COOPERATIVE SPECTRUM SENSING AND RADIO ENVIRONMENT MAP CONSTRUCTION IN COGNITIVE RADIO NETWORKS

by

Hüseyin Birkan Yılmaz

B.S., Mathematics, Bogazici University, 2002

M.S., Computer Engineering, Bogazici University, 2006

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Computer Engineering

Boğaziçi University

2012

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to all those who gave me the possibility to complete this thesis. To those individuals I neglect to mention here by name, I still offer my deepest thanks.

First and foremost, I would like to thank my thesis supervisor Assoc. Prof. Tuna Tuğcu for many insightful conversations during the development of the ideas in this thesis, and for helpful comments on the text. Without his support, it would be impossible for me to finish this thesis.

I am grateful to Prof. Sema Oktuğ, Prof. Özgür Barış Akan, Assoc. Prof. Fatih Alagöz, and Prof. Cem Ersoy for their participation in my thesis committee and their insightful comments for making this thesis better.

I would like to thank M. Kıvanç Mıhçak for his support via discussions during the early part of the thesis study. I am gratified to all my friends from Boğaziçi University especially Suzan Bayhan for her ideas, support during investigation of the mathematical background, and suggestions for the artwork of this thesis.

Last, but not least, I would like to thank my family and spouse, Nebahat. They gave me their love and support throughout all stages of my education.

This thesis is partially supported by the Scientific and Technical Research Council of Turkey (TUBITAK) under BIDEB 2211 (National Ph.D. Scholarship) and grant number 108E101, COST Action IC0902 - Cognitive Radio and Networking for Cooperative Coexistence of Heterogeneous Wireless Networks, and Turkish State Planning Organization (DPT) under grant number 07K120610.

ABSTRACT

COOPERATIVE SPECTRUM SENSING AND RADIO ENVIRONMENT MAP CONSTRUCTION IN COGNITIVE RADIO NETWORKS

In this thesis, we focus on both internal and external sensing in Cognitive Radio (CR) networks. In internal sensing, individual CRs discover spectrum opportunities via spectrum sensing whereas in external sensing, an external entity provides the spectrum occupancy and related information. For the first, we propose a novel cooperative spectrum sensing scheme, Uniform Quantization-based Cooperative Sensing (UniQCS) that uses uniform quantization and an effective fusion strategy. Numerical results demonstrate that under imperfect reporting channel and false reports, UniQCS performs better than hard decision algorithms such as Majority and M-of-N in terms of probability of detection and false alarm at the expense of a marginal increase in overhead bits. We demonstrate that the performance of UniQCS is very close to that of equal gain combiner, which constitutes the upper bound for the decision performance. Due to the challenges in internal sensing, external sensing recently has gained noticeable interest. In external sensing, CRs access spectrum through geolocation databases, which keep relatively static information. Radio Environment Map (REM) is a kind of improved geolocation database and an emerging topic with the latest regulations on TV white space communications. It constructs a signal power temperature map of the CR operation area via processing spectrum measurements collected from sensors dynamically. In this thesis, transmitter LocatIon Estimation based (LIvE) REM construction technique is proposed and compared with the well-known REM construction techniques in shadow and multipath fading channels. The simulation results suggest that the LIvE REM construction outperforms the compared methods in terms of root mean square error and correct detection zone ratio.

ÖZET

BİLİŞSEL RADYO AĞLARINDA İŞBİRLİKÇİ SPEKTRUM ALGILAMA ve FREKANS HARİTASI OLUŞTURMA

Bu tezde, Bilişsel Radyo (BR) ağlarında dahili ve harici algılamaya odaklanılmaktadır. Dahili algılamada BR'ler spektrum algılama ile fırsatları farkederler, halbuki harici algılamada spektrum kullanım bilgisi harici bir birim tarafından sunulur. İlki için, düzgün nicemleyici ve etkili bir kaynaştırma stratejisi kullanan yeni bir yardımlaşmalı spektrum algılama öneriyoruz (UniQCS). Önerilen yöntem hatalı raporlama kanalı ve yanlış raporların varlığında, pek az ek bit bedeliyle, çoğunluk ve N-den M gibi sıfır-bir kararı veren algoritmalardan algılama ve yanlış alarm olasılıkları açısından daha iyidir. Ayrıca önerdiğimiz yöntem Eşit Kazançlı Birleştirici (EKB) ile karşılaştırıldı ve algoritmamız EKB'ye yakın bir başarım sergilemektedir (EKB kullanıcılar arasında ayrım olmadığı durumlarda üst sınırdır). Dahili algılamadaki zorluklar sebebiyle harici algılama farkedilir bir ilgi kazanmıştır. Harici algılamada bilişsel radyo spektrum erişimini durağan bilgi tutan konum belirleme veritabanı üzerinden gerçekleştirmektedir. Radyo Ortam Haritası (ROH) daha gelişmiş bir konum belirleme veritabanıdır ve TV boş kanallarında haberleşme konusundaki en son düzenlemelerle güncel bir konu olmuştur. ROH, algılayıcılardan toplanan spektrum ölçümlerini işleyerek dinamik bir sinyal gücü sıcaklık haritası oluşturur. Bu tezde, aktif gönderici konumunun kestiriminden yararlanan, LIvE ROH oluşturma tekniği önerilmiştir ve bilinen tekniklerle kırınımlı ve sönümlü kanallarda karşılaştırılmıştır. Benzetim sonuçları önerilen yöntemin kanal bilgisini kullanarak diğer yöntemleri karesel ortalama hata ve doğru algılama bölge oranı açısından aştığını göstermektedir.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

A^T	Transpose of matrix A
$A(\mathrm{Z})$	Area of zone Z
$A(\mathrm{Z}_1^{CD})$	Area of correct detection zone type-1
$A(\mathrm{Z}_0^{FA})$	Area of false alarm zone
$A(\mathrm{Z}_0^{CD})$	Area of correct detection zone type-0
$A(\mathrm{Z}_1^{MD})$	Area of missed detection zone
B_k	kth bin of quantizer
B_c	Center bin of quantizer
$\overrightarrow{B}$	Bin vector at fusion center representing the number of nodes having observation in bins
$\overrightarrow{B}^{+fr}$	Bin vector with last component incremented by fr
$\mathcal{B}$	The set of all possible $\overrightarrow{B}$ vectors
$\mathcal{B}_0$	The set of $\overrightarrow{B}$ vectors where the final decision is H_0
$\mathcal{B}_1$	The set of $\overrightarrow{B}$ vectors where the final decision is H_1
$\mathcal{B}^{-fr}$	The set of all combinations of N-fr users distributed in quantization bins
$c(x, y)$	Covariance function
CDZR_1	Correct detection zone ratio type-1
$d(.,.)$	Distance between two points or vectors
$\widehat{d}_{(x,y)}$	Distance between point (x, y) and estimated transmitter location
d_{min}	The distance between primary transmitter and the closest MCD
$E_i(\mu)$	Exponential random variable with mean μ
f	Real valued (cost, measurement, random valued) function
$\hat{f}(x)$	Predicted value of f at point x
f_c	Center frequency
$f(O_i\|\gamma)$	Probability distribution function of O_i for a given SNR
$F(x)$	Random process generating $f(x)$

${}_1F_1(.;.;.)$	The confluent hypergeometric function
FAZR	False alarm zone ratio
fr	Number of false reports
h	The complex gain of the channel
H_0	Null hypothesis of the spectrum sensing problem (white space)
H_1	Alternative hypothesis of the spectrum sensing problem (channel occupied)
L	Number of i.i.d. diversity branches
L_i	Local measurement bin index of node i
$\mathrm{L}_n(.)$	The Laguerre polynomial of degree n
$\mathcal{L}(.,.)$	Lagrangian function
m	Nakagami m parameter
MDZR	Missed detection zone ratio
N	Number of cooperative users
n	Number of quantization bins
$n(t)$	The additive white Gaussian noise
N_{B_k}	Number of users having observation in bin B_k
N_s	Number of measurement points
N_m	Total number of samples for unit measurement
N_g	Number of grid points
O_i	Observation of node i
p	Point for interpolation technique
p_i	Point of node i
p_e	Total probability of error
$p_e^{i,j}$	Probability of receiving B_j at the fusion center when B_i is observed by CR node and sent over the reporting channel
$\mathbf{P_D}$	Probability of detection
$\mathbf{P_F}$	Probability of false alarm
$\mathbf{P_D^A}$	Probability of detection under AWGN channel
$\mathbf{P_F^A}$	Probability of false alarm under AWGN channel

$\mathbf{P}_{\mathbf{D}}^{E}$	Probability of detection under EGC
$\mathbf{P}_{\mathbf{F}}^{E}$	Probability of false alarm under EGC
$\mathbf{P}_{\mathbf{D}}^{R}$	Probability of detection under Rayleigh channel
$\mathbf{P}_{\mathbf{F}}^{R}$	Probability of false alarm under Rayleigh channel
$\mathbf{P}_{H_i}^{A}(B_k)$	Probability of having local observation O_i in bin B_k under hypothesis H_i and AWGN channel
$\mathbf{P}_{H_0}^{A}(B_k)$	Probability of having local observation O_i in bin B_k under hypothesis H_0 and AWGN channel
$\mathbf{P}_{H_1}^{A}(B_k)$	Probability of having local observation O_i in bin B_k under hypothesis H_1 and AWGN channel
$\mathbf{P}_{H_1}^{ch,e}(B_i)$	Probability of having report at fusion center (including error) in bin B_i under hypothesis H_1 and channel ch
$\mathbf{P}_{H_0}^{ch,e}(B_i)$	Probability of having report at fusion center (including error) in bin B_i under hypothesis H_0 and channel ch
$\mathbf{P}_{H_1}^{ch,e}(B_2)$	Probability of having report at fusion center (including error) in bin B_2 under hypothesis H_1 and channel ch
$\mathbf{P}_{H_0}^{ch,e}(B_2)$	Probability of having report at fusion center (including error) in bin B_2 under hypothesis H_0 and channel ch
$\mathbf{P}_{H_i}^{R,e}(B_k)$	Probability of having report at fusion center (including error) in bin B_k under hypothesis H_i and Rayleigh channel
$\mathbf{P}_{H_1}^{R,e}(B_k)$	Probability of having report at fusion center (including error) in bin B_k under hypothesis H_1 and Rayleigh channel
$\mathbf{P}_{H_0}^{R,e}(B_k)$	Probability of having report at fusion center (including error) in bin B_k under hypothesis H_0 and Rayleigh channel
$\mathbf{P}_{\mathbf{D,CSDV}}^{\mathbf{ch}}$	Probability of cooperative detection with CSDV under channel **ch**
$\mathbf{P}_{\mathbf{F,CSDV}}^{\mathbf{ch}}$	Probability of cooperative false alarm with CSDV under channel **ch**
$\mathbf{P}_{\mathbf{F},OR}^{ch}$	Probability of cooperative false alarm with Or logic under channel ch
$\mathbf{P}_{\mathbf{D,CSDV}}^{\mathbf{R}}$	Probability of cooperative detection with CSDV under Rayleigh channel
$\mathbf{P}_{\mathbf{F,CSDV}}^{\mathbf{R}}$	Probability of cooperative false alarm with CSDV under Rayleigh channel

P_i^{rx}	The ideal RSS at the ith MCD
$\overline{P_i^{rx}}$	Mean RSS at the ith MCD
$P_{i,j}^{rx}$	The jth measured RSS value at the ith MCD
P^{tx}	The transmit power of the transmitter
$\widehat{P}^{tx}$	Estimated transmit power of the transmitter
P_{L_0}	Path loss correction
$\mathbf{Q}_u$	Marcum Q-Function with degree of freedom u
$\mathbb{R}^n$	n dimensional Euclidean space
$s(t)$	The transmitted signal of the primary user
S_i	Gaussian random variable representing disturbance
T	Observation interval
TW	Time bandwidth product
$\overrightarrow{w}$	Weight vector for global decision function of UniQCS detector
w_i	Weight of ith measurement for interpolation
W	Bandwidth
x^*	Optimal decision of optimization function
X	Constraint or domain set of an optimization function
$\mathcal{X}$	Feasible set
(x_i, y_i)	Location of ith MCD
(x_{pu}, y_{pu})	Location of active primary user
Y_{ij}	$j$thsample of the received signal by the node i
$Y(t)$	The complex signal received by the cognitive radio device
z	Interpolated value of IDW
z_i	Measurement of ith node
$Z(x)$	Stochastic process describing random field
$\hat{Z}(x_0)$	Estimated value of stochastic process at point x_0
Z_1^{CD}	Correct detection zone type-1
Z_0^{CD}	Correct detection zone type-0
Z_1^{MD}	Missed detection zone
Z_0^{FA}	False alarm zone

α	False alarm constraint of optimization problem or path loss exponent
β	Power parameter of IDW
γ	SNR
$\overline{\gamma}$	Mean SNR
γ_t	Output SNR of diversity paths
$\gamma(x, y)$	Semivariogram
$\mathbf{\Gamma}(.)$	Gamma function
$\mathbf{\Gamma}(.,.)$	Incomplete Gamma function
$\delta_{\overrightarrow{w}}(.)$	Global decision function
Δ	Delta value of uniform quantizer
Δ_c	Delta of center bin of quantizer
Δ^*	Optimal value for delta of uniform quantizer
$\varepsilon(x)$	Error in prediction at point x
$\widehat{\theta}$	Estimated value for θ
λ	Detector threshold or Lagrange multiplier vector
λ_1	First threshold of quantizer
λ_1^*	Optimal value for first threshold of uniform quantizer
$\lambda_i(x^*)$	Weight of measurement at x_i for point x^*
$\mu(x)$	Mean of random process output at point x
$\sigma_k^2(x_0)$	Kriging variance at point x_0
σ_s^2	Signal power variance
ϕ	Exponential mean of received power
χ^2_{2TW}	Central chi square distribution with $2TW$ degree of freedom
$\chi^2_{2TW}(2\gamma)$	Non-central chi square distribution with $2TW$ degree of freedom and non-centrality parameter 2γ

LIST OF ACRONYMS/ABBREVIATIONS

AWGN	Additive White Gaussian Noise
AP	Access Point
BEP	Bit Error Probability
BR	Bilişsel Radyo
BS	Base Station
BSC	Binary Symmetric Channel
CapEx	Capital Expenditures
CogNeA	Cognitive Networking Alliance
CR	Cognitive Radio
CRN	Cognitive Radio Network
CSDV	Cooperative Sensing with Decision Vector
dB	Decibel
DB	Database
dBm	Decibel referenced to one miliwatt
DSA	Dynamic Spectrum Access
DSO	Digital Switch Over
ECC	Electronic Communications Committee
ECMA	European Computer Manufacturers Association
ED	Energy Detection
EGC	Equal Gain Combining
EKB	Eşit Kazançlı Birleştirici
FCC	Federal Communications Commission
GA	Genetic Algorithm
GHz	Giga-hertz
GSM	Global System for Mobile Communications
IC	Interference Cartography
IDW	Inverse Distance Weighted
IEEE	Institute of Electrical and Electronics Engineers

IETF	Internet Engineering Task Force
KKT	Karush-Kuhn-Tucker
LIvE REM	LocatIon Estimation based REM
LP	Linear Programming
M2M	Machine-to-Machine
MAC	Medium Access Control Sub-layer
MCD	Measurement Capable Device
MGC	Maximal Gain Combiner
MHz	Mega Hertz
MVPD	Multichannel Video Programming Distributor
NSF	National Science Foundation
OpEx	Operational Expenditures
PAWS	Protocol to Access White Space Database
PHY	Physical Layer
PMSE	Programme Making and Special Events
PU	Primary User
QoS	Quality of Service
REM	Radio Environment Map
RF	Radio Frequency
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
ROH	Radyo Ortamı Haritası
RSS	Received Signal Strength
SC	Selection Combining
SNR	Signal to Noise Ratio
SSC	Switch and Stay Combining
SU	Secondary User
SUN	Smart Utility Networks
TG	Task Group
TVDB	TV band Database
TVWS	TV White Spaces

UHF	Ultra High Frequency
UMTS	Universal Mobile Telecommunications System
UniQCS	Uniform Quantization for Cooperative Sensing
VHF	Very High Frequency
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WhiteFi	WiFi extension over TVWS
WG	Working Group
WLAN	Wireless Location Area Network
WRAN	Wireless Regional Area Network
WSD	White Space Device

1. INTRODUCTION

The present regulation framework limits the adaptation of wireless communication systems to fast changing needs of the wireless applications. The limited nature of radio resources in the current unlicensed frequency bands is the main reason for this scarcity. Many radio systems that support consumer electronics and personal high data-rate networks operate in these unlicensed frequency bands. The demand for additional spectrum is growing faster than the technology is able to improve the spectral efficiency, although the latest research has remarkable success to increase the spectral efficiency and capacity in radio communications.

Research performed by various organizations such as the Federal Communications Commission (FCC) indicates that the assumption of spectrum sufficiency is far from reality; there is available spectrum since most of the spectrum allocated sits underutilized. Measurement campaigns in various parts of the world have supported the observation that static spectrum access leads to some portions of the spectrum to be overcrowded while some other portions to be underutilized [1]. In a completed National Science Foundation funded study of allocated spectrum utilization, researchers at Kansas University found an average U.S. spectrum occupancy of 5.2% with a maximum occupancy of 13.2% in New York City [2]. Therefore, we can claim that static spectrum access cannot manage the spectrum effectively and results in a perception that the spectrum is scarce. To improve the spectrum utilization, opportunistic spectrum utilization has been proposed wherein devices occupy the spectrum that has been left vacant. An illustrative example of opportunistic spectrum utilization is shown in Figure 1.1. The example opportunistic spectrum utilization starts and finds new opportunity whenever the Primary Users (PU) appear.

The term *Cognitive Radio (CR)* is first mentioned by Mitola III in 1999 [3]. He defines CR as "A radio that employs model based reasoning to achieve a specified level of competence in radio-related domains". However, in [4] Simon Haykin

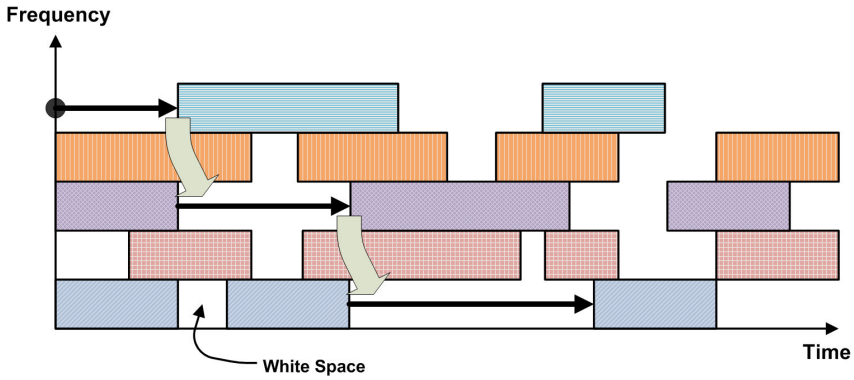


Figure 1.1. Conceptual example of opportunistic spectrum utilization.

defines a cognitive radio as "An intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming Radio Frequency (RF) stimuli by making the corresponding changes in certain operating parameters (e.g., transmit-power, carrier frequency, and modulation strategy) in real-time, with two primary objectives in mind:

- Highly reliable communications whenever and wherever needed.
- Efficient utilization of the radio spectrum."

We mainly adopt Haykin's definition, since it mentions being aware of surroundings, adapting according to environmental conditions, and learning mechanisms. Being aware of the surroundings is the most crucial part of the CR system since it includes sensing or querying the environment for primary activity, finding white spaces over spectrum bands and vacating the channel or adjusting the communication parameters, which is a must for not to disturb the PUs.

In this thesis, we focus on both internal and external sensing in Cognitive Radio Networks (CRN). In the internal sensing, CRs discover the spectrum opportunities via spectrum sensing whereas in the external sensing, an external entity provides the spectrum occupancy and related information. We propose a novel cooperative spectrum

sensing scheme, Uniform Quantization-based Cooperative Sensing (UniQCS) that uses uniform quantization and effective fusion strategy for internal sensing. The proposed method outperforms hard decision logic functions and achieves performance close to Equal Gain Combining (EGC) in terms of probability of detection, which constitutes an upper bound since we do not differentiate the nodes. In the external sensing, CRs access spectrum via querying geolocation databases, which keep the static information about radio frequency bands and the environment.

Another contribution of this thesis is the introduction of the PU LocatIon Estimation based (LIvE) *Radio Environment Map (REM)* (improved geolocation database) construction technique for external sensing. LIvE is compared with the well-known REM construction techniques in shadow and multipath fading channels. We focus on construction of the interference map of the network operation area just utilizing the measurements from few nodes and additionally the channel parameters such as path loss exponent and correction. LIvE REM construction technique utilizes channel parameters and knowing the channel parameters improves the quality of the constructed REM significantly.

1.1. Key Contributions

Thesis contributions are two fold; one is about internal sensing and the other is about external sensing.

The contributions on internal sensing can be summarized as:

(i) A cooperative sensing that uses a uniform quantizer is proposed.
(ii) Global decision techniques for systems that use quantization are generalized.
(iii) Probability of detection and false alarm formulations are derived for UniQCS.
(iv) An optimization problem that maximizes the probability of detection is solved.
(v) Robustness of UniQCS to false reports and an imperfect reporting channel is analyzed.

(vi) Detection performance of UniQCS is close to EGC (which provides an upper bound for homogeneous nodes).

The contributions on external sensing can be summarized as:

(i) We define a new useful metrics, correct detection and false alarm zone ratios, for the REM quality.
(ii) The proposed REM construction technique, LIvE, achieves high performance in terms of Root Mean Square Error (RMSE) and the correct detection zone ratio.
(iii) The time complexity of LIvE is half of that of Kriging while still performing better.
(iv) LIvE REM construction technique performance does not depend on the closeness of the PU to the measurement capable devices.

1.2. Thesis Outline

First, in Chapter 2 we review the related literature in order to clearly locate our main contributions in the literature. In this chapter, we summarize the outstanding works related to spectrum sensing and our proposed methods.

The contributions of the thesis is explained in two chapters: Cooperative spectrum sensing via uniform quantization in CRNs (Chapter 3) and REM construction via PU location estimation for external sensing in CRNs (Chapter 4).

Chapter 3 presents the proposed cooperative sensing that utilizes quantization. The proposed method can be easily applied to the existing CR systems for achieving performance that is close to EGC via marginal overhead bits. The performance of the proposed method is also analyzed under an imperfect reporting channel and when there are false reports. We mainly focus our analysis on the performance in terms of probability of detection under Rayleigh fading channel assumption. We also analyze the effect of number of cooperating nodes on the system performance. For the imperfect

reporting channel case we include the analysis on effect of probability of error in the reporting channel.

Chapter 4 first highlights the basic issues in REM construction from an estimation perspective. We also provide a useful metric for REM accuracy evaluation. Next, we present the mean RSS formulations depending on the channel model. After the mean RSS formulations, we propose the LIvE REM construction technique. The performance of the proposed technique is analyzed under log-normal shadowing and Rayleigh fading channel conditions.

Finally, Chapter 5 concludes this thesis by summarizing the key contributions and in addition it elaborates a discussion on the possible future directions.

2. RELATED WORK

2.1. Cognitive Radio

The need for higher data rates is increasing as a result of new wireless services and applications. However, spectrum allocation and management is still based on the old fashioned techniques from the very early days of wireless communications. Given the limitations of the natural frequency spectrum, it is obvious that the current static frequency allocation schemes cannot meet the requirements of this increasing demand. Measurement campaigns in various parts of the world have supported the finding that the static spectrum access leads to some portions of the spectrum to be overcrowded while some other to be underutilized [1]. So called the static spectrum access falls short of effective spectrum management and results in a perception that the spectrum is scarce.

