{i=k+1}^{L} g_i \; \frac{p_i}{\omega_1} \;=\; \frac{\mu_T - \mu\,(k)}{1 - \omega_k}$$

In the above expressions, we have denoted $\sum_{i=1}^{k} g_i \; p_i \;=\; \mu\,(k)$. The between-class

variance is defined as [39]:

$$\sigma_B^2 \;=\; \omega_0\,(\mu_0 - \mu_T)^2 + \;\omega_1\,(\mu_1 - \mu_T)^2 \tag{A.1}$$

$$=\; \omega_0\,\omega_1\,(\mu_1 - \mu_0)^2 \tag{A.2}$$

In terms of the threshold at the k^{th} gray level and using equations , the above expression for between - class variance can be rewritten as:

$$\sigma_B^2\,(k) \;=\; \frac{[\,\mu_T\,\omega_k - \mu(k)\,]^2}{\omega_k\,[1-\omega_k]} \tag{A.3}$$

A measure of class separability has been defined in [39] as shown by equation A.4.

$$\eta \;=\; \sigma_B^2/\sigma_T^2 \tag{A.4}$$

Otsu's algorithm is an optimization problem that involves determining the optimum threshold's gray level k^*, that maximizes the above defined measure of class separability i.e. mathematically,

$$\eta(k^*) \;=\; max_{1\le k<L}\,\sigma_B^2(k)/\sigma_T^2 \tag{A.5}$$

A simple way of determining the optimum threshold would be by varying the threshold in steps computing the measure of separability and then choosing the threshold that gives the maximum value for this measure.

Bibliography

[1] G. Minden and J. Roberts, "National Radio Networking Research Testbed Project Proposal Submitted to National Science Foundation."

[2] R. J. Matheson, "The Electrospace Model as a Frequency Management Tool," in *International Symposium on Advanced Radio Technologies*, (Boulder, CO), pp. 126–132, March 2003.

[3] A. Petrin and P. G. Steffes, "Analysis and Comparison of Spectrum Measurements Performed in Urban and Rural Areas to Determine the Total Amount of Spectrum Usage," in *International Symposium on Advanced Radio Technologies*, (Boulder, CO, USA), pp. 9–12, March 2005.

[4] A. Petrin and P. G. Steffes, "Potential Usability of Allocated but Unused Spectrum in the United States of America," in *27th Triennial General Assembly of the International Union of Radio Science*, (Maastricht, The Netherlands), August 2002.

[5] "US Frequency Allocation Chart." NTIA, October 2003.

[6] "The XG Vision: Request For Comments," tech. rep., BBN Technologies, version 2.0.

[7] S. Ellingson, "Spectral Occupancy at VHF: Implications for Frequency-Agile Cognitive Radios," in *IEEE Vehicular Technology Conference*, vol. 2, (Dallas), pp. 1379–82, September 2005.

[8] Federal Communications Commission, "FCC Chairman Michael K. Powell Outlines Critical Elements Of Future Spectrum Policy." Spectrum Speeches and Presentations - http://www.fcc.gov/sptf/speeches.html, August 2002.

[9] "Report of the Spectrum Efficiency Working Group," tech. rep., Federal Communications Commission Spectrum Policy Task Force, 15 November 2002.

[10] A. J. Petrin, *Maximizing the Utility of Radio Spectrum: Broadband Spectrum Measurements and Occupancy Model for Use by Cognitive Radio*. PhD

thesis, School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia, August 2005.

[11] M. Wylie-Green, "Dynamic Spectrum Sensing by Multiband OFDM Radio for Interference Mitigation," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 619 – 625, November 2005.

[12] Federal Communications Commission, "Unlicensed Operation in the TV Broadcast Bands." ET Docket No. 04-113, May 2004.

[13] C. Cordeiro, K. Challapali, D. Birru, and S. Shankar, "IEEE 802.22: The First Worldwide Wireless Standard Based on Cognitive Radios," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 328–337, November 2005.