This inefficiency in the spectrum usage and allocation necessitates a new communication paradigm to exploit wireless spectrum opportunities [5]. Hence, Dynamic Spectrum Access (DSA) is proposed, which allows wireless devices to operate opportunistically in spectrum holes/opportunities until the license holders, PUs, begin to use the band. The key enabling technology of DSA is the CR technology that is built on a software defined radio, and CR is defined as an intelligent wireless communication system that is aware of its environment and uses the methodology of understanding-by-building to learn from the environment and adapt to statistical variations in the input stimuli [4].

The description of CR by Mitola and Maguire in their seminal paper [3] focuses on the radio knowledge representation language and how the cognitive radio can enhance the flexibility of personal wireless services. CR is formally defined by the FCC [6] as a radio that can change its transmitter parameters based on interaction with its environment. The ultimate objective of the cognitive radio is to obtain the best available

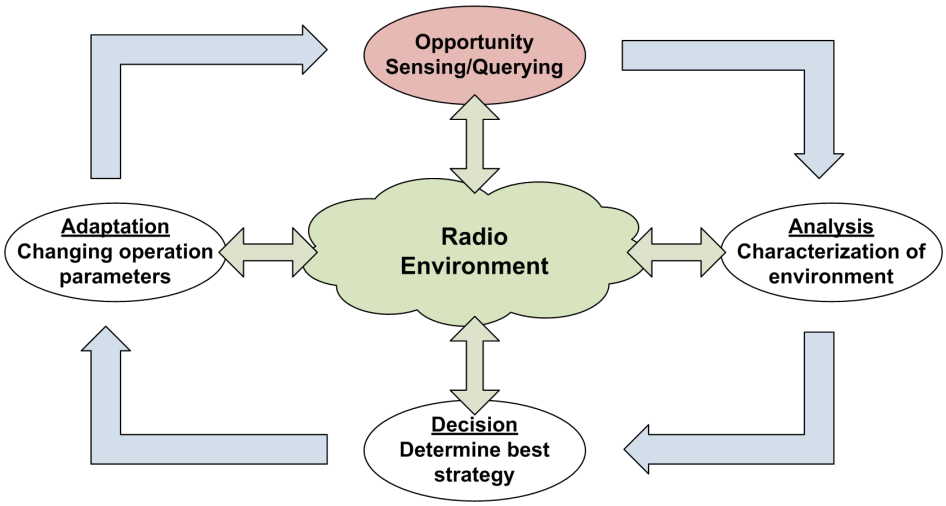


Figure 2.1. Simplified cognitive cycle.

spectrum through *cognitive capability* and *reconfigurability*. Tasks required for adaptive operation are: Spectrum sensing, spectrum analysis, spectrum decision [4, 7, 8]. Once the operating spectrum band is determined, the communication can be performed over this spectrum band. However, since the radio environment changes over time, space, and frequency, the cognitive radio device should keep track of the changes of the radio environment. If the current spectrum band in use becomes unavailable, the spectrum mobility function is performed to provide a seamless transmission. Any environmental change during the transmission, such as primary user appearance, user movement, or traffic variation, can trigger this adjustment. The main tasks of CR in cognitive cycle are summarized in Figure 2.1.

After the initial works in the literature on the architecture and basic concepts of CR, current research extensively focuses on how to realize DSA and more "cognitive" devices via artificial intelligence and machine learning techniques. Spectrum sensing, medium access and resource allocation, spectrum sharing among operators, and security are some of the topics that have attracted interest in the CR domain. We mainly focus on the spectrum sensing or querying part since it is the most crucial part of the cognitive cycle. Without finding the opportunities CR cannot utilize white spaces without harming the owner of the spectrum band.

2.2. Spectrum Sensing

Spectrum sensing is the most important task in the cognitive cycle for the realization of cognitive radio. Since cognitive radios are considered lower priority or Secondary Users (SU) of spectrum allocated to a primary user, a fundamental requirement is to avoid interference to potential PUs in their vicinity. On the other hand, PU networks are not required to change their infrastructure for spectrum sharing with cognitive networks. Therefore, cognitive radios should be able to independently detect PU presence through spectrum sensing schemes. Although spectrum sensing is traditionally considered as measuring the spectral content or measuring the interference over the spectrum, when the ultimate cognitive radio is considered, it is a more general term that involves obtaining the spectrum usage characteristics across multiple dimensions such as time, space, frequency, and code [9].

In a nutshell, the goal of spectrum sensing is to decide between two hypotheses

$$\begin{aligned} Y(t) &= n(t) && H_0 \ (white\ space), \\ Y(t) &= h \times s(t) + n(t) && H_1 \ (occupied) \end{aligned} \tag{2.1}$$

where $Y(t)$ is the complex signal received by the cognitive radio device, $s(t)$ is the transmitted signal of the primary user, $n(t)$ is the *Additive White Gaussian Noise (AWGN)*, and h is the complex gain of the ideal channel.

Common transmitter detection methods in the literature are *Energy Detection (ED)*, *Matched Filter Detection*, and *Cyclostationary Detection*. The ED based approach, also known as radiometry or periodogram, is the most common way of spectrum sensing in high SNR conditions since it does not require any a priori knowledge of primary signals and has much lower computational and implementation complexity [10–21]. We follow the ED approach due to these advantages with enhancing the detection process by using quantization and cooperation.

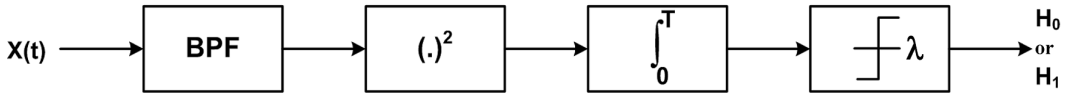


Figure 2.2. Block diagram of the conventional energy detector.

The conventional energy detector depicted in Figure 2.2 uses a single threshold to determine the presence or absence of the signal. The input band-pass filter selects the center frequency, f_c, and the bandwidth of interest in Hz, W. This filter is followed by a squaring device to measure the received energy and an integrator which determines the observation interval in seconds, T. Finally, the output is compared with a threshold, λ, to decide whether the signal is present. The presence of a primary user in AWGN as well as flat fading Rayleigh channels results in different ED outputs and also ED output is affected by the dynamic channel conditions.

2.2.1. Conventional Energy Detection in AWGN Channel

Under an AWGN channel, the energy received ($O_i = \sum_{j=1}^{2TW} Y_{ij}^2$) by a secondary user i follows the distribution

$$f(O_i|\gamma) \approx \begin{cases} \chi^2_{2TW} & H_0 \\ \chi^2_{2TW}(2\gamma) & H_1 \end{cases} \tag{2.2}$$

where χ^2_{2TW} and $\chi^2_{2TW}(2\gamma)$ represent the central and the non-central chi square distributions [10, 22, 23], TW and γ represent the time bandwidth product and SNR, respectively. Under AWGN channel conditions, the SNR value is fixed and it affects the separation between conditional probability distribution functions. Example probability distribution functions with different SNR values are depicted in Figure 2.3.

A high SNR separates the distributions enough to decide safely and with a reasonable probability of error. However, under low SNR conditions, it is difficult to distinguish between H_0 and H_1 probability distributions.

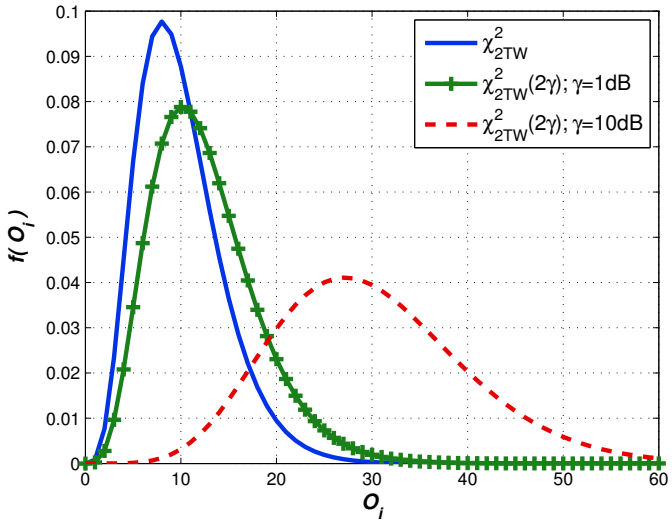


Figure 2.3. Example O_i probability distribution functions with different SNR values ($TW = 5$).

In the single threshold ED, the decision for H_0 and H_1 depends solely on λ. The probability of detection, $\mathbf{P_D^A}$, and the probability of false alarm, $\mathbf{P_F^A}$, for a single SU under AWGN channel can be calculated using λ with the exact closed-form equations

$$\mathbf{P_D^A}(\lambda) = \mathbf{P}\{O_i > \lambda \ \ |H_1\} = \mathbf{Q}_{TW}(\sqrt{2\gamma}, \ \sqrt{\lambda}) \tag{2.3}$$

$$\mathbf{P_F^A}(\lambda) = \mathbf{P}\{O_i > \lambda \ \ |H_0\} = \frac{\mathbf{\Gamma}(TW, \lambda/2)}{\mathbf{\Gamma}(TW)} \tag{2.4}$$

where $\mathbf{Q}_u$ is the Marcum Q-Function with degree of freedom u and $\mathbf{\Gamma}(.)$, $\mathbf{\Gamma}(.,.)$ represent the Gamma and Incomplete Gamma Functions, respectively [22, 23].

2.2.2. Conventional Energy Detection in Rayleigh Channel

The mobile radio channel is characterized by the multipath reception. The signal reaching to the receiver contains not only a direct line-of-sight radio wave, but also a large number of reflected radio waves [24]. Even worse in urban centers, the line-of-sight is often blocked by obstacles. The basic model of Rayleigh fading assumes the

received multipath signal to consist of a large number of reflected waves with i.i.d. inphase and quadrature amplitudes, and not the line-of-sight component. Therefore, we focus on the Rayleigh fading channel caused by multipath reception. Under the Rayleigh fading channel, the probability of false alarm, $\mathbf{P_F^R}$, remains the same as in the case of AWGN channel since it depends only on the distribution of noise. However, under the Rayleigh channel with no diversity, the signal amplitude follows Rayleigh distribution [22]. Therefore, SNR γ follows the exponential pdf

$$f(\gamma) = \frac{1}{\overline{\gamma}} \exp(-\gamma/\overline{\gamma}), \quad \gamma \geq 0. \tag{2.5}$$

The average $\mathbf{P_D}$ in this case, $\mathbf{P_D^R}$, can be calculated by averaging Equation 2.3 over Equation 2.5 [11, 22].

$$\mathbf{P_D^R}(\lambda) = \int_{\gamma} \mathbf{P_D^A} f(\gamma) d\gamma \tag{2.6}$$

$$\mathbf{P_D^R}(\lambda) = e^{-\lambda/2} \sum_{n=0}^{TW-2} \frac{1}{n!} \left(\frac{\lambda}{2}\right)^n + \left(\frac{1+\overline{\gamma}}{\overline{\gamma}}\right)^{TW-1} \left[e^{-\frac{\lambda}{2(1+\overline{\gamma})}} - e^{-\frac{\lambda}{2}} \sum_{n=0}^{TW-2} \frac{1}{n!} \left(\frac{\lambda\overline{\gamma}}{2(1+\overline{\gamma})}\right)^n \right] \tag{2.7}$$

Under the Rayleigh fading channel in which diversity paths are i.i.d., the output SNR, γ_t, is the sum of the SNRs on all branches and can be used for evaluating the average $\mathbf{P_D}$ for EGC scheme.

The pdf of γ_t for i.i.d. Rayleigh branches is given by

$$f(\gamma_t) = \frac{1}{(L-1)!\overline{\gamma}^L} (\gamma_t)^{L-1} \exp(-\gamma_t/\overline{\gamma}) \tag{2.8}$$

where L is the number of i.i.d. diversity branches. The pdf in Equation 2.8 is similar to the pdf of SNR in the Nakagami channel. The Nakagami parameter m can be viewed

as a diversity order. Hence, the average $\mathbf{P_D}$ for the EGC scheme, $\mathbf{P_D^E}$, is obtained by replacing m, $\overline{\gamma}$, and TW by L, $L\overline{\gamma}$, and LTW, respectively [22]. We take the diversity order L equal to the number of cooperating nodes for evaluating $\mathbf{P_D^E}$. Hence, the $\mathbf{P_D^E}$ formulation becomes

$$\mathbf{P_D^E}(\lambda) = \alpha \left[\Psi + \beta \sum_{n=1}^{LTW-1} \frac{(\lambda/2)^n}{2n!} \, {}_1F_1(L; n+1; \frac{\lambda}{2} \frac{\overline{\gamma}}{(1+\overline{\gamma})}) \right] \tag{2.9}$$

where ${}_1F_1(.;.;.)$ is the confluent hypergeometric function,

$$\alpha = \frac{1}{\Gamma(L) 2^{L-1}} \left(\frac{1}{\overline{\gamma}} \right)^L, \tag{2.10}$$

$$\beta = \Gamma(L) \left(\frac{2\overline{\gamma}}{1+\overline{\gamma}} \right)^L \exp(-\lambda/2), \tag{2.11}$$

and

$$\begin{aligned} \Psi \;=\; & \frac{2^{L-1}(L-1)! \, \overline{\gamma}}{(1/\overline{\gamma})^L (1+\overline{\gamma})} e^{-\frac{\lambda}{2(1+\overline{\gamma})}} \left[(1+\frac{1}{\overline{\gamma}}) (\frac{1}{1+\overline{\gamma}})^{L-1} \mathrm{L}_{L-1}(-\frac{\lambda\overline{\gamma}}{2\,(1+\overline{\gamma})}) + \right. \\ & \left. \sum_{n=0}^{L-2} (\frac{1}{1+\overline{\gamma}})^n \mathrm{L}_n(-\frac{\lambda\overline{\gamma}}{2\,(1+\overline{\gamma})}) \right] \end{aligned} \tag{2.12}$$

where $\mathrm{L}_n(.)$ is the Laguerre polynomial of degree n [22]. Since we do not differentiate the nodes with different SNR levels, EGC constitutes an upper bound for our optimization problems. These formulations are focusing on local ED, however cooperation increases the performance of the detection process.

2.2.3. Conventional Decision Function for Cooperative Sensing

To take advantage of the spatial diversity in the wireless channel, cooperative spectrum sensing methods have been proposed in [25–28]. It is shown analytically and through numerical results that cooperative sensing provides significant higher spectrum

capacity gains than local sensing [29]. The detection performance is determined by the quality of local observations and the quality of the information received by the fusion center where the cooperation is achieved. Therefore, the number of quantization bins, the number of bits sent for sensing reports, and the global decision logic affect the overall system performance. We provide background information on conventional decision functions for cooperative sensing schemes in this subsection.

In the case of cooperative sensing, sharing information among CRs and combining results from various measurements is a challenging task. The shared information can be soft or hard decisions made by each cognitive device [30]. The results presented in [30,31] show that the soft information-combining outperforms the hard information-combining method in terms of the probability of missed opportunity. On the other hand, hard-decisions are found to perform as good as soft decisions when the number of cooperating users is high [32]. *Therefore, our proposed detection scheme focuses on softened hard decision scheme via quantization for bandwidth limited reporting channel and incorporates the cooperation for increasing the performance of the detector under a low SNR regime.*

The optimum fusion rule for combining the sensing information is the Chair-Varshney rule which is based on log-likelihood ratio test [33]. Likelihood ratio tests are used for making classification using decisions from secondary users in [30], [34–36]. Various simpler techniques for combining sensing results are employed in [11]. The performances of EGC, Selection Combining (SC), and Switch and Stay Combining (SSC) are investigated for energy detector based spectrum sensing under Rayleigh fading. The EGC method is found to have a gain of approximately two orders of magnitude gain while SC and SSC having one order of magnitude gain.

When hard decisions are used, And, Or, Majority, and M-out-of-N methods can be used for combining the information from different cognitive radios [37]. In the And-rule, all sensing results should be H_1 for deciding H_1, where H_1 is the alternate hypothesis, i.e. the hypothesis that the observed band is occupied by a primary user.

In the Or-rule, the fusion center decides H_1 if any of the received decisions plus its own is H_1. In the Majority-rule, for deciding H_1 it is required to have majority of the nodes having decision H_1. M-out-of-N rule outputs H_1 when the number of H_1 decisions is equal to or larger then M where there are N cooperating nodes. Therefore, And, Or, and Majority rules may be considered as special cases of M-out-of-N rule.

A comprehensive classification of cooperative sensing is examined in [38], and research challenges are listed. Various cooperative sensing techniques are studied in [39–42]. The effect of user collaboration in Rayleigh fading channel is studied in [43]. Authors show that the increase in the number of cooperative users results in significant increase in the sensing performance and the spectrum utilization.

We also consider the cooperative sensing methods which utilize quantization or analyze the performance under an imperfect reporting channel. In [44], cooperative sensing and quantization schemes (1-bit or hard decision) are investigated for multiple primary bands. We also focus on quantization which is more general in terms of number of quantization bins compared to 1-bit quantization. The effects of the imperfect reporting channel to the sensing performance are analyzed for hard decision logic functions in [42]. We extend this study for functional global decision techniques. The spectrum sensing scheduling and sensing time are analyzed in [45, 46]. In [47], authors analyze cooperative sensing under bandwidth constraints. In [48], the authors examine the optimal quantizer for signal detection locally. The work is extended in [49], by using evidence theory based cooperative spectrum sensing with efficient quantization. We also focus on signal detection via quantization but we consider optimizing the overall process rather than local optimization. Similarly in [25], detection and false alarm probabilities are derived with consideration of errors in the reporting channel due to fading. Only the hard decision logic M-out-of-N is studied in the paper. In [50] and [51], cooperative sensing via quantization is studied for 2-bit and 3-bit quantization. However, the global decision logic is static in terms of weights. Our work generalizes the global decision logic, n-bit quantization, and improves weights by using a genetic algorithm. In [52] Bit Error Probability (BEP) wall is introduced and per-

formance analysis for M-out-of-N is done. The authors analyze the SNR loss due to BEP without considering the overall detection probabilities.

2.3. Geolocation Database and Radio Environment Map

Starting from 2009, in the USA and some parts of the Europe, analog TV broadcast service turning off has started, and has replaced by digital TV services [53]. Digital TV broadcasting has higher spectral efficiency compared to analog broadcasting as it compresses a multiple of TV channels into a single 6/7/8 MHz channel used in analog transmission. This migration process, so called Digital Switch Over (DSO), has resulted in a valuable portion of the spectrum to be freed up. Some of these freed bands are completely cleared from TV services and they are expected to be auctioned for alternate wireless applications, while the remaining portion may become vacant only in some geographical areas due to co-channel and adjacent channel interference issues. FCC, the regulatory agency in the USA, has carried out an auction for 700 MHz bands after the completion of DSO in 2009 in the USA while the remaining TV freed portion is opened for unlicensed operation enabled by dynamic spectrum access. Similarly, in UK the cleared spectrum of total 128 MHz bandwidth at 550-606 MHz and 790-862 MHz are auctioned by Ofcom after the completion of DSO in 2011. DSO is almost completed in Japan except the three prefectures devastated by the 11th March earthquake. Similarly, in Canada, and some parts of Europe DSO is almost completed. It is scheduled to be completely finished in most of Europe by 2012. South Korea, Australia, China, and Brazil are expected to switch off not later than 2015 [54]. The portion of the UHF and VHF frequency at 42-870 MHz emerging after DSO are referred to as TV White Spaces (TVWS). Initiated by the FCC in USA, use of TVWS for dynamic spectrum access has attracted worldwide interest such as the UK, Brazil, Japan, India, Singapore, and China [55]. Turkey launched trial transmissions in 2006 and originally planned to gradually do the switch by 2014 [56]. TVWS represents a frontier for wireless communications as it opens a valuable portion of the spectrum with favorable propagation characteristics for wireless communications. Main advantages of TVWS can be summarized as follows;

- Superior coverage compared to ISM at 2.4 GHz in the scale of kilometers, not in meters as in WLANs (4 times larger radius, 16 times larger coverage area)
- Decreased network Capital Expenditures (CapEx) and Operational Expenditures (OpEx) cost due to better areal capacity and fewer cell sites needed (less BS/AP is required for complete coverage compared to ISM)
- Lower power consumption due to operation at lower frequencies, and fewer network devices
- More appropriate for home/in-building networking due to its penetration capability through walls compared to WLANs.

Opening of TVWS has been an effective catalyst for the first practical but primitive applications of CRNs. As primary users of the TVWS, the TV broadcast stations are stationary and high power transmitters that do not change their transmission properties frequently. Therefore, cognitive access in these TVWS bands is less challenging compared to highly dynamic bands. Above all, the major challenge of the primary user detection is partly eliminated in TVWS. However, coexistence with other incumbents operating at TVWS bands is more challenging. For instance, equipments classified under Programme Making and Special Events (PMSE) devices like wireless microphones are difficult to detect compared to the stationary TV towers as they may be mobile and transmit at lower power levels. This issue is tackled by reserving a channel (e.g. channel 37) for the exclusive use of PMSE devices. FCC defined eight classes of services to be protected as follows [57]:

(i) Fixed Broadcast Auxiliary Service (BAS) links
(ii) Receive sites of TV translator
(iii) Low power TV
(iv) Class A TV stations and Multichannel Video Programming Distributors (MVPDs)
(v) Private land mobile and commercial mobile radio service operations
(vi) Offshore radio telephone service operations
(vii) Radio astronomy operations at specific sites
(viii) Low power auxiliary service operations

2.3.1. Potential of TVWS: How Much White Space is Available?

It is a fact that wireless technologies operating at unlicensed bands, e.g., ISM bands, have penetrated into every part of daily life. Hence, comparing the potential of TVWS with that of ISM bands can give an idea on the value of TVWS. There are three non-overlapping channels of 20 MHz each in ISM bands at 2.4 GHz, making up a total of 60 MHz. The TVWS (interleaved spectrum) has a bandwidth of 232 MHz in the UK. While this is the theoretical raw capacity, exact availability of the spectrum varies from place to place. White space measurements show that almost 50% of locations in the UK have more than 150 MHz bandwidth while 90% have more than 100 MHz capacity [53]. Similarly, measurement campaigns in many countries all over the world like the USA [58], Japan [59], and Spain indicate the potential of TVWS in meeting the escalating wireless connectivity requirements.

2.3.2. Milestones of TVWS Evolution

The idea of opportunistic access in UHF/VHF has its roots in 2004, dating back to FCC's notice of proposed rule making to open vacant TV white space to the unlicensed use of wireless devices. Till then, most of the research and regulations originated from the USA, and are followed by Ofcom, regulatory agency of the United Kingdom. UK, prominently faster in Europe, conducted research and analysis on the potential of TVWS in UK and released a consultation for possible use of TVWS in 2009. Spectrum regulators in Europe also consider TVWS as an opportunity for wireless communications and started a Working Group (WG SE 43 [60]) under Electronic Communications Committee (ECC) in 2009. The first standardization study was initiated under IEEE 802.22 [61] in November 2004, which specifically aims the use of TVWS for fixed, point-to-multipoint rural broadband access and defines PHY/MAC specifications. However, it proceeded relatively slow and only released IEEE 802.22-2011TM standard in July 2011 [62]. Cognitive Networking Alliance (CogNeA [63]), an industry-led alliance was formed in December 2008, and published the first standard ECMA 392 [64] in December 2009 for low power portable/personal devices to operate over TVWS [64]. IEEE

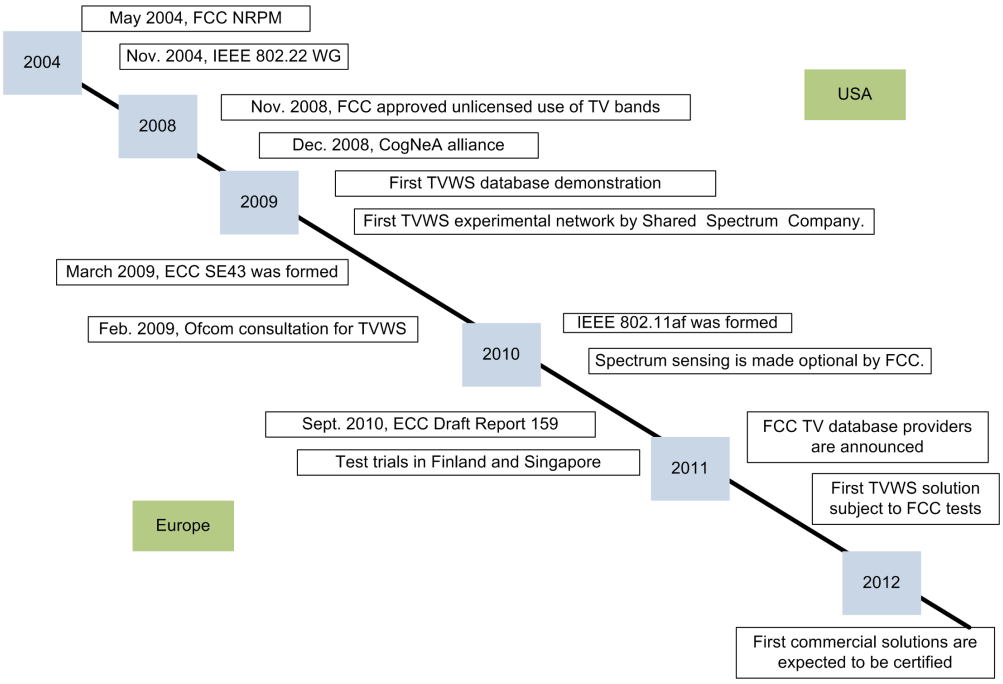


Figure 2.4. Milestones of TVWS evolution in the USA and Europe.

802.11af Task Group (TG) formed in September 2009 and expected to be finalized in 2012 examines the necessary changes for operation of WiFi on TV bands to realize the so called *WhiteFi* technology. Existing IEEE 802.11 protocols have to be adapted to the peculiarities of the white spaces. Basically, there are three fundamental characteristics of white spaces different from ISM: the narrower bands (6,7 or 8 MHz compared to the 20 MHz ISM bands) in the fragmented spectrum, communications on variable bandwidth channels, and temporal variations in spectral resources due to PUs of white spaces [65]. Milestones of TVWS evolution are presented in Figure 2.4.

IEEE 802.22 targets to decrease the "digital divide" while IEEE 802.11af aims to decrease the congestion in ISM bands by offloading some traffic to TVWS, and CogNeA mainly aims to realize the vision of fully connected homes via home networking on TVWS. Currently, as of December 2011, there are no commercial IEEE 802.22 device or TVWS devices. However, commercial solutions are expected to emerge in 2012.

2.3.3. Geolocation DB Architecture

A geolocation database, sometimes referred to as TV band Database (TVDB), aims to construct a realtime view of the spectrum occupancy at the TV bands at each location. TV devices query this database with their geolocation in terms of latitude and longitude, and in return receive the list of unoccupied frequencies before initiating a communication.

A TVDB must provide the available channel list possibly including time constraints (e.g. duration of availability) for the channel list and maximum transmission power. It keeps various information about PU (possibly CR) transmitters; TV tower location, antenna height, device ID, user type (PU or CR), device technology, device geolocation, transmitter power, technology, operation channel(s) and duration of use [66]. Basically, a TVDB provides three services: data repository, registrant service and query service. These three services can be provided by a single entity or multiple service providers can present one of them. The data repository is the database where permanent service and user information are stored while registrant service provides the registration of both the protected entities and the White Space Devices (WSDs). Query service provides both the WSD interface and public interface. The WSD query service receives the WSD related information and returns the channel available list while the public interface provides a web based system for free and public access. Figure 2.5 (adapted from [67]) depicts a simplified overview of TVDB architecture and its interfaces to both protected users and WSDs.

In December 22, 2011, FCC approved the Spectrum Bridge's database [57], and starting from January 2012, all incumbents including wireless microphones are mandated to register to this database in order to be protected from secondary access. Additionally, spectrum and country agnostic database messaging interface and query mechanisms need to be further defined. IETF PAWS (Protocol to Access White Space DB) [68], the standardization on how to query the database and data model for the queries and responses, is in development phase and is expected to be released in the

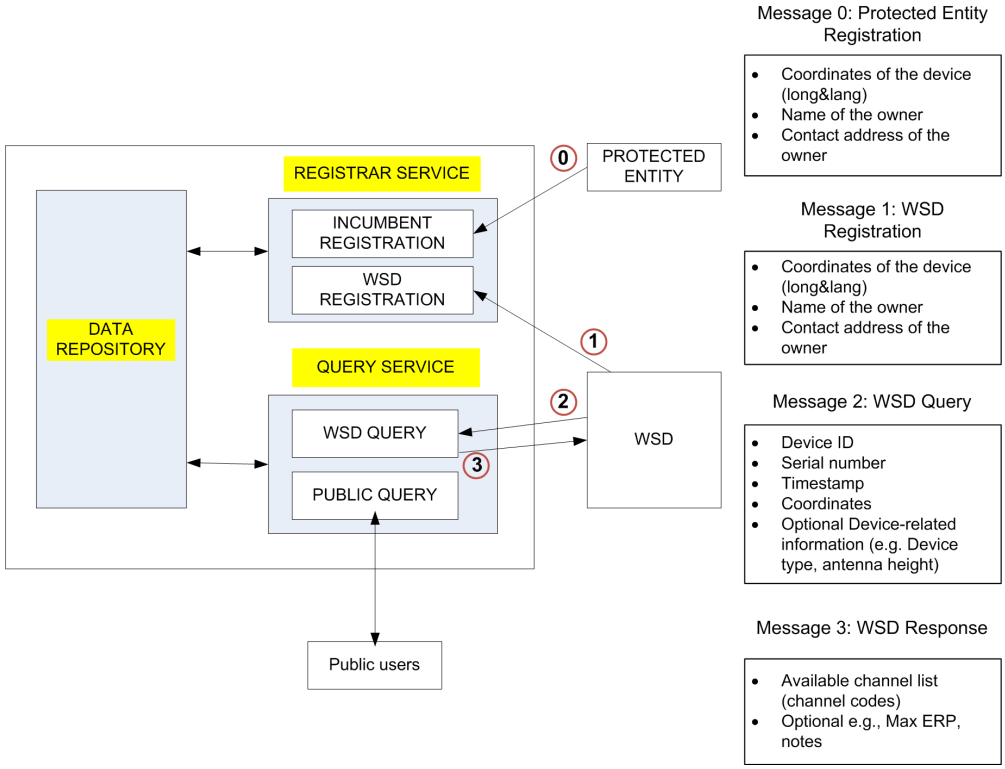


Figure 2.5. TVDB services and message formats.

first half of 2012. Several Internet web sites [69–72] provide the coverage area of TV stations at a specific location, visualize them on a map, and also present the list of available channels.

According to the FCC requirements, each database must periodically synchronize with other databases and the FCC's database itself in order to stay up-to-date, and disseminate the registered device information to each repository in the system. TVDB must be secure, spoofing-free, and must meet reliability requirements [73].

2.3.4. TVWS Application Proposals

Opening valuable portion of the TV bands to secondary access sparks the application proposals for these bands. There is much discussion about which applications that will be used in the TVWS. Obviously we have rural broadband and super-WiFi, but there are many others. In fact, most generally used applications for wireless com-

munications could potentially be used in TVWS. However, some applications are more attractive than others when considering commercial and technical aspects.

- *SUN, Machine-to-machine communications:* Efficient use and management of utilities such as electricity, natural gas and water has lately been more crucial with increasing concerns on ecology and energy resources. In this sense, Smart Utility Networks (SUN) deployed at houses provide an infrastructure to monitor, control, and possibly enable the best operation mode while decreasing waste of resources as well as decreasing the human cost of stepping to the utility meters and record readings. Good penetration properties make TVWS cut out for smart metering as these meters are mostly placed in buildings rather than outdoor cabinets. In addition, signals at TVWS propagate in a wider area making the smart metering data collectors deployed by the service providers less dense [74]. Spectrum Bridge, in partnership with Google and Plumas-Sierra Rural Electric Cooperative, in 2010 launched a trial Smart Grid network over TVWS [75] and demonstrated the viability of a SUN over TVWS in a cost-effective manner. Similarly, Machine-to-Machine (M2M) communications, traffic monitoring, and similar remote monitoring systems can enjoy the great bandwidth introduced by TVWS.
- *Broadband rural access:* In low population rural areas where it is difficult or expensive to deploy a cellular network, an infrastructure based TVWS network as defined in IEEE 802.22 can be deployed. It is a good choice since TVWS can propagate over long distances resulting in a coverage area considerably larger than that of Global System for Mobile Communications (GSM), Universal Mobile Telecommunications System (UMTS), and IEEE 802.16 WiMAX, e.g. a 802.22 BS has a coverage area of 30-100 km in radius [76].
- *WiFi extension over TVWS (WhiteFi):* Density of WLANs is increasing day by day, which increases the complexity of channel allocation and inter-AP coordination. This issue is mostly experienced in dense urban areas. WhiteFi can remedy this issue by offloading some portion of the traffic from ISM based WLANs to TVWS based WLANs. Similarly TVWS can help traffic offloading from already

congested WLANs and 3G/4G networks.

- *Home networking:* White spaces can be a good enabler for high speed home networking at low power. As many tools and equipments are going wireless in residential places, short range communications via Bluetooth or WiFi may experience distortions and may fall short of providing sufficient Quality of Service (QoS). Instead, as advocated by CogNeA and ECMA 392 standard [77], TVWS bands can be used at home networking.
- *Cognitive Femtocells:* Motivated by the fact that providing indoor cellular coverage is a burden on the operators, femtocells are proposed to close the coverage gap in indoor areas such as residential areas or small offices. Due to the transmitter and receiver being in close proximity, femtocells provide high data rate at low power levels, which makes them ecofriendly. However, from a technical perspective femtocells are challenging to operate due to macrocell-to-femtocell and femtocell-to-femtocell interference if they share all the operator's licensed spectrum. Instead, femtocells via cognitive functionalities, dubbed as *cognitive femtocells*, can discover the white spaces and operate through these bands without experiencing inter-tier interference. Similarly, they can coordinate spectrum sharing at the femtocell layer and thus intra-tier interference can be mitigated.

2.3.5. Basics of the Radio Environment Map Concept

REM was first defined as an abstraction of a real-world environment storing multi-domain information (e.g., PUs, policies, terrain data) [78]. However, it can also be considered more generally as an intelligent network entity that can further process the gathered information, inspect the spatio-temporal characteristics, and derive a map of the RF environment [79]. REM is a promising concept for efficient CRN operation without extensive burden on CR nodes as it can be considered as the *cognitive engine* located in the network.

REM introduces environment awareness that would be harder to acquire by individual CR capabilities via extensive spectrum analysis. Hence, REM can also be seen

as the network support turning simple nodes into intelligent CRs. Simplifying devices leads to cheaper devices which in turn eases the commercialization of CRs. Moreover, REM eases the adaptation of a CR to a new environment (e.g., a new country where regulations are different) in which the individual CR has not constructed any knowledgebase yet [80]. Currently, geolocation databases have come into reality by FCC's intent for TVWS access through databases. REM uses geolocation database hence it is a kind of improved geolocation that fuses geolocation information and much more. Additionally, REM stores more dynamic data hence more appropriate for dynamic CR operations which is not the case for TVWS. On the other hand, REM still needs to be explored further; what kind of data to store, how to store, and how to utilize data for cognition are yet to be answered. Despite significant potential of REM for cognitive networking, research on REM is still in an early stage.

2.4. Radio Environment Map Construction Techniques

REM construction is a wide concept referring to creating a complete map of CRN coverage area. However, we focus only on deriving interference level at each point of the CRN coverage area. This is also referred to as Interference Cartography (IC). IC forms a map of signal strength of RF environment as a function of spatial coordinates in a predefined area, measurements from CRs, and possibly time (for a dynamic environment).

Typically, a REM construction method performs the following tasks: data gathering, data processing to interpret the underlying model and deciding on the state of each pixel. If data is collected only internally from the nodes in CRN, REM may not be sufficiently accurate. Additionally, as REM construction costs time and power on CR nodes, it may become a burden on the network. Therefore, REM construction can be delegated to external entities such as sensor nodes dedicated to this task. In this case, CR nodes can also contribute to the environment modeling in their locality and fine tune the environment model based on their experience.

Some construction mechanisms may utilize propagation modeling, which is determined by the terrain properties between the transmitter and the location. If not modeled appropriately, a significant portion of the spectrum may be wasted due to false alarms while it may also lead to miss detections. Therefore, REM construction methods should retrieve terrain data stored internally or externally, and compute the spectrum availability using these terrain data.

In the literature, REM construction techniques can be broadly put into two classes: spatial statistics based methods [81] and transmitter location determination based methods [82], which are also referred to as direct and indirect methods, respectively [80].

2.4.1. Spatial Statistics Based Methods

Spatial statistics describes the statistical properties of a given area utilizing the spatial correlational structure of this region. Using spatial statistics and given the measurements at specific locations, missing data at areas without any measurements can be estimated as a function of measured data. The fusion of the data from the examined area is commonly based on different types of interpolation methods. Interpolation is based on the basic principle that geographically close locations are more related to each other compared to the more distant locations. Possible interpolation techniques in the literature are Kriging, Inverse Distance Weighted (IDW), and nearest neighbor interpolation. The frequency or volume of measurements, required accuracy in measurements, and density of measurement point are key factors determining the performance of each method. Although Kriging requires more measurement points, it is the most commonly applied technique in the literature [79, 83] due to its higher precision.

Specific fields of spatial statistics are random fields and point processes. Random fields model the phenomena under study as a continuous region while point processes characterizes the locations of *events*. In a network context, point processes can be used

to describe the location distribution of transmitters and receivers. Stochastic geometry of the wireless nodes is crucial to be represented in the design of network protocols. As PU density determines the interference, point processes are useful for the REM construction.

Instead of storing the raw data of locations, point processes describing the orientation of the nodes in the network with a number of parameters can be stored in the REM [81]. This descriptive and compact model reduces the storage requirements which is desirable especially for wide area REM operation. Additionally, compact representation of the system reduces the overhead of information exchange among different REMs.

2.4.1.1. Inverse Distance Weighted Interpolation. Inverse distance weighted interpolation basically gives weights to the measurements according to the inverse of the distance to the point of interest. A general form of finding an interpolated value z at a given point p based on measurements $z_i = z(p_i)$ for $i = 1, 2, 3, ..., N_s$ using IDW-β is an interpolating function:

$$z(p) = \sum_{i=1}^{N_s} \frac{w_i(p) z_i}{\sum_{j=1}^{N_s} w_j(p)} \tag{2.13}$$

where N_s is the number of measurement points,

$$w_i(p) = \frac{1}{d(p, p_i)^\beta} \tag{2.14}$$

is a simple IDW weighting function of power β, and $d(.,.)$ is a distance function.

Here, the weight decreases as the distance from the interpolated points increases. Larger values of β assign more influence to values closest to the interpolated point. For $0 < \beta < 1$, $z(p)$ has smooth peaks over the interpolated points p_i, while the peaks become sharp as $\beta > 1$. The choice of the value of β is, therefore, a function

of the degree of smoothing desired in the interpolation, the density and distribution of samples being interpolated, and the maximum distance over which an individual sample is allowed to influence the surrounding points.

2.4.1.2. Kriging. Kriging is a group of geostatistical techniques that interpolate the value of a random field at an unobserved location from observations of its value at nearby locations. The theory behind interpolation and extrapolation by kriging was developed by the French mathematician Georges Matheron based on the Master's thesis of Daniel Gerhardus Krige [84], the pioneering plotter of distance-weighted average gold grades at the Witwatersrand reef complex in South Africa.

Kriging belongs to the family of linear least squares estimation algorithms. The aim of kriging is to estimate the value of an unknown real-valued function, f, at a point, x^*, given the values of the function at some other points, $x_1, \ldots, x_n$. A kriging estimator is said to be linear because the predicted value $\hat{f}(x^*)$ is a linear combination that may be written as

$$\hat{f}(x^*) = \sum_{i=1}^{n} \lambda_i(x^*) f(x_i). \tag{2.15}$$

The weights $\lambda_i(x^*)$ are solutions of a system of linear equations, which are obtained by assuming that f is a sample-path of a random process $F(x)$, and that the error of prediction

$$\varepsilon(x) = F(x) - \sum_{i=1}^{n} \lambda_i(x) F(x_i) \tag{2.16}$$

is to be minimized in some sense. For instance, the so-called simple kriging assumption is that the mean and the covariance of $F(x)$ is known and then, the kriging predictor is the one that minimizes the variance of the prediction error.

This method, originally used in mining exploration, has been applied in multitude of domains consisting of environmental science, meteorology, agriculture, and remote sensing. Central to geostatistics is the variogram, a function that models the variance between two points in space as a function of the distance between them. In the case of grid-sampled fields, the distance between measurements is a fixed lag distance. Randomized and optimized sampling schemes produce variable lag distances [85]. The weights used in linear combination depend on spatial correlation derived from *semivariogram* (half of the variogram) model of the measurement data. Semivariogram determines the correlation between any two points in the considered system based on their distance separation. In spatial statistics, the theoretical variogram $2\gamma(x, y)$ is a function describing the degree of spatial dependence of a spatial random field or stochastic process $Z(x)$. It is defined as the variance of the difference between field values at two locations across realizations of the field [86].

$$2\gamma(x, y) = \mathrm{var}(Z(x) - Z(y)) = E\left(|(Z(x) - \mu(x)) - (Z(y) - \mu(y))|^2\right). \tag{2.17}$$

If the spatial random field has constant mean μ, this is equivalent to the expectation for the squared increment of the values between locations x and y [87]:

$$2\gamma(x, y) = E\left(|Z(x) - Z(y)|^2\right), \tag{2.18}$$

where $\gamma(x, y)$ itself is called the semivariogram.

Kriging interpolates the value $Z(x_0)$ of a random field $Z(x)$ at an unobserved location x_0 from observations $z_i = Z(x_i),\ i = 1, \ldots, n$ of the random field at nearby locations $x_1, \ldots, x_n$. Kriging computes the best linear unbiased estimator $\hat{Z}(x_0)$ of $Z(x_0)$ based on a stochastic model of the spatial dependence quantified either by the variogram $\gamma(x, y)$ or by expectation $\mu(x) = E[Z(x)]$ and the covariance function $c(x, y)$ of the random field [88, 89].

Therefore, more formally, Kriging estimator is given by a linear combination

$$\hat{Z}(x_0) = \sum_{i=1}^{n} w_i(x_0) Z(x_i) \tag{2.19}$$

of the observed values $z_i = Z(x_i)$ with weights $w_i(x_0),\ i = 1, \ldots, n$ chosen such that the variance (also called Kriging variance or Kriging error):

$$\begin{aligned} \sigma_k^2(x_0) &\triangleq \operatorname{Var}\left(\hat{Z}(x_0) - Z(x_0)\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} w_i(x_0) w_j(x_0) c(x_i, x_j) + \operatorname{Var}\left(Z(x_0)\right) - 2\sum_{i=1}^{n} w_i(x_0) c(x_i, x_0) \end{aligned} \tag{2.20}$$

is minimized subject to the unbiasedness condition:

$$\mathrm{E}[\hat{Z}(x) - Z(x)] = \sum_{i=1}^{n} w_i(x_0)\mu(x_i) - \mu(x_0) = 0 \tag{2.21}$$

The Kriging variance must not be confused with the variance

$$\operatorname{Var}\left(\hat{Z}(x_0)\right) = \operatorname{Var}\left(\sum_{i=1}^{n} w_i Z(x_i)\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j c(x_i, x_j) \tag{2.22}$$

of the Kriging predictor $\hat{Z}(x_0)$ itself.

The interpolation by simple Kriging is given by:

$$\hat{Z}(x_0) = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}^T \begin{pmatrix} c(x_1, x_1) & \cdots & c(x_1, x_n) \\ \vdots & \ddots & \vdots \\ c(x_n, x_1) & \cdots & c(x_n, x_n) \end{pmatrix}^{-1} \begin{pmatrix} c(x_1, x_0) \\ \vdots \\ c(x_n, x_0) \end{pmatrix} \tag{2.23}$$

2.4.2. Transmitter Location Determination Based Methods

Aforementioned statistical methods directly approximate the signal strength without being concerned about the sources of the power. However, if information on PU locations are already available or can be approximated, they ease the process of REM construction. Transmitter location determination based methods use this approach and first focus on localizing the transmitter(s) and deriving their properties. Subsequently, they estimate the signal strength at each location by applying the propagation modeling. However, this approach has more degrees of freedom: multiple transmitters, transmitter properties such as antenna propagation pattern, and accurate characterization of the propagation environment.

After the potential transmitters are located, appropriate modeling of the signal strength calls for appropriate propagation modeling, e.g., channel gains between each pair of transmitter and receiver. In [81], Riihijärvi *et al.* study the effect of transmitter properties on the signal strength in a CRN using the second order statistics. In [82], authors applies an image processing based technique, which identifies the transmitters in the system and estimates their parameters based on Received Signal Strength (RSS) from sensors.

2.5. Optimization Techniques

In this section we give background information on the optimization techniques we use in our works. Mathematical models of optimization can be generally represented by a constraint set X and a cost function f that maps elements of X in to real numbers [90]. These kinds of models and problems are often used in communication systems to achieve a given objective. We mainly want to find an optimal decision (point in X with the minimum cost for minimization problems), i.e., an $x^* \in X$ such that

$$f(x^*) \leq f(x), \quad \forall x \in X. \tag{2.24}$$

2.5.1. Unconstrained Optimization

In this subsection, we consider unconstrained optimization problems and formulate them as presented in Equation 2.25.

$$\text{minimize} f(x) \quad \text{s.t. } x \in \mathbb{R}^n \tag{2.25}$$

For the most part, we assume that f is a continuously differentiable function, and we often also assume that f is twice continuously differentiable. The first and second derivatives of f play an important role in the characterization of optimal solutions via necessary and sufficient conditions. The first and second derivatives are also central in numerical algorithms for computing approximately optimal solutions.

A vector x^* is an *unconstrained local minimum* of f if it is no worse than its neighbors. More formally,

$$f(x^*) \leq f(x), \quad \forall x \text{ with } d(x, x^*) < \epsilon \tag{2.26}$$

where $d(x_1, x_2)$ stands for Euclidean norm.

A vector x^* is an *unconstrained global minimum* of f if it is no worse than all other vectors.

$$f(x^*) \leq f(x), \quad \forall x \in \mathbb{R}^n \tag{2.27}$$

Necessary conditions for optimality are

$$1. \quad \nabla f(x^*) = 0, \tag{2.28}$$

$$2. \quad \nabla^2 f(x^*) \text{ is positive semi-definite.} \tag{2.29}$$

Sufficient conditions for optimality are

$$1. \quad \nabla f(x^*) = 0, \tag{2.30}$$

$$2. \quad \nabla^2 f(x^*) \text{ is positive definite.} \tag{2.31}$$

Second condition is equivalent to the condition that Hessian is positive definite. Where positive definite and semi-definite stands for $n \times n$ matrices with $z^T M z > 0$ and $z^T M z \geq 0$ for all non-zero vectors z with real entries, respectively.

In many cases it is useful to know that there exists at least one global minimum of a function f over a set X. Generally, such a minimum need not exist. Given the range of values that $f(x)$ takes as x ranges over X, that is, the set of real numbers $\{f(x)|x \in X\}$, there are two possibilities:

(i) The set $\{f(x)|x \in X\}$ is bounded below; that is, there exists a scalar M such that $M < f(x)$ for all $x \in X$. In this case, the greatest lower bound of $\{f(x)|x \in X\}$ is a real number, which is denoted by $\inf_{x \in X} f(x)$.