[14] J. R. Hoffman and R. J. Matheson, "RSMS Measurement and Analysis of LMR Channel Usage," in *International Symposium on Advanced Radio Technologies*, (Boulder, CO, USA), pp. 13–19, March 2005.

[15] "The XG Architectural Framework: Request For Comments," tech. rep., BBN Technologies, version 1.0.

[16] H. Tang, "Some Physical Layer Issues of Wide-Band Cognitive Radio Systems," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 151 – 159, November 2005.

[17] S. D. Jones, N. Merheb, and I.-J. Wang, "An Experiment for Sensing-Based Opportunistic Spectrum Access in CSMA/CA Networks," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 593– 596, November 2005.

[18] A. Ghasemi and E. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Envionments," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 131–136, November 2005.

[19] A. Petrin and P. G. Steffes, "Study of Spectrum Usage and Potential Interference to Passive Remote Sensing Activities in the 4.5 cm and 21 cm Bands," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 3, (Anchorage, Alaska, USA), pp. 1679– 1682, September 2004.

[20] M. A. McHenry, "NSF Spectrum Occupancy Measurements Project Summary," tech. rep., Shared Spectrum Company, August 2005.

[21] A. Petrin and P. G. Steffes, "Measurement and Analysis of Urban Spectrum Usage," in *Proceedings of International Symposium on Advanced Radio Technologies*, (Boulder, CO, USA), pp. 45–48, NTIA Special Publication SP-04-409, March 2004.

[22] A. Rogers, J. Salah, D. Smythe, P. Pratap, J. Carter, and M. Derome, "Interference Temperature Measurements from 70 to 1500 MHz in Suburban and Rural Environments of the Northeast," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 119–123, November 2005.

[23] D. Cabric, S. M. Mishra, and R. W. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*, vol. 1, (Pacific Grove, California, USA), pp. 772– 776, November 2004.

[24] A. E. Leu, K. Steadman, M. McHenry, and J. Bates, "Ultra Sensitive TV detector Measurements," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD), pp. 30–36, November 2005.

[25] R. Tandra and A. Sahai, "Fundamental Limits on Detection in Low SNR Under Noise Uncertainty," in *International Conference on Wireless Networks, Communications and Mobile Computing*, vol. 1, pp. 464– 469, 2005.

[26] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Experimental Study of Spectrum Sensing based on Energy Detection and Network Cooperation," in *Proceedings of the 2006 International Workshop on Technology and Policy for Accessing Spectrum (TAPAS)*, (Boston, MA, USA), August 2006.

[27] D. Cabric, A. Tkachenko, and R. W. Brodersen, "Spectrum Sensing Measurements of Pilot, Energy, and Collaborative Detection," in *IEEE Military Communications Conference*, (Washington DC, USA), October 2006.

[28] Mark McHenry (Shared Spectrum Company), "Spectrum Measurements to Support Dynamic Spectrum Sharing Research and Development." Presentation to NSF Future Spectrum Technology and Policy Workshop, May 25-27 2005.

[29] G. F. Gott, S. K. Chan, C. A. Pantjiaros, and P. J. Laycock, "High Frequency Spectral Occupancy at the Solstices," in *IEE Proceedings - Communications*, vol. 144, pp. 24–32, February 1997.

[30] C. A. Pantjiaros, P. J. Laycock, G. F. Gott, and S. K. Chan, "Development of the Laycock-Gott Occupancy Model," in *IEE Proceedings - Communications*, vol. 144, pp. 33–39, February 1997.

[31] D. J. Percival, M. Kraetzl, and M. S. Britton, "A Markov Model for HF Spectral Occupancy in Central Australia," in *7th International Conference on High Frequency Radio Systems and Techniques*, (Nottingham, UK), pp. 14–18, July 1997.