(ii) The set $\{f(x)|x \in X\}$ is unbounded below. In this case $\inf_{x \in X} f(x) = -\infty$

Existence of at least one global minimum is guaranteed if f is a continuous function and X is a compact subset of $\mathbb{R}^n$ (Weierstrass theorem).

The most straightforward method to use optimality conditions to solve an optimization problem, is as follows: First, find all points satisfying the first order necessary condition $\nabla f(x) = 0$; then (if f is not known to be convex), check if $\nabla^2 f$ is positive

definite, in which case we are sure that they are strict local minima. Finding global minimum is more complicated, you have to know that it exists.

The conceptual framework of this subsection is fundamental in nonlinear programming and applies to constrained optimization methods as well.

Consider the problem of unconstrained minimization of a continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$. Most of the interesting algorithms for this problem rely on an important idea called iterative descent that works as follows: We start at some point x_0 (an initial guess) and successively generate vectors $x_1, x_2, \ldots$, such that f is decreased at each iteration, that is

$$f(x_{k+1}) < f(x_k), \qquad k = 0, 1, ..., \tag{2.32}$$

In doing so, we successively improve our current solution estimate and we hope to decrease f all the way to its minimum.

Choosing diminishing step size simplifies the algorithm. This step size rule does not guarantee descent at each iteration, although descent becomes more likely as the step size diminishes. One difficulty with a diminishing step size is that it may become so small that substantial progress cannot be maintained, even when far from a stationary point. For this reason, diminishing strategy must be chosen with care. Generally, the diminishing step size rule has good theoretical convergence properties. We use diminishing step size algorithms for optimizing thresholds of the uniform quantizer for detecting signals.

We deal with location estimation problems which are a kind of Least Squares problems of the form

$$\text{minimize } f(x) = \frac{1}{2}\|g(x)\|^2, \qquad \text{s.t. } x \in \mathbb{R}^n. \tag{2.33}$$

where g is a continuously differentiable function with component functions $g_i, ..., g_m$, where $g_i : \mathbb{R}^n \to \mathbb{R}^{r_i}$. Usually $r_i = 1$, but it is sometimes convenient to consider the more general case. Least squares problems are very common in practice. A principal case arises when g consists of n scalar-valued functions and we want to solve the system of n equations with n unknowns $g(x) = 0$. We can formulate this as the least squares optimization problem.

Least squares problem is simply an optimization problem with no constraints and an objective which is a sum of squares of terms of the form $a_i^T x - b_i$ that can be written as

$$\text{minimize } f(x) = \|Ax - b\|^2 \quad \text{s.t. } x \in \mathbb{R}^n \tag{2.34}$$

that is known as matrix form.

The solution of a least squares problem in Equation 2.34 can be reduced to solving a set of linear equations,

$$(A^T A)x = A^T b, \tag{2.35}$$

so we have the analytical solution $x = (A^T A)^{-1} A^T b$. For least squares problems we have good algorithms for solving the problem to high accuracy, with very high reliability. The least squares problem can be solved in a time approximately proportional to $n^2 k$, with a known constant where k is the number of equations (number of rows of A).

2.5.2. Constrained Optimization

The design of the communication systems in order to achieve a given objective subject to various constraints is an essential task that appears often. Consider the

optimization problem in the form

$$\text{minimize } f(x) \quad \text{s.t. } f_i(x) = 0; i = 1, ..., m \quad x \in X. \tag{2.36}$$

where $x = [x_1, ..., x_n]$ is the optimization variables, $f(.)$ and $f_i(.)$ are the objective function and the equality constraints functions, respectively. x is called feasible point if it satisfies the constraints. The set of all feasible points is called feasible set and denoted by $\mathcal{X}$.

If the objective and constraints functions are all linear, the problem is called a Linear Programming (LP) problem and the global optimal point is easy to be found. Simplex algorithm is one of the most popular LP algorithms. Since the LP problem having a solution must have an optimal value that falls on the boundary of the feasible region, the algorithm starts with a given initial solution and moves to the neighboring vertex that best improve the objective function value. These movements are performed until obtaining the optimal point [90].

When the optimization problem is convex, the global optimal solution is equal to local optimal point. LP problem is a special kind of the convex optimization problem. Different methods can be used to find the global optimal point. For the unconstrained convex problem, gradient and Newtons method are the known ones. Gradient method moves from an initial feasible point towards the optimal value by updating iteratively the current optimization variables values in the direction of the gradient. Although the gradient method is simple and it guarantees locating the optimal point (if exists), it has relatively slow convergence [91].

In a constrained convex optimization problems, projected gradient algorithm, interior point method, and ellipsoid method can be applied. In projected gradient algorithm, the search direction is projected into the subspace tangent to the active constraints. Ellipsoid method generates a sequence of ellipses inside the feasible set whose volumes decreases at each iteration to enclose the maximum of the convex func-

tion. Ellipsoid method is used in low-dimensional problems due its poor performance in large ones. Interior point method is a search algorithm that adds a penalty to the objective function when the search point approaches the boundary of the feasible set.

Lagrange dual problem for constrained problem is a method to solve these kind of problems. Consider the problem defined in Equation 2.36, the basic idea in Lagrangian duality is to take the constraints into account by augmenting the objective function with a weighted sum of the constraint functions. The Lagrangian is defined as

$$\mathcal{L}(x, \lambda) = f(x) - \sum_{i=1}^{m} \lambda_i f_i(x) \tag{2.37}$$

where $\lambda = [\lambda_1, ..., \lambda_m]$ is the Lagrange multipliers vector.

Lagrange dual function is defined as the minimum value of the Lagrangian:

$$g(\lambda) = \inf_{x \in X} \mathcal{L}(x, \lambda) = \inf_{x \in X} (f(x) - \sum_{i=1}^{m} \lambda_i f_i(x)) \tag{2.38}$$

When the Lagrangian is unbounded below in x, the dual function takes on the value $-\infty$. Since the dual function is the pointwise infimum of a family of affine functions of λ, it is concave, even when the problem is not convex.

The dual function yields lower bounds on the optimal value p^* of the problem in Equation 2.36. For any $\lambda_i \geq 0$ we have $g(\lambda) \leq p^*$. Thus we have a lower bound that depends on some parameter λ. A natural question is: What is the best lower bound that can be obtained from the Lagrange dual function? This leads to the optimization problem

$$\text{maximize } g(\lambda) \quad \text{s.t. } \lambda \geq 0. \tag{2.39}$$

This problem is called the Lagrange dual problem associated with the problem stated in Equation 2.36. The Lagrange dual problem in Equation 2.39 is a convex optimization

problem, since the objective to be maximized is concave and the constraint is convex. This is the case whether or not the primal problem in Equation 2.36 is convex.

Based on the Lagrangian function(we assume that the functions f, $f_1, ..., f_m$ are differentiable, but we make no assumptions yet about convexity), the following necessary and sufficient conditions are formed to find the global maximum of the problem in Equation 2.36 as follows:

- *Karush-Kuhn-Tucker (KKT) necessary condition:* Let x^* and λ^* be any primal and dual optimal points with zero duality gap. Since x^* minimizes $\mathcal{L}(x, \lambda^*)$ over x, it follows that its gradient must vanish at x^*, i.e., there exists unique Lagrange multiplier vector $\lambda^* = [\lambda_1^*, ..., \lambda_m^*]$ such that

$$\begin{aligned} \frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_i} &= 0, \quad i = 1, ..., n \\ \lambda_i^* &\geq 0, \quad i = 1, ..., m \\ \lambda_i^* f_i(x^*) &= 0, \quad i = 1, ..., m \end{aligned} \tag{2.40}$$

  Note that the necessary condition means that if a given point satisfies the KKT conditions, it might not be a local minimum of the problem. To summarize, for any optimization problem with differentiable objective and constraint functions for which strong duality obtains, any pair of primal and dual optimal points must satisfy the KKT conditions.
- *General sufficient condition:* if x^* is a feasible point together with the Lagrange multipliers vector λ^* satisfies $\lambda_i^* f_i(x^*) = 0$, $i = 1, ..., m$ and maximizes the Lagrangian function $\mathcal{L}(x, \lambda^*)$ over $x \in \mathcal{X}$, i.e., $x^* = \arg\max_{x \in X} \mathcal{L}(x, \lambda^*)$, then x^* is a global maximum of the optimization problem.

If $f(.)$ and $f_i(.)$ are convex functions, the Lagrangian function is convex function as well and the necessary conditions become also sufficient. Therefore, the global

maximum x^* can be found by solving the system of equations formed by

$$\frac{\partial \mathcal{L}(x^*, \lambda^*)}{\partial x_i} = 0, \quad i = 1, ..., n \tag{2.41}$$

2.5.3. Genetic Algorithm

In the computer science field of artificial intelligence, a Genetic Algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms, which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover.

In a genetic algorithm, a population of strings (called chromosomes), which encode candidate solutions (called individuals) to an optimization problem, evolves toward better solutions. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.

A typical genetic algorithm requires a genetic representation of the solution domain, and a fitness function to evaluate the individual. The fitness function is defined over the genetic representation and measures the quality of the represented solution. The fitness function is always problem dependent. Once the genetic representation and the fitness function are defined, a GA proceeds to initialize a population of solutions (usually randomly) and then to improve it through repetitive application of the

mutation, crossover, inversion and selection operators.

During each successive generation, a proportion of the existing population is selected to breed a new generation. Individual solutions are selected through a fitness-based process, where fitter solutions (as measured by a fitness function) are typically more likely to be selected. Certain selection methods rate the fitness of each solution and preferentially select the best solutions.

The next step is to generate a second generation population of solutions from those selected through genetic operators: crossover (also called recombination), and/or mutation.

For each new solution to be produced, a pair of "parent" solutions is selected for breeding from the pool selected previously. By producing a "child" solution using the above methods of crossover and mutation, a new solution is created which typically shares many of the characteristics of its "parents". New parents are selected for each new child, and the process continues until a new population of solutions of appropriate size is generated [92].

Although Crossover and Mutation are known as the main genetic operators, it is possible to use other operators such as regrouping, colonization-extinction, or migration in genetic algorithms [92].

3. COOPERATIVE SENSING VIA UNIFORM QUANTIZATION

3.1. Incentives and Contributions

The sharing of local observations between the secondary users and the fusion center (that fuses the local decisions of cooperating nodes) is the most crucial factor that determines the performance of cooperative sensing. Detection performance is determined by the quality of local observations and the quality of the information received by the fusion center. Therefore, the number of quantization bins, the number of bits sent for sensing reports, and the global decision logic affect the system performance. Furthermore, the imperfections in the reporting channel and the erroneous reports due to malfunctioning or malicious secondary devices should also be considered. Distorting impacts of these factors on the global decision logic must be analyzed for optimizing the sensing performance.

There is a intricate interplay among the period and size of the sensing reports and bandwidth resources. Decreasing the number of bits for sensing reports with acceptable performance enables increasing the number of sensing periods and better performance. Furthermore, having a bandwidth-limited reporting channel does not allow sending the whole observation and using complicated protocols for sending the sensing reports to fusion center. Hence, the nodes should quantize their observations in an optimal manner rather than sending the exact observation values. Using more quantization bins increases the quality of the information sent at the cost of limited congestion in the reporting channel.

In [44], cooperative sensing and quantization schemes (1-bit or hard decision) are investigated for multiple primary bands. We also focus on quantization, but consider a more generalized version of quantization which enables multiple quantization bins. The effects of imperfect reporting channel to sensing performance is analyzed for hard

decision logic functions in [42]. We extend this study for functional global decision techniques. Spectrum sensing scheduling and sensing time are analyzed in [45, 46]. In [47], authors analyze the cooperative sensing under bandwidth constraints. In [48], the authors examine the optimal quantizer for signal detection locally. We also focus on signal detection via quantization, however, we consider the overall process rather than just focusing on the quantization part. Similarly in [25], detection and false alarm probabilities are derived with consideration of errors in the reporting channel due to fading. Only the hard decision logic M-out-of-N is studied in the paper. In [52], BEP wall is introduced and performance analysis for M-out-of-N is done. Authors analyze the SNR loss due to BEP without considering the overall detection probabilities.

In this thesis, a novel cooperative sensing scheme is proposed, which uses uniform quantization named *Uniform Quantization-based Cooperative Sensing (UniQCS)*, and optimizes the parameters of the proposed method to maximize the probability of detection while using a uniform quantizer. Most of the works on quantization focus on just the local quantization process. On the contrary, we optimize the overall performance in terms of cooperative detection. The sensing performance of UniQCS is close to EGC (which provides an upper bound for the schemes that do not differentiate the nodes) and better than conventional hard decision algorithms at the cost of a marginal increase in overhead bits. Furthermore, the robustness of UniQCS is verified under an imperfect reporting channel and false reports.

Contributions of UniQCS can be listed as:

(i) A cooperative sensing method that uses uniform quantizer is proposed.
(ii) Global decision techniques for systems that use quantization are generalized.
(iii) Probability of detection and false alarm formulations are derived for the perfect reporting channel and when no false reports exist.
(iv) Probability of detection and false alarm formulations are derived for an imperfect reporting channel and when no false reports exist.
(v) Probability of detection and false alarm formulations are derived for the perfect

reporting channel and when false reports exist.

(vi) Optimization problem that maximizes the probability of cooperative detection is solved.

(vii) Robustness of UniQCS to false reports and an imperfect reporting channel is analyzed.

(viii) Detection performance of UniQCS is close to the EGC (which provides an upper bound for the schemes that do not differentiate the nodes).

3.2. Uniform Quantizer for Cooperative Sensing (UniQCS)

The conventional ED is a quantizer with two bins that uses single threshold to determine the presence or absence of the signal. By using more quantization bins at the sensing nodes, the information gain accrued by cooperation can be increased. Thresholds divide the observation space into bins and the sensing node determines the bin into which the observation falls.

We propose to quantize the observed energy O_i by secondary user i locally and collect such information from all secondary users at the fusion center and give a global decision in an optimal manner. Most of the works on quantization focus on just the local quantization process. We do not just optimize the local quantization process; instead we optimize the overall performance in terms of cooperative detection. Therefore, the proposed method has two parts: local quantization and global decision logic.

3.2.1. Local Quantization

The proposed local quantization supports variable number of quantization bins. In Figure 3.1, quantization levels with four bins are depicted as an example. There are three thresholds that are determined by the first threshold, λ_1, since the distance between consecutive thresholds is fixed and denoted by Δ (i.e., uniform quantizer). The observation of node i, O_i, is greater than or equal to zero, and the observation space is divided into bins where B_k denotes the kth bin.

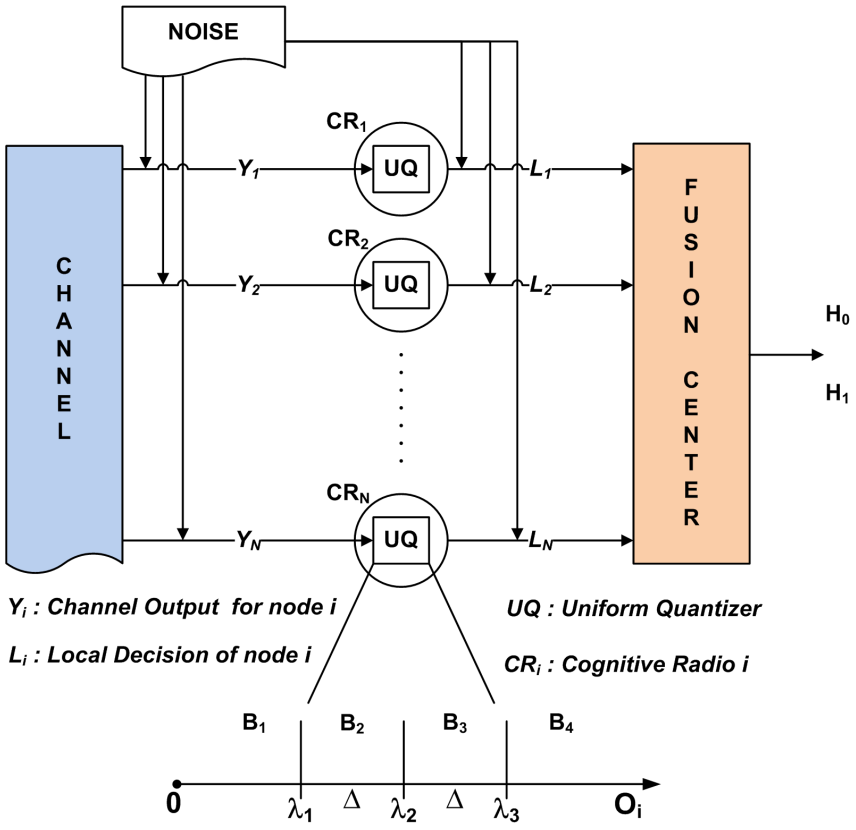


Figure 3.1. System model of cooperative methods and local quantization.

For a given λ_1, other thresholds can be found by adding Δ for the next threshold at each step. Hence, the degree of freedom is just one for a given Δ. For a given set of thresholds $\{\lambda_1, \lambda_2, ..., \lambda_{n-1}\}$, we can evaluate the probability of having observation O_i in B_k under H_0 and H_1, respectively. $\mathbf{P}^A_{H_i}(B_k)$ denotes the probability of having local observation in bin B_k under hypothesis H_i and AWGN channel

$$\mathbf{P}^A_{H_0}(B_k) = \begin{cases} 1 - G_{TW}(\lambda_k) & \text{if } k = 1 \\ G_{TW}(\lambda_{k-1}) & \text{if } k = n \\ G_{TW}(\lambda_{k-1}) - G_{TW}(\lambda_k) & \text{otherwise} \end{cases} \tag{3.1}$$

$$\mathbf{P}^A_{H_1}(B_k) = \begin{cases} 1 - Q_{TW}(\gamma,\, \lambda_k) & \text{if } k = 1 \\ Q_{TW}(\gamma,\, \lambda_{k-1}) & \text{if } k = n \\ Q_{TW}(\gamma,\, \lambda_{k-1}) - Q_{TW}(\gamma,\, \lambda_k) & \text{otherwise} \end{cases} \tag{3.2}$$

where $G_{TW}(\lambda)$ denotes $\frac{\mathbf{\Gamma}(TW,\lambda/2)}{\mathbf{\Gamma}(TW)}$ and $Q_{TW}(\gamma,\ \lambda)$ denotes $\mathbf{Q}_{TW}(\sqrt{2\gamma},\ \sqrt{\lambda})$.

Under a Rayleigh channel and conditioned on H_0, there is no difference with the AWGN case since there is no transmission. For evaluating bin probabilities conditioned on H_1, Equation 2.7 of $\mathbf{P}_{\mathbf{D}}^{R}$ is used. A similar bin probability formula can be easily obtained for a Rayleigh channel by replacing $Q_{TW}(\gamma,\ \lambda)$ in Equation 3.2 with the $\mathbf{P}_{\mathbf{D}}^{R}$ formulation.

$$\mathbf{P}_{H_1}^{R}(B_k) = \begin{cases} 1 - \mathbf{P}_{\mathbf{D}}^{R}(\lambda_k) & \text{if } k = 1 \\ \mathbf{P}_{\mathbf{D}}^{R}(\lambda_{k-1}) & \text{if } k = n \\ \mathbf{P}_{\mathbf{D}}^{R}(\lambda_{k-1}) - \mathbf{P}_{\mathbf{D}}^{R}(\lambda_k) & \text{otherwise} \end{cases} \tag{3.3}$$

3.2.2. Global Decision Function

To take advantage of the spatial diversity in the wireless channel, cooperative spectrum sensing methods have been proposed in [26–28]. We provide a cooperative sensing scheme for similar reasons.

The system model for the proposed method is depicted in Figure 3.1. Each cooperating secondary user senses the spectrum and sends its "quantized" local measurement as L_i, (index of the quantization bin) to the fusion center at the cognitive base station. The fusion center makes a global decision according to L_i values.

The global decision logic schemes in the literature can be classified into two main categories: soft decision and hard decision logic. In hard decision methods, the fusion center collects local decisions consisting of 1-bit information. In soft decision methods, the exact measurements are reported to the fusion center. Well known soft decision methods are EGC and Maximal Gain Combiner (MGC). EGC uses fixed weights for measurements reported to the fusion center. All received measurements are summed coherently and compared against one global threshold. In MGC, the received measurements are weighted with respect to their SNR values, and then summed and compared

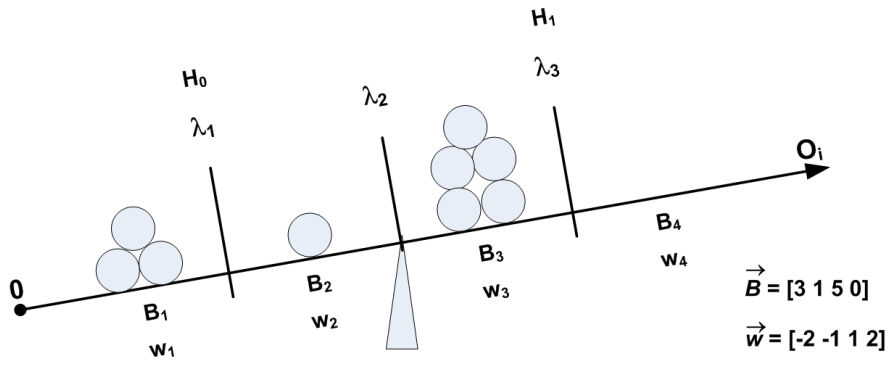


Figure 3.2. Seesaw analogy of the global decision logic.

with one predetermined global threshold.

If the quantization is done using basically two bins, hard decision logic functions, such as Or and Majority Logic, can be applied as the global decision logic at the fusion center. Considering 0 and 1 for L_i trivially suggests hard decision logic for cooperation. However, in the case of more quantization bins none of the hard decision logic functions can be used. In such a situation, the global decision logic must have a functional form.

We propose the global decision technique, *Cooperative Sensing with Decision Vector (CSDV)* [93] which is inspired by the analogy of a seesaw. The example in Figure 3.2 contains three measurements in B_1, one measurement in B_2, and five measurements in B_3, hence the final decision is clearly H_0 with the given weights, even though the majority of the nodes are at the right side.

For the given 4-bin example in Figure 3.2, the fusion center receives the quantized measurements and counts the number of users in each quantization bin. Having three reports in B_1, one report in B_2, and five reports in B_3 implies $\overrightarrow{B} = [3\ 1\ 5\ 0]$. The decision function, $\delta_{\overrightarrow{w}}(.)$, is evaluated with the help of the weights and the number of users in the bins as

$$\delta_{\overrightarrow{w}}(\overrightarrow{B}) = \begin{cases} 1 & \text{if } \overrightarrow{B} \cdot \overrightarrow{w} > 0 \\ 0 & \text{otherwise} \end{cases} . \tag{3.4}$$

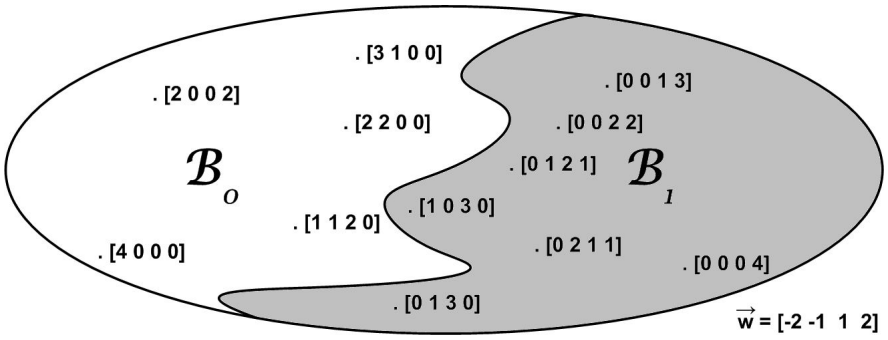


Figure 3.3. Set of local decisions for four nodes and quantization bins.

In the proposed method, the global decision depends on the threshold values and the weight vector. The reader should note that the weights are assigned to the quantization bins, not to the reporting nodes.

Having a weight vector enables us to tune the system performance and the system behavior. Clearly, $\overrightarrow{w}$ partitions $\mathcal{B}$ (the set of all possible $\overrightarrow{B}$ vectors as depicted in Figure 3.3) into two disjoint sets as $\mathcal{B}_0$ and $\mathcal{B}_1$, where the final decision is H_0 and H_1, respectively. $\delta_{\overrightarrow{w}}(.)$ evaluates to 0 in $\mathcal{B}_0$, and its corresponding terms in the probability evaluations become 0. Both $\mathbf{P_D}$ and $\mathbf{P_F}$ are evaluated by summing the probabilities of observing the cases in $\mathcal{B}_1$ since $\mathbf{P_D}$ and $\mathbf{P_F}$ are the probabilities of deciding H_1 as the final decision conditioned on having primary communication and no primary communication, respectively.

$$\mathbf{P^{ch}_{F,CSDV}} = \sum_{\overrightarrow{B}\in\mathcal{B}_1} (\mathbf{P}^{ch}_{H_0}(\overrightarrow{B})) \tag{3.5}$$

$$\mathbf{P^{ch}_{D,CSDV}} = \sum_{\overrightarrow{B}\in\mathcal{B}_1} (\mathbf{P}^{ch}_{H_1}(\overrightarrow{B})) \tag{3.6}$$

Example 3.2.1. Let's consider four nodes that are sensing the spectrum via ED with four quantization bins. We present the situations that evaluate to H_1 in cooperative sensing via the proposed global decision logic with $\overrightarrow{w} = [-2,\ -1,\ 1,\ 2]$ in Table 3.1. The table also includes the probability of occurrence of each situation.

Table 3.1. Example $\mathcal{B}_1$ elements with four nodes and quantization bins and their probabilities under H_0 and H_1.

$\overrightarrow{B} \in \mathcal{B}_1$	Conditioned on H_0	Conditioned on H_1
$[0,0,0,4]$	$\mathbf{P}_{H_0}^{ch}(B_4)^4$	$\mathbf{P}_{H_1}^{ch}(B_4)^4$
$[0,0,1,3]$	$\binom{4}{3}\mathbf{P}_{H_0}^{ch}(B_4)^3\mathbf{P}_{H_0}^{ch}(B_3)$	$\binom{4}{3}\mathbf{P}_{H_1}^{ch}(B_4)^3\mathbf{P}_{H_1}^{ch}(B_3)$
$[0,1,0,3]$	$\binom{4}{3}\mathbf{P}_{H_0}^{ch}(B_4)^3\mathbf{P}_{H_0}^{ch}(B_2)$	$\binom{4}{3}\mathbf{P}_{H_1}^{ch}(B_4)^3\mathbf{P}_{H_1}^{ch}(B_2)$
$[1,0,0,3]$	$\binom{4}{3}\mathbf{P}_{H_0}^{ch}(B_4)^3\mathbf{P}_{H_0}^{ch}(B_1)$	$\binom{4}{3}\mathbf{P}_{H_1}^{ch}(B_4)^3\mathbf{P}_{H_1}^{ch}(B_1)$
$[0,0,2,2]$	$\binom{4}{2}\mathbf{P}_{H_0}^{ch}(B_4)^2\mathbf{P}_{H_0}^{ch}(B_3)^2$	$\binom{4}{2}\mathbf{P}_{H_1}^{ch}(B_4)^2\mathbf{P}_{H_1}^{ch}(B_3)^2$
$[0,2,0,2]$	$\binom{4}{2}\mathbf{P}_{H_0}^{ch}(B_4)^2\mathbf{P}_{H_0}^{ch}(B_2)^2$	$\binom{4}{2}\mathbf{P}_{H_1}^{ch}(B_4)^2\mathbf{P}_{H_1}^{ch}(B_2)^2$
$[0,1,1,2]$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_0}^{ch}(B_4)^2\mathbf{P}_{H_0}^{ch}(B_3)\mathbf{P}_{H_0}^{ch}(B_2)$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_1}^{ch}(B_4)^2\mathbf{P}_{H_1}^{ch}(B_3)\mathbf{P}_{H_1}^{ch}(B_2)$
$[1,0,1,2]$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_0}^{ch}(B_4)^2\mathbf{P}_{H_0}^{ch}(B_3)\mathbf{P}_{H_0}^{ch}(B_1)$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_1}^{ch}(B_4)^2\mathbf{P}_{H_1}^{ch}(B_3)\mathbf{P}_{H_1}^{ch}(B_1)$
$[1,1,0,2]$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_0}^{ch}(B_4)^2\mathbf{P}_{H_0}^{ch}(B_2)\mathbf{P}_{H_0}^{ch}(B_1)$	$\binom{4}{2}\binom{2}{1}\mathbf{P}_{H_1}^{ch}(B_4)^2\mathbf{P}_{H_1}^{ch}(B_2)\mathbf{P}_{H_1}^{ch}(B_1)$
$[0,0,3,1]$	$\binom{4}{1}\mathbf{P}_{H_0}^{ch}(B_4)\mathbf{P}_{H_0}^{ch}(B_3)^3$	$\binom{4}{1}\mathbf{P}_{H_1}^{ch}(B_4)\mathbf{P}_{H_1}^{ch}(B_3)^3$
$[0,1,2,1]$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_0}^{ch}(B_4)\mathbf{P}_{H_0}^{ch}(B_3)^2\mathbf{P}_{H_0}^{ch}(B_2)$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_1}^{ch}(B_4)\mathbf{P}_{H_1}^{ch}(B_3)^2\mathbf{P}_{H_1}^{ch}(B_2)$
$[1,0,2,1]$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_0}^{ch}(B_4)\mathbf{P}_{H_0}^{ch}(B_3)^2\mathbf{P}_{H_0}^{ch}(B_1)$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_1}^{ch}(B_4)\mathbf{P}_{H_1}^{ch}(B_3)^2\mathbf{P}_{H_1}^{ch}(B_1)$
$[0,2,1,1]$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_0}^{ch}(B_4)\mathbf{P}_{H_0}^{ch}(B_3)\mathbf{P}_{H_0}^{ch}(B_1)^2$	$\binom{4}{1}\binom{3}{2}\mathbf{P}_{H_1}^{ch}(B_4)\mathbf{P}_{H_1}^{ch}(B_3)\mathbf{P}_{H_1}^{ch}(B_1)^2$
$[0,0,4,0]$	$\mathbf{P}_{H_0}^{ch}(B_3)^4$	$\mathbf{P}_{H_1}^{ch}(B_3)^4$
$[0,1,3,0]$	$\binom{4}{3}\mathbf{P}_{H_0}^{ch}(B_3)^3\mathbf{P}_{H_0}^{ch}(B_2)$	$\binom{4}{3}\mathbf{P}_{H_1}^{ch}(B_3)^3\mathbf{P}_{H_1}^{ch}(B_2)$
$[1,0,3,0]$	$\binom{4}{3}\mathbf{P}_{H_0}^{ch}(B_3)^3\mathbf{P}_{H_0}^{ch}(B_1)$	$\binom{4}{3}\mathbf{P}_{H_1}^{ch}(B_3)^3\mathbf{P}_{H_1}^{ch}(B_1)$

There are 16 situations that result in H_1 for a global decision with the given $\overrightarrow{w}$. Probabilities are presented in the second and third column of the table. Probabilities of observing the $[0,0,0,4]$ outcome at the fusion center after nodes report their local observations under H_0 and H_1 condition are $\mathbf{P}_{H_0}^{ch}(B_4)^4$ and $\mathbf{P}_{H_1}^{ch}(B_4)^4$, respectively. These probabilities are summed up to evaluate the $\mathbf{P_D}$ and $\mathbf{P_F}$ for the global decision logic. $\mathbf{P_D}$ is evaluated via summing the probabilities conditioned on H_1 since it represents the probability of deciding H_1 globally when the true state is H_1. Similarly, $\mathbf{P_F}$ is evaluated via summing the probabilities conditioned on H_0 since it represents the probability of deciding H_1 globally when the true state is H_0.

3.2.3. Formulation of P_D and P_F

We mainly focus on the Rayleigh channel since it includes multipath effects as mentioned in Section 2.2, and using CSDV method as the fusion strategy leads to the following probabilities of cooperative detection and false alarm under a Rayleigh channel:

$$\begin{aligned}\mathbf{P^R_{F,CSDV}} &= \sum_{\overrightarrow{B}\in\mathcal{B}_1} \mathbf{P}\{\text{Collected local decisions at fusion center} = \overrightarrow{B}|H_0\} \\ &= \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}^R_{H_0}(B_k))^{N_{B_k}} \end{aligned} \tag{3.7}$$

$$\begin{aligned}\mathbf{P^R_{D,CSDV}} &= \sum_{\overrightarrow{B}\in\mathcal{B}_1} \mathbf{P}\{\text{Collected local decisions at fusion center} = \overrightarrow{B}|H_1\} \\ &= \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}^R_{H_1}(B_k))^{N_{B_k}} \end{aligned} \tag{3.8}$$

where N is the number of cooperative users, N_{B_k} is the number of users having observation in bin B_k, and n is the number of quantization bins.

3.2.4. Formulation of P_D and P_F with False Reports

When one of the CR nodes, intentionally or not, reports the presence of a primary user by reporting H_1 erroneously all the time, Or logic always decides the presence of a primary user. In this case, $\mathbf{P}^{ch}_{\mathbf{F},OR} = 1$ under any given channel condition and it is not possible to utilize the white spaces at all.

Using the CSDV method as the fusion strategy with fr false reports leads to the following probabilities of cooperative detection and false alarm under a Rayleigh

channel:

$$\mathbf{P_{F,CSDV}^{R}} = \sum_{\overrightarrow{B}\in\mathcal{B}^{-fr}} \left\{ \delta_{\overrightarrow{w}}(\overrightarrow{B}^{+fr}) \prod_{k=1}^{n} \binom{N\text{-}fr\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}_{H_0}^{R}(B_k))^{N_{B_k}} \right\} \tag{3.9}$$

$$\mathbf{P_{D,CSDV}^{R}} = \sum_{\overrightarrow{B}\in\mathcal{B}^{-fr}} \left\{ \delta_{\overrightarrow{w}}(\overrightarrow{B}^{+fr}) \prod_{k=1}^{n} \binom{N\text{-}fr\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}_{H_1}^{R}(B_k))^{N_{B_k}} \right\} \tag{3.10}$$

where N is the number of cooperative users, N_{B_k} is the number of users having observation in bin B_k without considering fr nodes, n is the number of quantization bins, $\mathcal{B}^{-fr}$ represents all combinations of N-fr users distributed in the quantization bins, and $\overrightarrow{B}^{+fr} = \overrightarrow{B} + [0\ 0\ 0\ fr]$ for the 3-threshold (4-bin) case. We do not have any assumptions on the false reporting nodes; the formulation depends only on the number of false reporting nodes. Increasing the number of false reporting nodes degrades the system performance.

3.2.5. Formulation of P_D and P_F in Imperfect Reporting Channel

When the reporting channel is imperfect, errors may occur on the reported local decisions. In case of 1-bit quantization, L_i can only be 0 or 1 (corresponding to B_1 or B_2). Thus, it can be modeled as a *Binary Symmetric Channel (BSC)* with cross-over probability p_e, which is equal to the bit error rate of the channel.

In the 1-bit quantization (hard decision) case, under H_1, the fusion center can receive 1 from a CR node in two situations: In the first case, CR node decides H_1 (B_2) locally and transmits it to the fusion center without any errors, which occurs with probability $\mathbf{P}_{H_1}^{ch}(B_2)(1-p_e)$. The second situation occurs when CR node decides H_0 (B_1) locally and transmits it to the fusion center with error, which occurs with probability $\mathbf{P}_{H_1}^{ch}(B_1)p_e$. Therefore, the detection probability for a single node with an

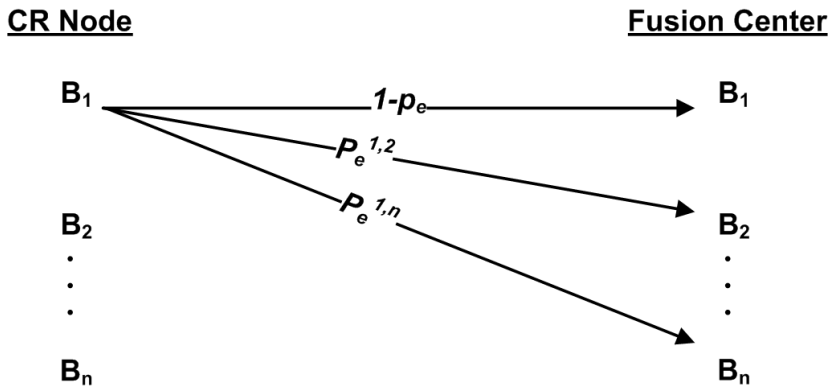


Figure 3.4. Report channel model of n-bin quantization system.

imperfect reporting channel becomes

$$\mathbf{P}_{H_1}^{ch,e}(B_2) = \mathbf{P}_{H_1}^{ch}(B_2)(1-p_e) + \mathbf{P}_{H_1}^{ch}(B_1)p_e. \tag{3.11}$$

Using similar reasoning and formulation, the local false alarm probability with an imperfect reporting channel is given as

$$\mathbf{P}_{H_0}^{ch,e}(B_2) = \mathbf{P}_{H_0}^{ch}(B_2)(1-p_e) + \mathbf{P}_{H_0}^{ch}(B_1)p_e. \tag{3.12}$$

In case of more quantization bins, the problem gets complicated. An example system model for n-bin quantization is depicted in Figure 3.4. Only the arrows originating from B_1 are shown to make the figure more readable. The sum of error cases is denoted by p_e while $p_e^{i,j}$ represents the probability of receiving B_j at the fusion center when B_i is observed by a CR node and sent over the reporting channel. Here, the symbol errors occur and the symbol error rate for each symbol pair may be different according to the coding used.

Generalizing the formulation in Equation 3.11 and 3.12, we denote the probability of receiving B_i at the fusion center through the imperfect reporting channel as $\mathbf{P}_{H_1}^{ch,e}(B_i)$

and $\mathbf{P}_{H_0}^{ch,e}(B_i)$ under H_1 and H_0, respectively, and evaluate

$$\mathbf{P}_{H_1}^{ch,e}(B_i) = \mathbf{P}_{H_1}^{ch}(B_i)(1-p_e) + \sum_{k\neq i} \mathbf{P}_{H_1}^{ch}(B_k) p_e^{k,i}, \tag{3.13}$$

$$\mathbf{P}_{H_0}^{ch,e}(B_i) = \mathbf{P}_{H_0}^{ch}(B_i)(1-p_e) + \sum_{k\neq i} \mathbf{P}_{H_0}^{ch}(B_k) p_e^{k,i}. \tag{3.14}$$

After deriving these equations, we plug $\mathbf{P}_{H_0}^{ch,e}(B_i)$ and $\mathbf{P}_{H_1}^{ch,e}(B_i)$ into $\mathbf{P_F}$ and $\mathbf{P_D}$ formulation in Equation 3.5 and 3.6, respectively. We evaluate $\mathbf{P_F}$ and $\mathbf{P_D}$ for the CSDV under a Rayleigh and imperfect reporting channel

$$\mathbf{P}_{\mathbf{F,CSDV}}^{\mathbf{R}} = \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}_{H_0}^{R,e}(B_k))^{N_{B_k}} \tag{3.15}$$

$$\mathbf{P}_{\mathbf{D,CSDV}}^{\mathbf{R}} = \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}_{H_1}^{R,e}(B_k))^{N_{B_k}} \tag{3.16}$$

where $\mathcal{B}$ is the set of all possible $\overrightarrow{B}$s, N is the number of cooperative users, N_{B_k} is the number of reports in bin B_k from the fusion center's perspective (with errors), $\mathbf{P}_{H_i}^{R,e}(B_k)$ is the probability of having a report in bin B_k under H_i in a Rayleigh channel, and n is the number of quantization bins.

3.3. Problem Formulation

Recall that the main goal of spectrum sensing is to decide between two hypotheses that the channel state is empty (H_0) or it is actively used (H_1). If the conditional distributions of O_i are known, the most appropriate optimality criterion for the decision is Neyman-Pearson optimality that maximizes $\mathbf{P}_{\mathbf{D,CSDV}}^{\mathbf{ch}}$ subject to the $\mathbf{P}_{\mathbf{F,CSDV}}^{\mathbf{ch}} < \alpha$

constraint.

Proposed quantization and global decision logic need to optimize λ_1, Δ, and $\overrightarrow{w}$. We portray two problem definitions and the solution methodologies. The first problem aims to optimize the thresholds of the quantizer for a given $\overrightarrow{w}$ and α. The second problem aims to optimize the $\overrightarrow{w}$, however $\mathbf{P^{ch}_{D,CSDV}}$ is not differentiable with respect to $\overrightarrow{w}$ and also it is not 1-1 in terms of $\overrightarrow{w}$. Therefore, we embed a genetic algorithm into the optimization process.

3.3.1. Threshold Optimization

The corresponding optimization problem for optimizing λ_1 and Δ under a Rayleigh channel when there are no false reports and the perfect reporting channel is formulated in Equation 3.17 when $\overrightarrow{w}$ is given. When there are false reports or the reporting channel is imperfect, we can obtain similar optimization problem formulations by using Equation 3.9, 3.10 and 3.15, 3.16, respectively.

$$
\begin{aligned}
UniQCS1: \quad & \underset{\Delta,\lambda_1}{maximize} \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}^R_{H_1}(B_k))^{N_{B_k}} \\
& subject\ to \sum_{\overrightarrow{B}\in\mathcal{B}} \delta_{\overrightarrow{w}}(\overrightarrow{B}) \prod_{k=1}^{n} \binom{N\text{-}\sum_{j=1}^{k-1} N_{B_j}}{N_{B_k}} (\mathbf{P}^R_{H_0}(B_k))^{N_{B_k}} \leq \alpha \qquad (3.17)
\end{aligned}
$$

Different (Δ, λ_1) values change the quantization bins and accordingly $(\mathbf{P}^R_{H_1}(B_k))^{N_{B_k}}$ and $(\mathbf{P}^R_{H_0}(B_k))^{N_{B_k}}$. If $\lambda_1 = 0$ and $\Delta = 0$, then we have $\mathbf{P_F} = 1 = \mathbf{P_D}$. For a fixed Δ, shifting λ_1 and consequently the other thresholds to higher values results in a decrease in $\mathbf{P_F}$ and $\mathbf{P_D}$. Similarly, for a fixed λ_1, shifting Δ to higher values results in the same behavior. This behavior with the principle of Neyman-Pearson that says that the maximum $\mathbf{P^{ch}_{D,CSDV}}$ is achieved when $\mathbf{P^{ch}_{F,CSDV}} = \alpha$ are the key points for solving the maximization problem. We seek for λ_1^* and Δ^* satisfying $\mathbf{P_F} = \alpha$, which gives

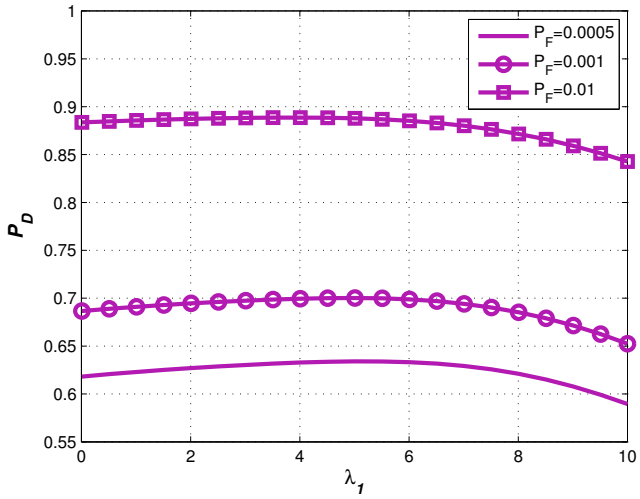


Figure 3.5. $\mathbf{P_D}$ versus λ_1 (SNR=5 dB, TW=5, N=10) under AWGN channel.

us the maximum $\mathbf{P_D}$. In Figure 3.5, increasing λ_1 and solving the corresponding Δ gives us the maximum $\mathbf{P_D}$ for a given λ_1. Since it is a concave function, its maximum can be found by numerical methods easily. Therefore, we use the diminishing step size optimization method that increases λ_1 and finds the corresponding Δ.

The algorithm for finding the solution to the given formulation with given parameters is depicted in Figure 3.6. The function $findDeltaForPFEqualsAlpha(\lambda_1, \alpha, \overrightarrow{w})$ simply solves equation for the corresponding Δ that makes $\mathbf{P_F} = \alpha$ by shifting Δ to higher values with decreasing increments.

3.3.2. Improving Weights and Threshold Optimization

We introduce another optimization problem for optimizing $\overrightarrow{w}$ similar to UniQCS1.

$$UniQCS2: \quad \underset{\overrightarrow{w}}{maximize} \ \mathbf{P_D} \ subject \ to \ \mathbf{P_F} \leq \alpha \tag{3.18}$$

Different $\overrightarrow{w}$ vectors change $\mathcal{B}_1$, hence $\delta_{\overrightarrow{w}}(\overrightarrow{B})$ values. This change further affects $\mathbf{P_D}$ values evaluated optimally by UniQCS1. We optimize $\overrightarrow{w}$ via embedding a genetic algorithm into the optimization procedure of the proposed method since the objective

Require: $\overrightarrow{w}$, α

Initialize $maxPD$, λ_1, $\lambda_{increment}$, λ_1^*, Δ^*
while $\lambda_{increment} > precision$ **do**
  $\Delta = findDeltaForPFEqualsAlpha(\lambda_1, \alpha, \overrightarrow{w})$
  $\mathbf{P_D} = evaluatePD(\lambda_1, \Delta, \overrightarrow{w})$
  if $(\mathbf{P_D} > PD_{prev})$ **then**
    $\lambda_1 += \lambda_{increment}$
    $PD_{prev} = \mathbf{P_D}$
  else
    $\lambda_1 -= \lambda_{increment}$
    Decrease $\lambda_{increment}$
    $\lambda_1 += \lambda_{increment}$
  end if
  if $(\mathbf{P_D} > maxPD)$ **then**
    $maxPD \leftarrow \mathbf{P_D}$
    $\lambda_1^* \leftarrow \lambda_1$
    $\Delta^* \leftarrow \Delta$
  end if
end while

Figure 3.6. Threshold optimization algorithm.

function is not differentiable in terms of $\overrightarrow{w}$ and it is more appropriate for this kind of problems. The initial phase of the genetic algorithm selects initial weights. For different $\overrightarrow{w}$ values, thresholds are optimized by UniQCS1 and rated according to $\mathbf{P_D}$ values and the fit $\overrightarrow{w}$ individuals survive to the next generation. The crossover strategy averages the corresponding parent $\overrightarrow{w}$'s to form new individuals. In all generations five elites, which give offsprings to new population, are selected according to $\mathbf{P_D}$ values, and one individual experiences mutation. After several iterations, $\overrightarrow{w}$ improvement according to the parameters becomes insignificant and the algorithm stops. The flowchart of the genetic algorithm for weight improvements is presented in Figure 3.7.