[32] M.-H. Chang and K.-H. Lin, "A Comparative Investigation on Urban Radio Noise at Several Specific Measured Areas and its Applications for Communications," *IEEE Transactions on Broadcasting*, vol. 50, pp. 233 – 243, September 2004.

[33] F. Weidling, D. Datla, V. Petty, P. Krishnan, and G. J. Minden, "A Framework for RF Spectrum Measurements and Analysis," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 573–576, November 2005.

[34] A. D. Spaulding and G. H. Hagn, "On the Definition and Estimation of Spectrum Occupancy," *IEEE Transactions on Electromagnetic Compatibility*, vol. 19, pp. 269–280, August 1977.

[35] J. S. Bendat and A. G.Piersol, *Random Data: Analysis and Measurement Procedures*. John Wiley & sons, 1971.

[36] S. R. Fleurke, H. G. Dehling, H. K. Leonhard, A. D. Brinkerink, and R. den Besten, "Measurement and Statistical Analysis of Spectrum Occupancy," *European Transactions on Telecommunications*, vol. 15, pp. 429–436, 2004.

[37] R. H. Olson, "On the Use of Baye's Theorem in Estimating False Alarm Rates," *American Meteorological Society, Monthly Weather Review*, vol. 93, pp. 557–558, September 1965.

[38] W. K. Pratt, *Digital Image Processing*. John Wiley & Sons, 3 ed.

[39] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC - 9, pp. 62–66, January 1979.

[40] M. Sonka, V. Hlavac, and R. Boyl, *Image Processing , Analysis and Machine Vision*. Brooks/Cole Publishing Company, second ed.

[41] M. Cheriet, J. N. Said, and C. Y. Suen, "A Recursive Thresholding Technique for Image Segmentation," *IEEE Transactions on Image Processing*, vol. 7, pp. 918–921, June 1998.

[42] E. Visotsky, S. Kuffner, and R. Peterson, "On Collaborative Detection of TV Transmissions in Support of Dynamic Spectrum Sharing," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 338–343, November 2005.

[43] B. Wild and K. Ramchandran, "Detecting Primary Receivers for Cognitive Radio Applications," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 124–130, November 2005.

[44] M. McHenry, "The Probe Spectrum Access Method," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD,USA), pp. 346–351, November 2005.

[45] American Federation of Musicians, "Comments in response to ET Docket No. 04-186," October 2006.

[46] Wireless Internet Service Provider's Association, "Comments in response to ET docket no. 04-186," March 2006.

[47] C. W. Rhodes, "Interference Between Television Signals due to Intermodulation in Receiver Front - Ends," *IEEE Transactions on Broadcasting*, vol. 51, pp. 31–37, March 2005.

[48] "H.R.2015 - Balanced Budget Act of 1997." http://thomas.loc.gov/cgi-bin/query/z?c105:H.R.2015.ENR:.

[49] T. X. Brown, "An Analysis of Unlicensed Device Operation in Licensed Broadcast Service Bands," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 11–29, November 2005.

[50] R. Rajbanshi, V. R. Petty, D. Datla, F. Weidling, D. DePardo, P. Kolodzy, M. J. Marcus, A. M. Wyglinski, J. B. Evans, G. J. Minden, and J. A. Roberts, "Feasibility of Dynamic Spectrum Access in Underutilized Television Bands," in *Submitted to Second IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, 2007.

[51] D. Prendergast, M. Guillet, B. Caron, Y. Wu, X. Wang, B. Ledoux, and S. Lafleche, "The Effects of Public Safety Mobile Systems Operations (in TV Channels 63/68) on DTV and NTSC Broadcasting," *IEEE Transactions on Broadcasting*, vol. 51, pp. 43– 50, March 2005.

[52] J. D. Poston and W. D. Horne, "Discontiguous OFDM Considerations for Dynamic Spectrum Access in Idle TV Channels," in *IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, (Baltimore, MD, USA), pp. 619– 625, November 2005.

[53] J. G.Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 4 ed., 2001.