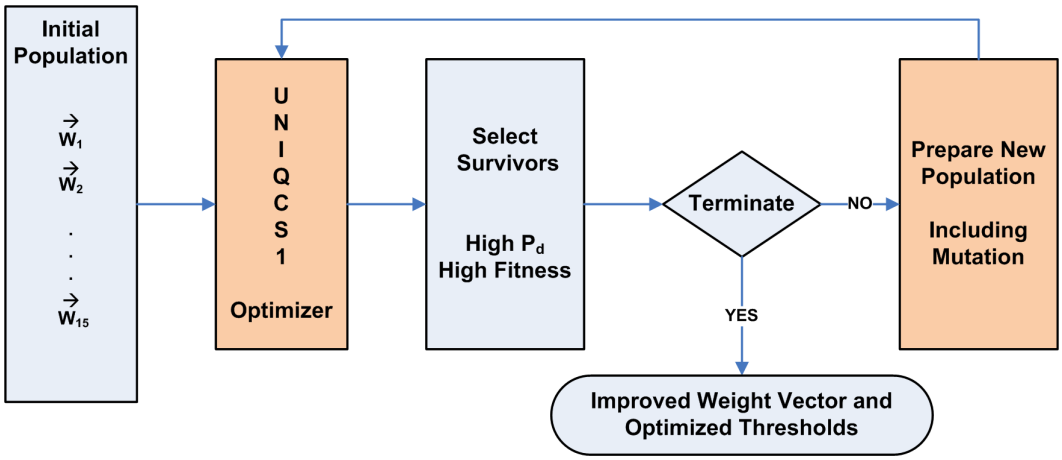


Figure 3.7. Flowchart of the genetic algorithm.

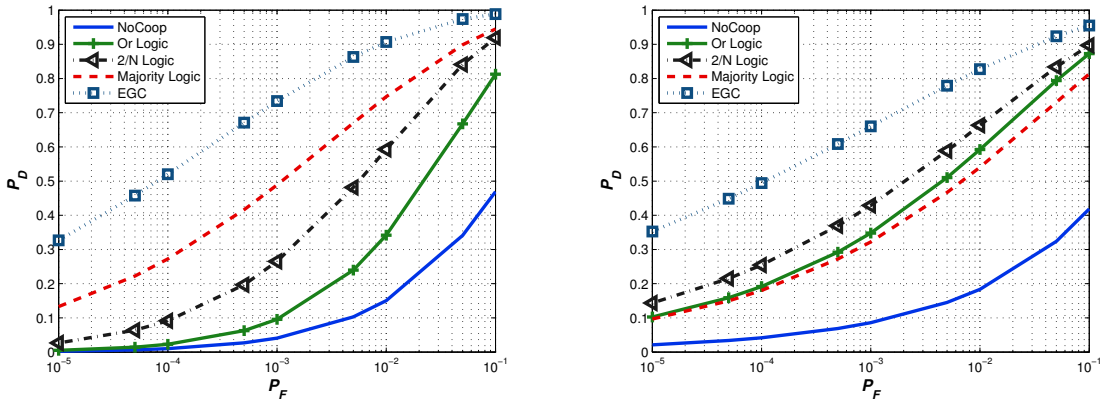
The genetic algorithm improves $\overrightarrow{w}$, and the thresholds are optimized by UniQCS1 according to the system parameters such as mean SNR, number of cooperating nodes, and channel conditions. Hence, we can evaluate the optimization variables for predetermined system parameters offline. If system parameters change, new values of optimization variables are evaluated by using UniQCS.

3.4. Results

We analyze the performance of UniQCS mostly under a flat fading Rayleigh channel. We evaluate the maximum $\mathbf{P_D}$ for a given $\mathbf{P_F}$ constraint. For comparisons, we include the "no cooperation" case and EGC when necessary. After an initial investigation, the performance is evaluated under no false reports and the perfect reporting channel conditions. Then, we introduce false reports to analyze the performance loss. Similarly, we introduce an imperfect reporting channel without false reports to analyze the effect of an imperfect reporting channel only.

3.4.1. Performance Metrics and System Parameters

P_d is an important metric of the system since it directly affects the primary users. A low probability of detection implies high missed detection, which results in a failure of the CR node in vacating the channel during primary access. Therefore, stringent



(a) AWGN channel (b) Rayleigh channel

Figure 3.8. ROC curves of different global decision logic functions (SNR=5 dB, TW=5, N=10).

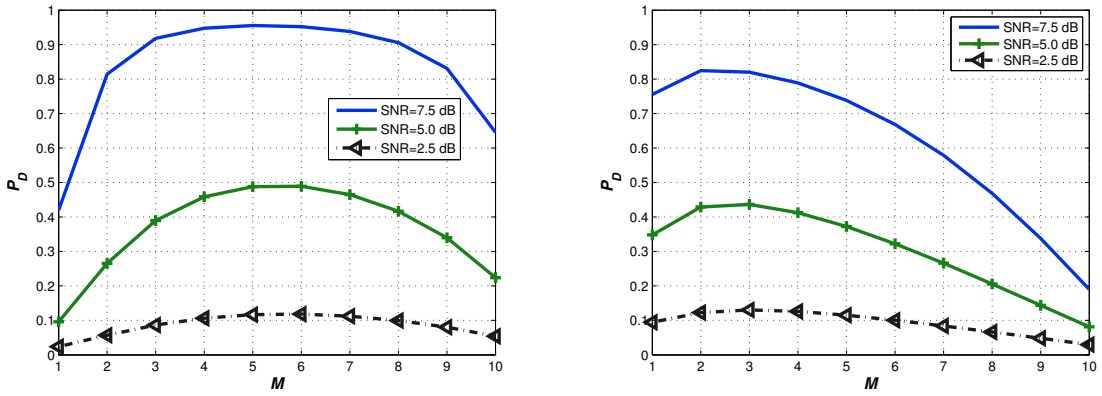
conditions are imposed on the value of P_d. On the other hand, having a high probability of false alarm, P_f, results in low utilization of white spaces. Therefore, the tradeoff between these two metrics must be considered. *Receiver Operating Characteristic (ROC)* curves are used for the tradeoff analysis between $\mathbf{P_D}$ and $\mathbf{P_F}$.

We consider the following system parameters in our analysis: average SNR, TW, and the number of cooperating users denoted by N. The UniQCS method optimizes its parameters by using the algorithms explained in Section 3.3 for UniQCS1 and UniQCS2. UniQCS2 elevates $\overrightarrow{w}$ in a suboptimal manner using the genetic algorithm while UniQCS1 finds (λ_1, Δ) in an optimal manner using numerical methods.

3.4.2. Initial Investigation

The initial investigation consists of two main parts. First, we analyze the hard decision logic functions under AWGN and Rayleigh channels as the preliminary work. We also include the EGC case to analyze how much room we have for improvement.

In Figure 3.8, ROC curves of hard decision logic functions and EGC are depicted. In the Rayleigh channel, the Majority logic falls behind the Or logic in terms of $\mathbf{P_D}$ since



(a) AWGN channel

(b) Rayleigh channel

Figure 3.9. M versus $\mathbf{P_D}$ curves with different SNRs ($\mathbf{P_F} = 0.001$, TW=5, N=10).

most of the nodes have SNR values below the mean SNR, which is 5 dB. Under AWGN and Rayleigh channels, we have room for improvement with the given parameters since EGC performs far better than the other methods.

In Figure 3.9, M versus $\mathbf{P_D}$ curves are depicted for different SNRs. M=1 corresponds to the Or logic since M-out-of-N with M being equal to 1 is equivalent to 1/N logic. We analyze the hard decision logic functions under different SNRs. For the AWGN channel, the maximum $\mathbf{P_D}$ is attained at approximately 6, 5, and 4 for mean SNR equal to 2.5 dB, 5.0 dB, and 7.5 dB, respectively, which clearly suggests us to use different global decision logic functions for different channel conditions. For the Rayleigh channel, the performance in terms of $\mathbf{P_D}$ increases up to M=3 since most of the nodes have an SNR below the mean SNR. Therefore, a hard decision logic similar to Majority logic is not favored.

Investigations suggest us that the quantization and the proposed global decision promise significant improvements. For the preliminary studies, we consider a quantizer as depicted in Figure 3.10. We consider the following system parameters in our analysis: channel condition, average SNR, TW, and number of cooperating users denoted by N. The CSDV method also has its own parameters such as Δ and Δ_c. We use weight vectors $\overrightarrow{w}$ = [-2 -1 0 1 2] for the 4-threshold case and $\overrightarrow{w}$ = [-3 -2 -1 0 1 2 3] for the

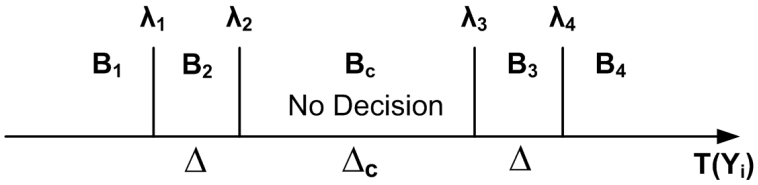


Figure 3.10. Multi threshold ED with four thresholds for preliminary analysis.

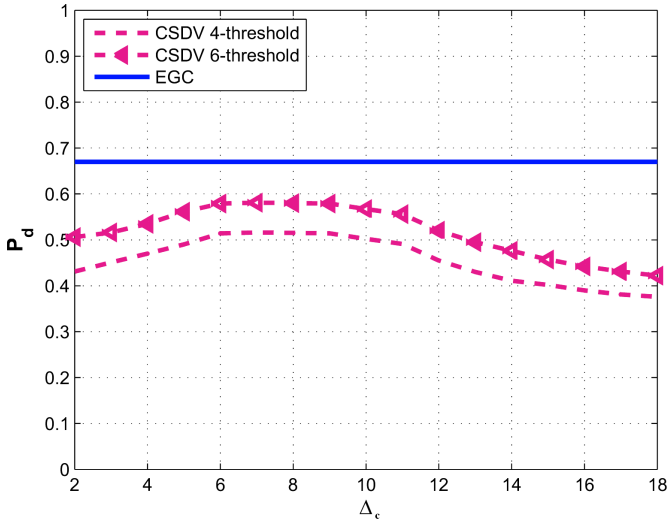


Figure 3.11. Effect of Δ_c on $\mathbf{P_D}$ under a Rayleigh channel (SNR=5 dB, TW=5, $\mathbf{P_F} = 0.001$).

6-threshold case. We use $\Delta = 4$ and number of cooperating nodes is set to 10 for all cases. Therefore, we predetermine Δ and simplify the optimization problem and transform it to linear search of λ_1 for the preliminary study.

The first parameter we study is the Δ_c parameter of CSDV. In Figure 3.11, x-axis corresponds to different Δ_c values and the y-axis represents the corresponding $\mathbf{P_D}$ values. It is apparent that for both 4-threshold and 6-threshold cases, having $\Delta_c = 6$ gives the best result. We use $\Delta_c = 6$ for the rest of the preliminary investigation.

We also analyze the effect of number of thresholds on the sensing performance with the help of ROC curves. Figure 3.12 depicts the ROC curves of CSDV for different number of thresholds under a Rayleigh channel. The sensing performance of CSDV

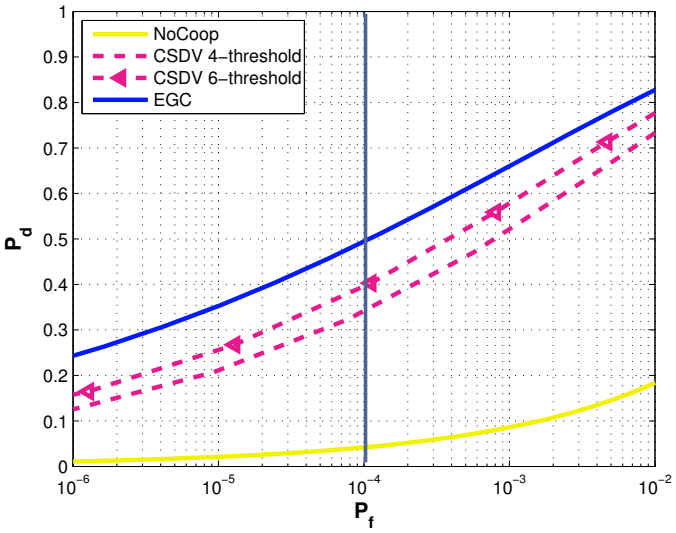


Figure 3.12. ROC curves for CSDV for different number of thresholds under a Rayleigh channel (SNR=5 dB, TW=5).

with 6-threshold is better than that of the 4-threshold case. In a Rayleigh channel, varying SNR value degrades the sensing performance. The CSDV performance is not so close to the EGC performance, hence, we have room for a performance gain via increasing the number of thresholds. Using 6-threshold CSDV improves the system performance considerably.

Effect of false reports is analyzed in Figure 3.13. In the AWGN channel condition, CSDV performs better than Majority and M/N Logic even if there are false reporting nodes. M/N Logic has the highest performance loss since half of the required reports are provided by false reports. Table 3.2 depicts the performance gain of CSDV compared to Majority and M/N Logic methods when there are two false reports in the AWGN channel.

In the preliminary results, we focus on collaborative ED and uniform quantizer improvements against low SNR conditions. By transmitting 2 or 3-bit information, the CSDV method performs very close to EGC scheme which transmits exact measurements and outperforms the hard decision methods. Thus, we achieve close to the EGC performance at the expense of transmitting one or two extra bits compared to

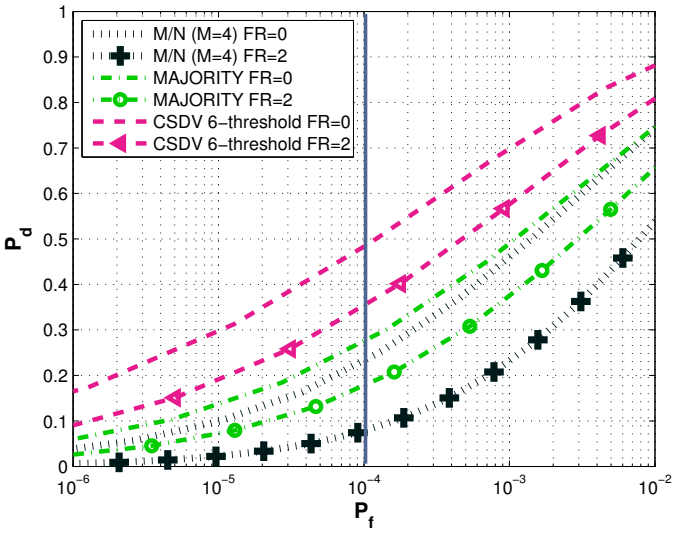


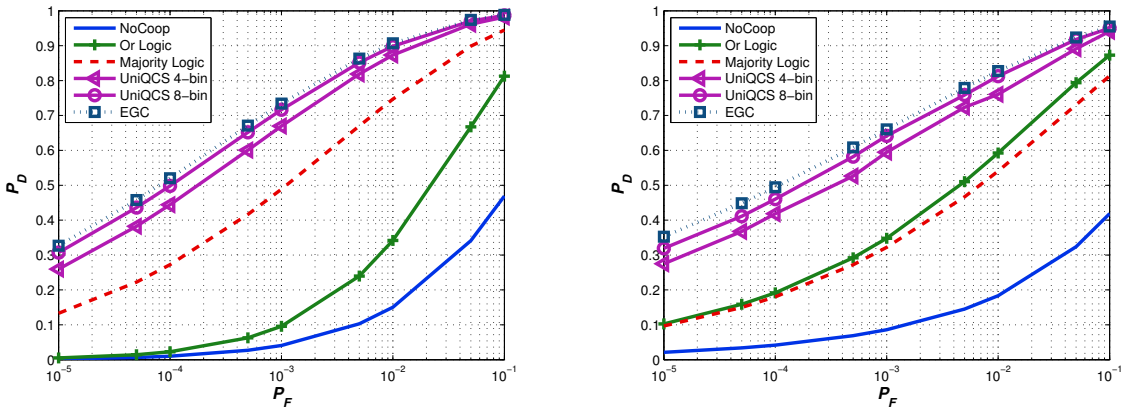
Figure 3.13. ROC curves for different cooperation schemes including false reports under AWGN channel (SNR=5 dB, TW=5).

Table 3.2. Robustness against false reports in AWGN channel.

	Perf. gain of CSDV compared to	
Parameters	Majority	M/N
$fr = 0$, $\mathbf{P_F} = 10^{-5}$	105%	200%
$fr = 2$, $\mathbf{P_F} = 10^{-5}$	150%	400%
$fr = 0$, $\mathbf{P_F} = 10^{-4}$	52%	92%
$fr = 2$, $\mathbf{P_F} = 10^{-4}$	100%	300%
$fr = 0$, $\mathbf{P_F} = 10^{-3}$	40%	48%
$fr = 2$, $\mathbf{P_F} = 10^{-3}$	50%	135%

the hard decision methods. The CSDV achieves performance gain compared to Majority Logic, between 40% and 100% in an AWGN channel for the operating region of $\mathbf{P_F} \in (0.00001,\ 0.001)$. We also analyze the robustness to false reports and the CSDV outperforms Majority and M/N Logic in the AWGN channel. The CSDV method performs better than Majority and M/N Logic (without false reports) even if there are two false reports for CSDV.

The next steps in our research focus on optimizing the quantizer thresholds and



(a) AWGN channel (b) Rayleigh channel

Figure 3.14. ROC curves of different global decision logic functions (SNR=5 dB, TW=5, N=10).

improving the weights. We are motivated with the preliminary results to work on the general problem. Thresholds and the weights cannot be selected dynamically as in the case of preliminary results and they require optimization.

3.4.3. No False Reports and Perfect Reporting Channel

After the preliminary investigation, we solve the optimization problems and optimize the quantizer thresholds while improving the weights of the global decision function via a genetic algorithm as explained in Section 3.3. First, we examine the performance of the proposed method under the perfect reporting channel when there are no false reports.

3.4.3.1. Analysis of the Global Decision Logic on $\mathbf{P_D}$. In Figure 3.14a and 3.14b, we analyze the performance of different global decision logic functions by using ROC curves under AWGN and Rayleigh channels, respectively. In the figures, $\mathbf{P_D}$ values corresponding to the given $\mathbf{P_F}$ constraint are plotted. UniQCS performs close to EGC with 8-bin quantization and far better than the hard decision logic functions. For $\mathbf{P_F} = 10^{-2}$, UniQCS 8-bin achieves 0.9 and 0.81 $\mathbf{P_D}$ under AWGN and Rayleigh channels, respectively. For the same $\mathbf{P_F}$, Majority logic achieves 0.75 and 0.54, respectively.

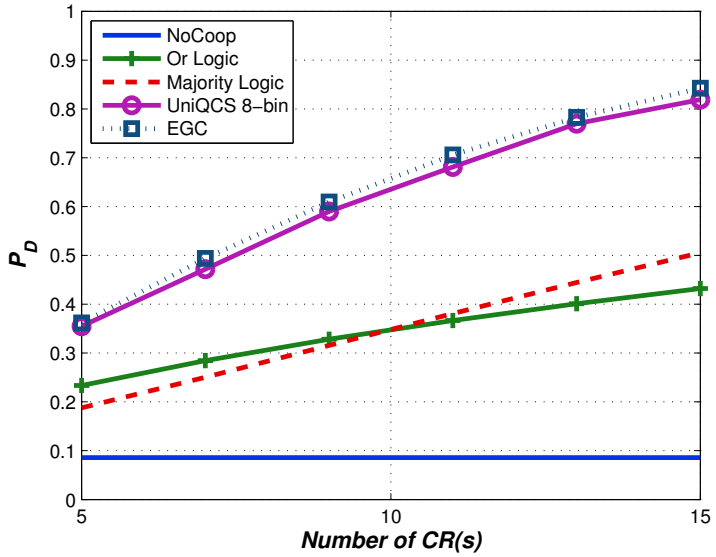


Figure 3.15. N versus $\mathbf{P_D}$ curves of different global decision logic functions under Rayleigh channel ($\mathbf{P_F} = 0.001$, SNR=5 dB, TW=5).

The performance gain of UniQCS compared to Majority in this situation is approximately 21% and 50%, respectively. In a Rayleigh channel, Or logic performs better than the Majority logic for 5 dB SNR since there are fewer nodes with a high SNR value, favoring Or logic. Starting from this point, in this thesis, we consider the Rayleigh channel, and for the UniQCS case, we analyze the system performance considering only 8-bin quantization.

3.4.3.2. Effect of Number of Users on $\mathbf{P_D}$. In Figure 3.15, we analyze the effect of number of users on $\mathbf{P_D}$ for a given $\mathbf{P_F}$. In the figure, the x-axis and the y-axis correspond to the number of nodes and $\mathbf{P_D}$, respectively. Increase in the number of nodes results in higher $\mathbf{P_D}$. UniQCS and EGC exhibit similar performances where EGC is an upper bound for the case since we do not distinguish users according to their SNR values, i.e. no differentiation among the nodes. UniQCS 8-bin improves the system performance and performs close to EGC by just using 3-bit reporting bits. For achieving $\mathbf{P_D}$ higher than 0.9, UniQCS needs at least 14 nodes. None of the hard decision logic functions can achieve $\mathbf{P_D} = 0.9$ constraint with the given system parameters. Majority and Or logic functions have a crossing point since the SNR is exponentially

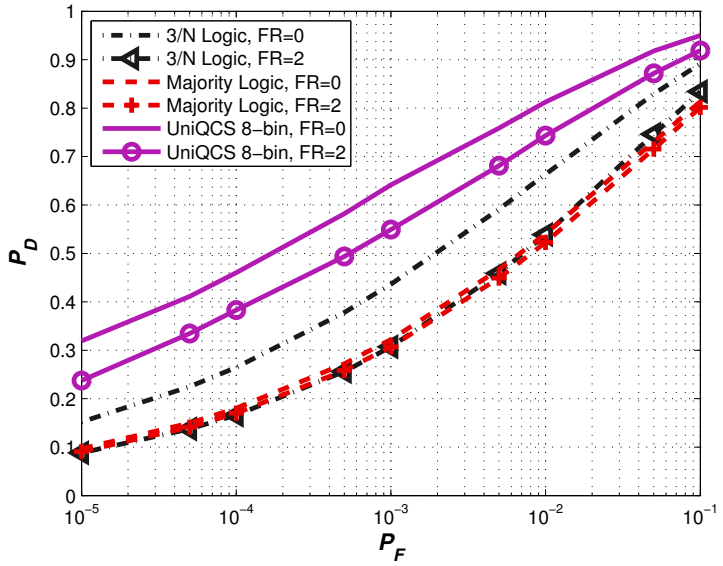


Figure 3.16. ROC curves of different global decision logic functions including false reports (SNR=5 dB, TW=5, N=10).

distributed and there are not enough CR nodes to take advantage of the Majority logic up to 10 CR nodes. The performance gain of UniQCS is at least 60% compared to the hard decision methods for 15 CR nodes.

3.4.4. Robustness Against False Reports

In this subsection, we examine the performance of the proposed method under the perfect reporting channel while there are two false reporting nodes. Since there are two false reports, we do not include the Or logic in the performance figures. Or logic and 2/N logic clearly fail ($\mathbf{P_D} = 1 = \mathbf{P_F}$). Therefore, the utilization of white spaces becomes impossible. Hence, we include 3/N logic instead of Or logic and 2/N logic for comparisons.

3.4.4.1. Analysis of Global Decision Logic on $\mathbf{P_D}$. False reports adversely affect 3/N logic more than any other logic in Figure 3.16. UniQCS is also affected by false reports, but UniQCS with two false reports still performs better than other methods without any false reports. The Majority logic is the least affected method due to false reports.

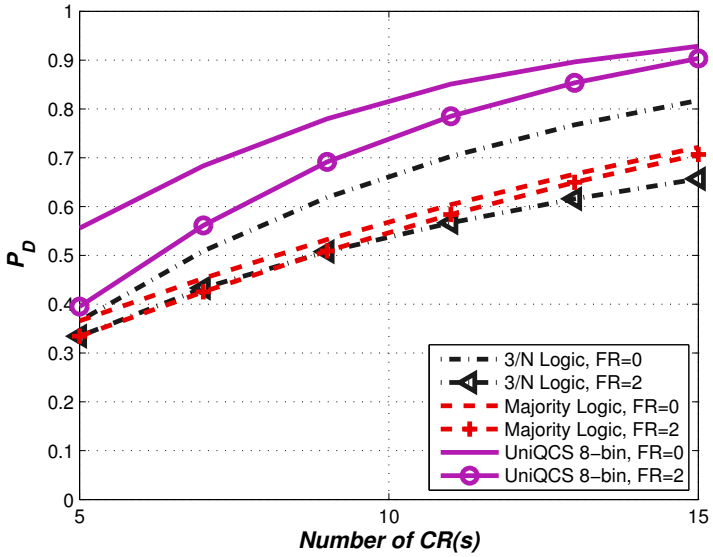


Figure 3.17. N versus $\mathbf{P_D}$ curves of different global decision logic functions including false reports ($\mathbf{P_F} = 0.01$, SNR=5 dB, TW=5).

However, its performance is far below the UniQCS. For $\mathbf{P_F} = 10^{-2}$, the performance loss due to false reporting nodes is approximately 7%, 18%, and 2% for UniQCS, 3/N, and Majority logic, respectively. For the same $\mathbf{P_F}$, the performance gain of UniQCS is between 10% and 140% compared to the other methods under false reports.

3.4.4.2. Effect of Number of Users on $\mathbf{P_D}$. In Figure 3.17, a similar behavior is observed as in the previous false report analysis figures. UniQCS is also affected by false reports, but UniQCS with two false reports again performs better than other methods even if they do not experience any false reports. For fewer nodes the gain of UniQCS over other algorithms decreases. Increasing the number of cooperating nodes results in increase in robustness against false reports since the performance loss decreases. For 15 cooperative nodes, the performance loss due to false reports for UniQCS, 3/N, and Majority logic are approximately 1.6%, 19%, and 1.3%, respectively. The Majority logic performance loss is less than the others, but its performance is not better than UniQCS. None of the methods other than UniQCS can achieve $\mathbf{P_D}$ higher than 0.9 with the given parameters. UniQCS is robust against false reports and performs better than other methods.

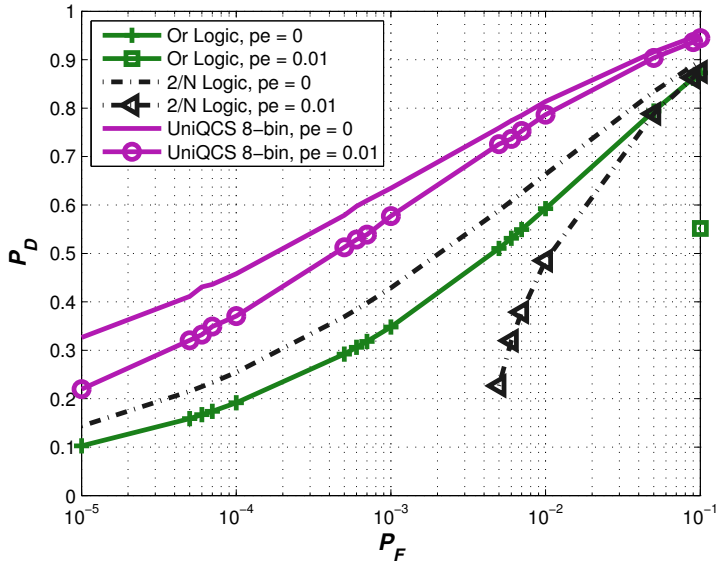


Figure 3.18. ROC curves of different global decision logic functions including imperfect reporting channel (SNR=5 dB, TW=5, N=10).

3.4.5. Robustness Against Imperfect Reporting Channel

In this subsection, we examine the performance of the proposed method under an imperfect reporting channel while there are no false reports.

3.4.5.1. Analysis of Global Decision Logic on $\mathbf{P_D}$. In Figure 3.18, we compare the perfect and imperfect channel cases. Some of the curves are incomplete since the target $\mathbf{P_F}$ is not achievable. Under an imperfect channel, Or logic is the most adversely affected method. Smallest $\mathbf{P_F}$ that is achievable for Or Logic and 2/N logic are 0.1 and 0.005, respectively, and smaller $\mathbf{P_F}$ values are not achievable. The performance loss due to imperfect reporting channel in UniQCS decreases as $\mathbf{P_F}$ increases. Again, UniQCS performs better than other hard decision logic functions even if they have the perfect reporting channel. All the false alarm probabilities in the evaluation range are attainable by UniQCS with an acceptable performance loss. For $\mathbf{P_F} = 0.05$, UniQCS achieves $\mathbf{P_D} > 0.9$ while 2/N achieves $\mathbf{P_D} = 0.8$. Therefore, the performance gain of the proposed algorithm is higher than the other methods explicitly.

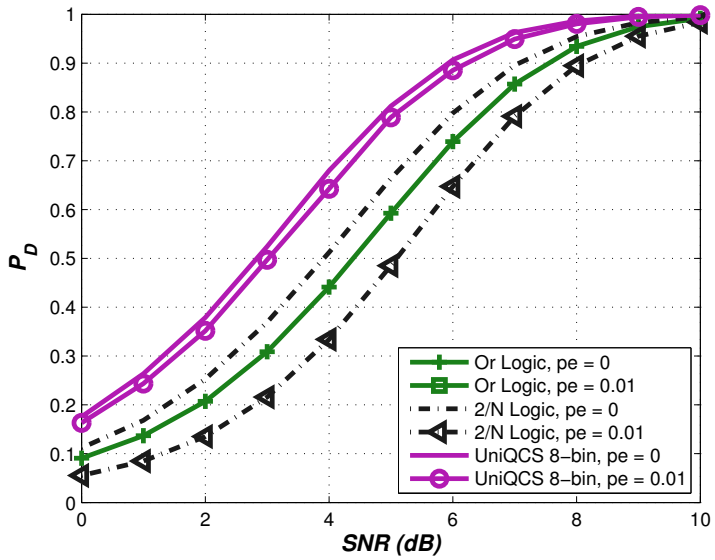


Figure 3.19. SNR versus $\mathbf{P_D}$ curves of different global decision logic functions including imperfect reporting channel ($\mathbf{P_F} = 0.01$, TW=5, N=10).

3.4.5.2. Effect of SNR on $\mathbf{P_D}$. Effect of SNR on $\mathbf{P_D}$ is analyzed in Figure 3.19. As SNR increases, $\mathbf{P_D}$ increases, and we observe that Or logic cannot achieve $\mathbf{P_F} = 0.01$ at all in case of the imperfect reporting channel. 2/N logic is also adversely affected by $\mathbf{P}_e = 0.01$ while UniQCS is robust against reporting channel errors. If we consider the 5 dB case, UniQCS performance gain is at least 50% compared to the other methods under imperfect reporting channel conditions. Increasing SNR results in higher $\mathbf{P_D}$ for 2/N, but Or logic cannot achieve $\mathbf{P_F} = 0.01$ even for high SNR values.

The performance loss due to $\mathbf{P}_e$ for different SNR values are also depicted in Figure 3.19. UniQCS 8-bin is also adversely affected. However, the adverse effects of the imperfect channel are limited. For the 5 dB case, UniQCS 8-bin performance loss is approximately 3% while 2/N performance loss is approximately 29%. We cannot consider the Or logic performance loss since it cannot satisfy the $\mathbf{P_F}$ constraint. Hence, the results strongly suggest the use of the UniQCS method for the given system parameters in case of the bandwidth limited reporting channel.

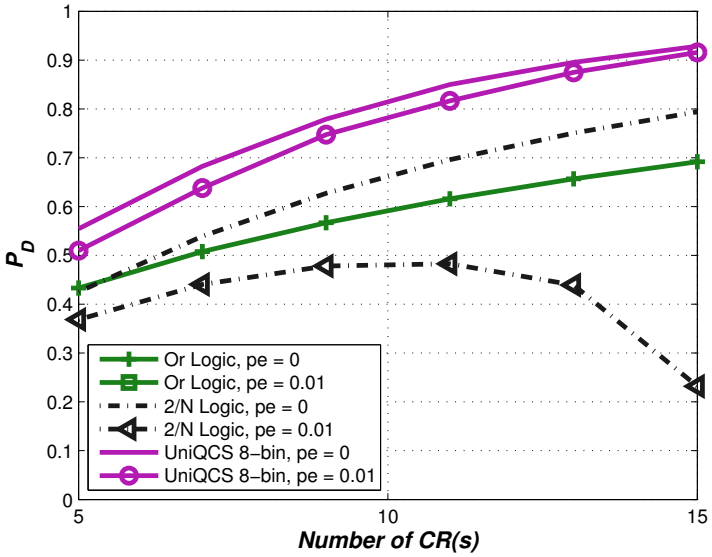


Figure 3.20. N versus $\mathbf{P_D}$ curves of different global decision logic functions including imperfect reporting channel ($\mathbf{P_F} = 0.01$, SNR=5 dB, TW=5).

3.4.5.3. Effect of Number of Users on $\mathbf{P_D}$. In Figure 3.20, the effect of number of users on $\mathbf{P_D}$ is analyzed while the target $\mathbf{P_F}$ is 0.01. Under an imperfect channel, Or logic is the most adversely affected method. Moreover, Or logic cannot achieve $\mathbf{P_F} = 0.01$ for the given parameters under an imperfect channel, hence it is not included as a curve in the figure. As the number of CR nodes increases, 2/N logic performs worse since it is sufficient to have two nodes deciding H_1 and the probability of having two errors increases with more nodes when the average SNR evaluated after the introduction of the bit error probability is below the SNR wall [94]. Once again, UniQCS is robust against the imperfect reporting channel and performs better than other methods even if the reporting channel is perfect for other methods.

The performance loss due to $\mathbf{P}_e$ for different number of users are also depicted in Figure 3.20. UniQCS 8-bin is also adversely affected. However, the adverse effects of the imperfect channel is limited. For the 15 CR case, UniQCS 8-bin performance loss is approximately 1% while 2/N performance loss is approximately 75%. We cannot consider the Or logic performance loss since it cannot satisfy the $\mathbf{P_F}$ constraint. Since the cooperation gain for 2/N is not enough to resist to adverse effects of the SNR

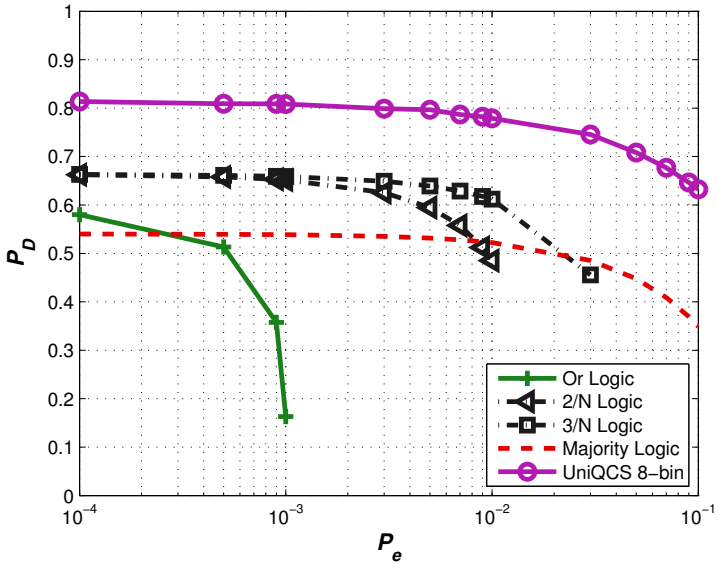


Figure 3.21. $\mathbf{P}_e$ versus $\mathbf{P_D}$ curves of different global decision logic functions including imperfect reporting channel ($\mathbf{P_F} = 0.01$, SNR=5 dB, TW=5, N=10).

wall, increasing the number of nodes decreases the performance of the system. In contrary, UniQCS 8-bin cooperation scheme beats the SNR wall. Therefore, increasing the number of nodes positively affects the system performance.

3.4.5.4. Effect of $\mathbf{P}_e$ on $\mathbf{P_D}$. Effect of $\mathbf{P}_e$ on $\mathbf{P_D}$ is analyzed in Figure 3.21. Increasing $\mathbf{P}_e$ decreases $\mathbf{P_D}$, if $\mathbf{P_F} = 0.01$ is achievable. UniQCS and Majority logic are robust against an imperfect reporting channel. However, the Majority logic performance is much lower than that of UniQCS. The proposed method performance gain is at least 15% compared to other methods for $\mathbf{P}_e = 10^{-3}$. Moreover, BEP wall for UniQCS is higher than the other methods as seen from the figure. All the false alarm probabilities in the evaluation range are attainable by UniQCS with an acceptable performance loss. Majority logic and UniQCS 8-bin behave similarly with performance difference. UniQCS 8-bin outperforms the other methods in terms of $\mathbf{P_D}$ with given parameters. Up to $\mathbf{P}_e = 10^{-2}$, the UniQCS 8-bin performance loss is negligible. Or logic can handle at most $\mathbf{P}_e = 10^{-3}$ with $\mathbf{P_D} = 0.18$, which is not acceptable. For higher $\mathbf{P}_e$, Or logic cannot satisfy the $\mathbf{P_F} = 0.01$ constraint with given system parameters. Similar behavior is exhibited by 2/N and 3/N for $\mathbf{P}_e \approx 10^{-2}$.

3.5. Chapter Summary

Awareness about the wireless environment is the most crucial part of the CR system since it includes sensing the environment for the primary activity, finding white spaces, and vacating the channel or adjusting the communication parameters which is a must for not disturbing the primary users or constructing the radio environment map.

In this chapter, we focus on cooperative ED and uniform quantizer improvements against low SNR conditions under imperfect reporting channel and false reports. By transmitting 2- or 3-bit information, the proposed UniQCS method performs very close to that of the EGC scheme which transmits exact measurements and outperforms the hard decision methods. Our proposed method, UniQCS, achieves a performance close-to EGC scheme, which is an upper bound since the nodes are not differentiated, at the expense of transmitting one or two extra bits compared to the hard decision methods. UniQCS is also robust against imperfect reporting channels and false reports.

For future directions, we will work on adaptive weights, multi-hop performance analysis for reporting the quantized measurement, and on developing an optimal sensing framework.

4. LOCATION ESTIMATION BASED REM GENERATION

4.1. Incentives and Contributions

The research on CR has matured from early works on the idea of CRs to today's multidimensional proposals such as signal processing techniques, machine learning algorithms, and regulations for cognitive radio. However, there still remain many technical challenges to overcome for evolving CRs into practical networks. TVWS, the portion of the UHF and VHF bands emerging after DSO, has been an effective catalyst for the first practical applications of CRNs. Initiated by the FCC in the USA, the use of TVWS for dynamic spectrum access has attracted worldwide interest such as the UK, Brazil, Japan, India, Singapore, and China. DSO is almost completed in Japan except for the three prefectures devastated by the 11th March earthquake, in Canada, and some parts of Europe. It is scheduled to be completely finished in most of the Europe by 2012. South Korea, Australia, China, and Brazil are expected to switch off not later than 2015 [54]. Turkey launched trial transmissions in 2006 and originally planned to gradually do the switch by 2014 [56].

White space measurements show that almost 50% of locations in the UK have at least 150 MHz bandwidth while 90% have at least 100 MHz capacity [53]. Similarly, measurement campaigns in many countries all over the world like the USA [58], Japan [59], and Spain indicate the potential of TVWS in meeting the escalating wireless connectivity requirements. Opening TVWS for secondary access accelerates the practical implementations of CRNs.

As the first step of opportunistic spectrum access, a TVWS device must establish the available frequencies at its location just before communications and must ensure the PUs are not harmed. Due to the concerns mainly raised by the broadcasting industry on the proper protection of the incumbent TV services, regulatory bodies have brought

strict rules on the white space devices. The rules demanded by the regulators such as the detection of a primary signal at -114 dBm by FCC are over-conservative. Such regulations reduce the available white spaces drastically (approximately by a factor of three) [58]. Instead, accessing a centralized database, which keeps track of the available spectrum based on geographical coordinates or the properties of registered PU transmitters, has been accepted more promising for mitigating the strict sensing challenges while meeting the regulatory obligations [95].

The Geolocation database stores TV base station locations and the related parameters (static data) and constructs a real-time view of the spectrum occupancy at the TV bands at each location [95], which is a quasi-static information. Currently, the database based TVWS operation is being pursued by regulators in the USA and the UK. In November 2009, FCC opened a call for database administrators and approved ten companies (Google, Spectrum Bridge, Telcordia, Microsoft, etc.) as administrators in 2010. A step further from the geolocation database lies the Radio Environment Map (REM). The REM was first defined as an abstraction of a real-world environment storing multi-domain information dynamically [78]. The REM can act as an enabler for cognition in radios; it can store PU activity statistics as well as RF environment information such as propagation characteristics even if it is dynamic information.

Essential functionality of a REM is the construction of a dynamic interference map for each frequency at each location of interest which is volatile data. Since it is impractical to have measurements at each location in the CR operation area, the REM fuses the available measurements to estimate the interference level at locations with no measurement data. The radio interference map interpolation is the most critical part of a REM construction method.

In this thesis, we also propose a primary user *LocatIon Estimation based (LIvE)* REM construction technique in fading channels. We mainly focus on the radio interference map interpolation by utilizing the measurements from known locations. The proposed technique outperforms other REM construction techniques and its time com-

plexity is better than Kriging and inverse distance weighted interpolation which are commonly used techniques.

Contributions of LIvE REM construction can be listed as:

(i) We propose a new useful metric, correct detection zone ratio, for the REM quality.
(ii) LIvE REM construction technique achieves high performance in terms of RMSE and correct detection zone ratio.
(iii) The time complexity of LIvE is half of that of Kriging while performing better.
(iv) LIvE REM construction technique performance does not depend on the closeness of the PU to the measurement capable devices.

4.2. Radio Environment Map

A geolocation database stores various information about PUs, and using this information, it derives the available channel list at a location. CRs consult this database and retrieve the available channel list, including time constraints and the maximum transmission power. Various information kept by the geolocation database on PU (and possibly CR) transmitters should include TV tower location, antenna height, user type (PU or CR), device transmit power, technology, operation channel(s), and duration of use [66].

The REM was first defined as an abstraction of a real-world environment storing multi-domain information [78]. However, it can also be considered more generally as an intelligent network entity that can further process the gathered information, inspect the spatio-temporal characteristics, and derive a map of the RF environment [79]. The REM is a promising concept for efficient CRN operation without extensive burden on CRs as it can be considered as the *cognitive engine* located at the network. The REM introduces environment awareness that would be harder to acquire by individual CR capabilities via extensive spectrum analysis. Hence, the REM can also be seen as the network support turning simple nodes into intelligent CRs. The design of the

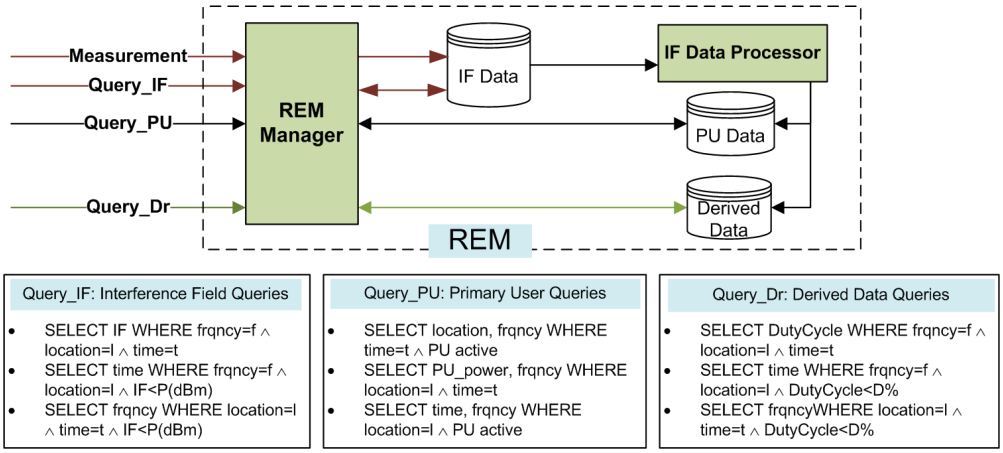


Figure 4.1. REM data flow and example queries.

REM relies on the following related processes: data gathering and representation, data processing/fusion, and data retrieval/query as depicted in Figure 4.1.

4.2.1. REM Data and System Model

The data stored in a REM can be classified into three as static, volatile, and derived [80]. Static data corresponds to the information that does not change frequently, e.g., the locations of TV towers, or the operators in a region similar to the stored data on the geolocation database. On the contrary, volatile data represents the information that is highly dynamic. CR nodes in a network and their spectrum use information can be categorized under this class. The REM keeps its information about the environment up to date by dynamically tracking the changes. Lastly, the derived data is interpreted from static and volatile data by the processing techniques depicted in Figure 4.1.

The data stored in a REM depends on the measurements collected from CR nodes in the network and other Measurement Capable Devices (MCD). Each measurement report also has the geolocation and time stamps. MCDs can be specific purpose dedicated sensor nodes, mobile phones, or similar devices.

The sensing data records should contain the following information:

Table 4.1. The model of the data reported by the MCDs.

Parameter	Size	Units
RSS	Vector	[dBm]
Frequency Band	Vector	[MHz]
Time Instance	Scalar	[seconds]
x-coordinate	Scalar	[Latitude]
y-coordinate	Scalar	[Longitude]
z-coordinate	Scalar	[Altitude]
MCD Identifier	Label	[no dimension]

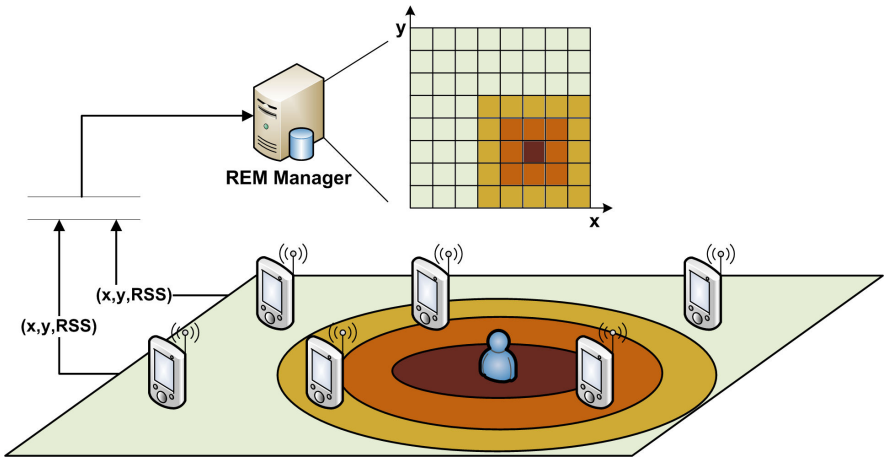


Figure 4.2. System model of CR operation area.

- Received primary (and/or secondary) signal levels
- Location of the measurement (longitude, latitude, altitude)
- Time stamp
- Inspected frequency bands

The raw sensing data depicted in Table 4.1 (adapted from [80]) are reported to functional entity REM Manager that exploits the reported information and information from the geolocation DB to build a complete map by interpolating the geo-localized measurements. The REM Manager persists the raw data and the interpolated data to a database for responding queries easily and fast.

Our system model is depicted in Figure 4.2. MCDs are placed in a grid-like layout in the operation area and a primary user is actively using a frequency band. These MCDs may be both devices deployed specifically by the cognitive radio operator or secondary user devices. Circular regions represent the different interference level contours. REM manager collects the $(x, y, RSS,$ Time stamp $)$ values from the measuring nodes and constructs the radio interference map via interpolating the RSS values at locations without any measurements.

RSS values depend on the characteristics of the propagation environment. We consider the *log-normal shadowing* and *Rayleigh fading* environments for the fading channel analysis. For the analysis, we propose new useful quality metrics for REM evaluation such as correct detection and false alarm zone ratios.

4.2.2. REM Quality Metrics

How accurate the REM describes the real operation environment is measured by the REM quality metrics. These quality metrics can be evaluated by comparing the estimated REM with the true REM in terms of the correct detection zone, false alarm zone, and RMSE. Number of sensors, distribution of sensors in the field, capability of sensors, dynamics of the propagation environment, and accuracy of the propagation modeling are the key parameters affecting the performance of the REM.

The intersection of the zones of the true REM and the estimated REM defines the zones that are correctly and incorrectly determined. In Figure 4.3, the true REM and the estimated REM contours are depicted. The solid line divides the area into two subareas in which the transmission by secondary users is banned and allowed, denoted by dark and light shaded areas. The subscripts of the zones depict the actual state with respect to the true REM. The areas where the true REM and the estimated REM contours both determine H_1 or H_0 zones correctly are called the Correct Detection Zone type-1 (Z_1^{CD}) and Correct Detection Zone type-0 (Z_0^{CD}), respectively. False Alarm Zone (Z_0^{FA}) is the area in which transmission is forbidden due to estimation

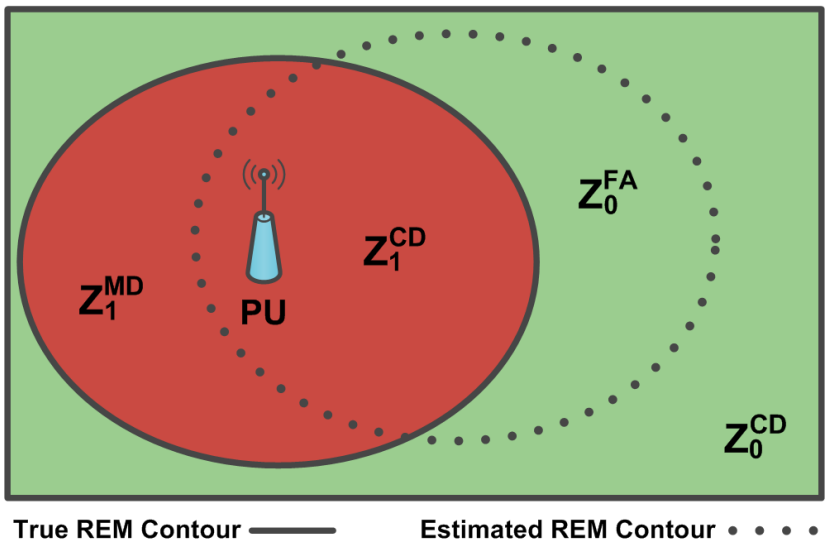


Figure 4.3. False alarm, missed detection, and correct detection zones.

errors though it would not harm the primary users. Finally, Missed Detection Zone (Z_1^{MD}) is the area where the transmission is not allowed but the estimated REM infers the opposite. *We propose new metrics for the performance evaluation of the REM construction techniques. We normalize the regions according to true REM and determine the Z_1^{CD} and Z_0^{FA} ratios (CDZR$_1$ and FAZR), which present a better understanding of the performance of the REM construction algorithm;*

$$\text{FAZR} = \frac{A(Z_0^{FA})}{A(Z_0^{FA}) + A(Z_0^{CD})}$$

$$\text{CDZR}_1 = \frac{A(Z_1^{CD})}{A(Z_1^{CD}) + A(Z_1^{MD})}$$

where $A(Z)$ represents the area of zone Z. FAZR and CDZR$_1$ are analogous to $\mathbf{P_F}$ and $\mathbf{P_D}$ in the spectrum sensing scope. Similarly, we define MDZR as

$$\text{MDZR} = \frac{A(Z_1^{MD})}{A(Z_1^{CD}) + A(Z_1^{MD})}.$$

Analogy of $\mathbf{P_D} + \mathbf{P_M} = 1$ is reflected as $\text{CDZR}_1 + \text{MDZR} = 1$.

4.3. RSS Measurements in Fading Channels

RSS measurements in fading channels are mainly characterized by the fading type. We use log-normal shadowing and Suzuki process for slow fading and fast fading channels, respectively. Suzuki process includes log-normal shadowing and Rayleigh fading hence, we use shadowing/fading shorthand for the fast fading channel.

We give the RSS measurements dynamics for slow and fast fading channel conditions. Consequently, we model the mean RSS in dB scale for the MCDs and depending on the fading type. The fading can cause poor performance in a communication system because it can result in a loss of signal power without reducing the power of the noise. This signal loss can be over some or all of the signal bandwidth. Therefore, it is important to model the channel and mean RSS since it affects the system performance while estimating the location of the primary signal and constructing the REM.

4.3.1. RSS Measurements in Slow Fading Channel

Long term fading arises when the coherence time of the channel is large relative to the delay constraint of the channel. In this regime, the amplitude and phase change imposed by the channel can be considered roughly constant over the period of use. Slow fading can be caused by events such as shadowing, where a large obstruction such as a hill or large building obscures the main signal path between the transmitter and the receiver. The amplitude change caused by shadowing is often modeled using a log-normal distribution.

Under a log-normal channel, RSS measurements in dB obey normal distribution. Hence, the analysis in dB is more appropriate. The ideal RSS at the ith MCD, denoted by P_i^{rx}, is expressed as

$$P_i^{rx}[dB] = P^{tx}[dB] - P_{L_0} - 10\log_{10} d^{\alpha}_{(x_i,y_i)} + S_i \tag{4.1}$$

where P_{L_0} and α are path loss correction and path loss exponents, respectively. P^{tx} is the transmit power of the transmitter and $d_{(x_i,y_i)}$ is the distance between active primary user and the ith secondary user or MCD. S_i is a Gaussian random variable with mean zero and variance σ_s^2 for expressing the effect of log-normal shadowing.

RSS values at each MCD has severe disturbances caused by the shadowing effect, so we measure raw RSS values and evaluate sample mean to reduce the shadowing effect. Mean RSS is formulated as

$$\overline{P_i^{rx}}[dB] = \sum_{j=1}^{N_m} P_{i,j}^{rx}[dB]/N_m \tag{4.2}$$

where $P_{i,j}^{rx}$ is the jth measured RSS value at the ith MCD and N_m is the total number of samples for unit measurement. Since the mean of S_i is zero, mean RSS value at ith MCD has the mean $P^{tx}[dB] - P_{L_0} - 10\log_{10} d_{(x_i,y_i)}^{\alpha}$.

4.3.2. RSS Measurements in Fast Fading Channel

The short term fading occurs when the coherence time of the channel is small relative to the delay constraint of the channel. In this regime, the amplitude and phase change imposed by the channel fluctuates considerably over the period of use. We use Rayleigh fading model over log-normal shadowing. Rayleigh fading is most applicable when there is no dominant propagation along a line of sight between the transmitter and receiver. In Rayleigh fading environment, instantaneous power exhibits an exponential distribution (not in dB). We consider shadowing/fading (log-normal shadowing plus Rayleigh fading) effects and use Suzuki distribution. The mobile channels modeled by Suzuki process already include multipath propagation and shadowing [96].

RSS measurements in Watts obey an exponential distribution with mean received from log-normal distribution. Hence, the analysis in Watts is more appropriate in this case (shadowing/fading channel). The ideal RSS at the ith MCD, denoted by P_i^{rx}, is

expressed as

$$P_i^{rx}[Watts] = E_i(10^{\frac{\phi}{10}}) \tag{4.3}$$

where $E_i(\mu)$ is an exponential random variable with mean μ and

$$\phi = P^{tx}[dB] - P_{L_0} - 10\log_{10} d^{\alpha}_{(x_i,y_i)} + S_i \tag{4.4}$$

where S_i is a Gaussian random variable with mean zero and variance σ_s^2 for expressing the effect of log-normal shadowing.

Since Rayleigh fading and shadowing cause statistically independent multiplicative fluctuations of the received power, attenuation in dB-values caused by shadowing and Rayleigh fading can be added. The results indicate that if logarithmic moments of the approximate log-normal distribution would be matched to the exact Suzuki distribution, the mean would be found to be 2.5 dB below the local-mean caused by shadowing [96]. Hence to reduce the fading effects, we evaluate the sample mean of measurements and consider the loss due to short term fading. The mean power in this case is

$$\overline{P_i^{rx}}[dB] = \sum_{j=1}^{N_m} P_{i,j}^{rx}[dB]/N_m + 2.5[dB] \tag{4.5}$$

where N_m is the total number of samples for unit measurement, $P_{i,j}^{rx}$ is the jth measured RSS value at the ith MCD in Watts and follows an exponential distribution. We derive RSS value from Exponential distribution in Watts and convert them in dB to evaluate the mean in dB scale.

Channel model affects the RSS measurements and it depends on the coefficients α and σ_s. These parameters are modeled for some environments and situations via extensive measurements and data analysis. We do not assume we have the knowledge of σ_s, but we assume we know α and P_{L_0} for REM construction.

Table 4.2. Empirical measurements for indoor and outdoor propagation.

Building Type	α	σ_s[dB]
Vacuum, infinite space	2.0	0
Retail store (indoor)	2.2-2.5	8.7
Office (indoor)	2.4-3.3	7.0-14.1
Urban cellular radio	2.7-3.5	3.0-8.0
Shadowed urban cellular radio	3.0-5.0	4.0-8.0
Suburban cellular radio	3.0-5.0	6.0-8.0

4.3.3. Empirical Measurements

Empirical measurements of coefficients α and σ_s in dB depicted in Table 4.2 shows the following values for a number of indoor and outdoor wave propagation cases [80,97]:

4.4. LIvE REM Construction Technique

REM construction techniques mainly focus on the estimating the received power for unmeasured locations via utilizing measurements. In our proposed technique, we also assume that we have the exact knowledge about α and P_{L_0}, which can be estimated effectively via MCDs by signaling periodically.

LIvE REM construction consists of two main parts; the first part focuses on the PU location and power estimation via using the channel properties and the measurements from MCDs, while the second part uses the information from the first part and constructs the REM.

4.4.1. Location Estimation Method

First, we estimate the primary user location and the transmit power. After appropriate averaging, the channel disturbance is minimized. If there is no disturbance

in the measurements in [dB], then

$$10\alpha \log_{10} d_{(x_i,y_i)} \approx P^{tx} - P_{L_0} - \overline{P_i^{rx}} \ . \tag{4.6}$$

Hence,

$$\sqrt{(x_{pu} - x_i)^2 + (y_{pu} - y_i)^2} \approx 10^{\frac{P^{tx} - P_{L_0} - \overline{P_i^{rx}}}{10\alpha}} \ . \tag{4.7}$$

Therefore,

$$(x_{pu} - x_i)^2 + (y_{pu} - y_i)^2 \approx 10^{\frac{P^{tx} - P_{L_0} - \overline{P_i^{rx}}}{5\alpha}} \ ,$$

and,

$$(x_{pu}^2 - 2x_{pu}x_i + x_i^2) + (y_{pu}^2 - 2y_{pu}y_i + y_i^2) \approx 10^{\frac{P^{tx} - P_{L_0} - \overline{P_i^{rx}}}{5\alpha}} \ .$$

Which results in,

$$x_i^2 + y_i^2 \approx 2x_{pu}x_i + 2y_{pu}y_i + 10^{\frac{P^{tx} - P_{L_0} - \overline{P_i^{rx}}}{5\alpha}} - R^2 \tag{4.8}$$

where $R^2 = x_{pu}^2 + y_{pu}^2$, (x_i, y_i) and (x_{pu}, y_{pu}) are the locations of ith MCD and active primary user, respectively. Hence, we can express Equation 4.8 in the matrix form $A\theta \approx b$ where

$$A = \begin{pmatrix} 2x_1 & 2y_1 & 10^{\frac{-P_{L_0} - \overline{P_1^{rx}}}{5\alpha}} & -1 \\ 2x_2 & 2y_2 & 10^{\frac{-P_{L_0} - \overline{P_2^{rx}}}{5\alpha}} & -1 \\ . & . & . & . \\ 2x_{N_{MCD}} & 2y_{N_{MCD}} & 10^{\frac{-P_{L_0} - \overline{P_{N_{MCD}}^{rx}}}{5\alpha}} & -1 \end{pmatrix}$$

$$\theta = \begin{pmatrix} x_{pu} \\ y_{pu} \\ 10^{\frac{P^{tx}}{5\alpha}} \\ R^2 \end{pmatrix}, \quad b = \begin{pmatrix} x_1^2 + y_1^2 \\ x_2^2 + y_2^2 \\ . \\ x_{N_{MCD}}^2 + y_{N_{MCD}}^2 \end{pmatrix}$$

The estimation problem formulated in the matrix form can be solved easily using the least squares method [98, 99]. Let $\widehat{\theta}$ be estimated value for θ. Then, the solution is computed as

$$\begin{aligned} \widehat{\theta} &= \arg\min \|A\theta - b\| \\ &= (A^T A)^{-1} A^T \, b \end{aligned} \tag{4.9}$$

where $\|A\theta - b\| = (A\theta - b)^T (A\theta - b)$. Hence, we first estimate the location and the power of the active primary user by the help of the channel information.

Introducing the non-linear constraint $R^2 = x_{pu}^2 + y_{pu}^2$ results in

$$\begin{aligned} \widehat{\theta} &= \arg\min \|A\theta - b\| \\ s.t. \quad & r^T\theta + \theta^T P \theta = 0 \end{aligned} \tag{4.10}$$

where

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad r = \begin{pmatrix} 0 \\ 0 \\ 0 \\ -1 \end{pmatrix}$$

We solve the constrained optimization problem by transforming the original prob-

lem into Lagrange dual problem:

$$\mathcal{L}(\theta, \lambda) = (A\theta - b)^T (A\theta - b) + \lambda(r^T\theta + \theta^T P\theta) \tag{4.11}$$

where λ is the Lagrange multiplier. We differentiate the Lagrangian with respect to θ and λ and get

$$\frac{\partial \mathcal{L}}{\partial \theta} = 2\theta^T(A^T A + \lambda P) - 2b^T A + \lambda r^T \tag{4.12}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = r^T\theta + \theta^T P\theta \tag{4.13}$$

To satisfy the optimality condition, Equation 4.12 and 4.13 must be zero. Setting Equation 4.12 equal to zero leads to

$$\theta = (A^T A + \lambda P)^{-1}(A^T b - \frac{\lambda}{2} r) \tag{4.14}$$

Setting Equation 4.13 equal to zero and substituting θ found in Equation 4.14 results in a single unknown variable equation in terms of λ only, which can be easily solved by numerical methods. After finding λ, we solve the optimization problem in Equation 4.10 by solving the Lagrange dual problem in Equation 4.11 using the least squares method [100].

4.4.2. REM Construction Method

After solving for $\widehat{\theta}$, we have the estimated values for (x_{pu}, y_{pu}) and P^{tx}. We proceed with the REM construction method by utilizing these estimated values $(\widehat{x}_{pu}, \widehat{y}_{pu})$, $\widehat{P}^{tx}$, and the channel parameters (α and P_{L_0}). We evaluate the estimated power at any location in Equation 4.15 by using the estimated data and the channel information.

$$\overline{P^{rx}}(x, y)[dB] = \widehat{P}^{tx}[dB] - P_{L_0} - 10\alpha \log_{10} \widehat{d}_{(x,y)} \tag{4.15}$$

where $\widehat{d}_{(x,y)} = \sqrt{(x - \widehat{x}_{pu})^2 + (y - \widehat{y}_{pu})^2}$.

The estimated received power is evaluated for a specific grid resolution on the operation area and persisted to the database. As the time progresses, this operation is repeated, the raw data and the estimated interference data are persisted to the database. Recording this data with advancing time provides a vital information and enables the cognition in the network. Different types of queries can be run on the persisted data.

4.4.3. Comparison of REM Construction Techniques

Spatio-temporal interpolation approaches require the evaluation of weights at each point of interest. A popular and widely used interpolation approach is Kriging. Kriging has the advantage that it provides an unbiased estimate with the least error variance. However, Kriging based weight evaluation have a cubic time complexity in terms of N_s for each point of interest (N_g), thus resulting in a time complexity of $O(N_s^3 N_g)$ [101], where N_g is the number of grid points.

Similarly the LIvE REM construction technique involves with the matrix inversion of size $N_s \times N_s$ which gives us a time complexity of $O(N_s^3 N_g)$ without any optimizations. If we utilize the Coppersmith-Winograd algorithm, we have $O(N_s^{2.376} N_g)$ [102].

Table 4.3. Elapsed time for REM construction techniques ($N_g = 10000$).

REM Construction Technique	Duration (s)
IDW1 (Inverse Distance Weighted Interpolation)	0.1146
IDW2 (Inverse Distance Square Weighted Interpolation)	0.0462
Kriging	0.1020
LIvE REM	0.0656

In Table 4.3 (for $N_g = 10000$), the measured elapsed time for IDW1, IDW2, Kriging, and LIvE REM construction techniques are listed. Simulations are run on a laptop

Table 4.4. REM construction techniques comparison.

Information Parameter	**IDW**	**Kriging**	**LIvE REM**
N_s requirement	Low	High	Low
Update frequency	Real time	Real time	Real time
Channel knowledge	No	No	Yes
Interpolation method	Multiple linear regression/ Distance based	Variogram/ Distance based	Location estimation based
Processing requirements	Low	High	Moderate
Precision	Low	Moderate	High

machine with Core2Duo 2.2GHz CPU and 3GB RAM. IDW2 is the inverse distance square weighted interpolation technique and its time complexity is the smallest among the measured techniques. Our proposed method has the second smallest duration and its performance is better than others.

In Table 4.4, (adapted from [80]), we compare the characteristics of the REM construction techniques. We compare IDW, Kriging, and LIvE REM construction techniques. The comparison matrix clearly suggests us using the LIvE REM construction technique with its moderate processing requirements, high precision, and low N_s requirement. Utilizing the channel knowledge (which is possible to acquire with measuring and communicating between MCDs periodically) comes with these performance improvements. Results Section elaborates the quality of REM constructed via the LIvE REM construction technique.

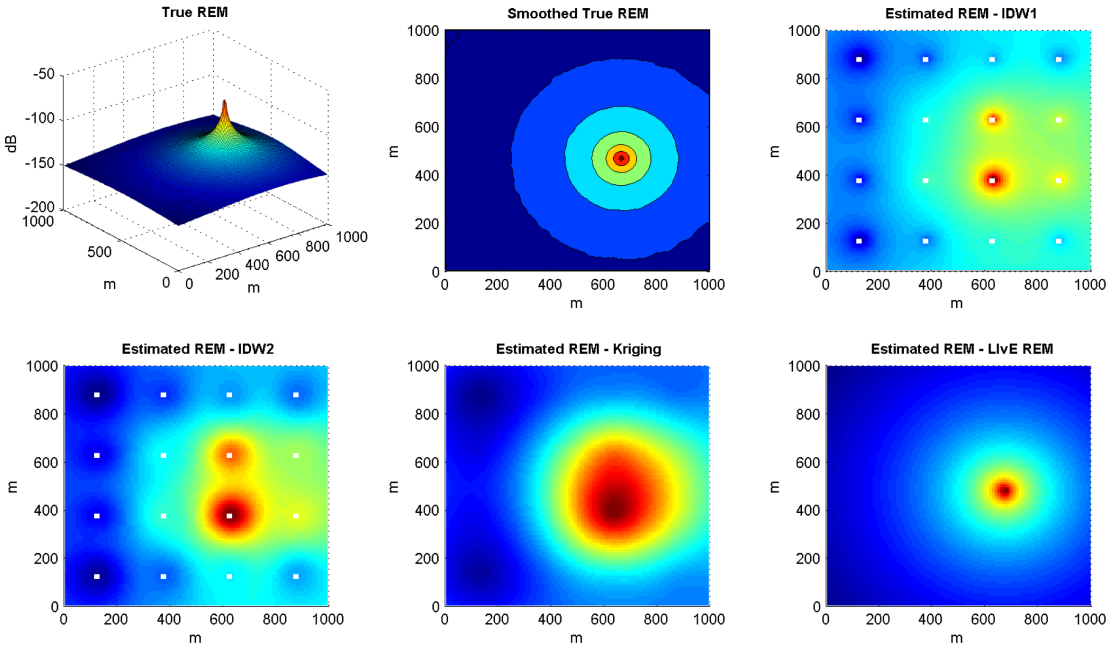


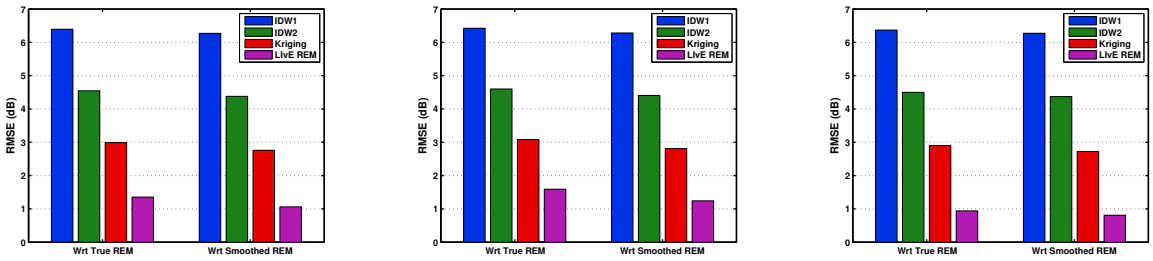
Figure 4.4. Generated REMs under log-normal shadowing ($N_m = 100$).

4.5. Results

We consider an area of 1000m$\times$1000m divided into subregions. In each subregion, an MCD at the center measures the average RSS and reports it to the REM manager. The path loss exponent, the path loss correction, and σ_s are selected as 3.5, 38.4, and 8, respectively for an urban area [80]. IDW1 and IDW2 use inverse distance and inverse distance square weighted interpolation techniques.

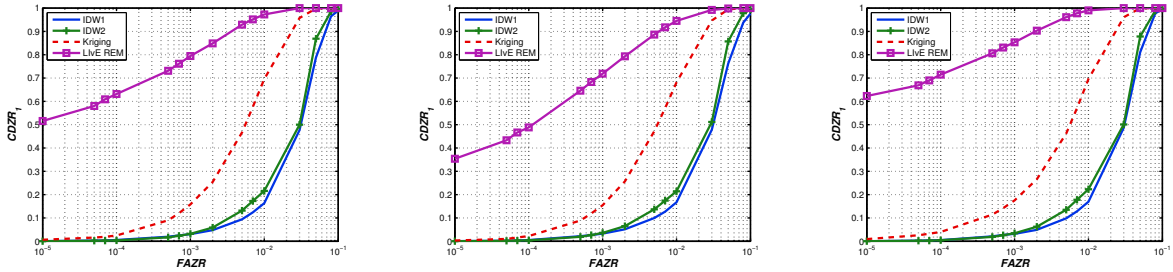
4.5.1. Effect of Channel Conditions

In this subsection, we consider 16 MCDs (4×4 subregions with edge size 250m $\times$ 250m) in an area with log-normal shadowing and shadowing/fading channel conditions and compare the methods in terms of RMSE and ROC curves. In Figure 4.4, generated REMs with different techniques under log-normal shadowing channel are presented. Estimated REMs via Kriging and LIvE REM construction technique are much more similar to the true REM.



(a) $N_m = 100$, shadowing (b) $N_m = 100$, shadowing/fading (c) $N_m = 300$, shadowing/fading

Figure 4.5. RMSE values for REM construction techniques.



(a) $N_m = 100$, shadowing (b) $N_m = 100$, shadowing/fading (c) $N_m = 300$, shadowing/fading

Figure 4.6. ROC curves of REM construction techniques.

For analyzing the performance under different channels, we consider log-normal shadowing and combined shadowing/fading (log-normal shadowing plus Rayleigh fading) channel conditions. First, we compare the methods in terms of RMSE with respect to the true REM and the smoothed REM.

In all three conditions, the proposed method achieves better performance than other methods since RMSE shows the deviation of the constructed REM from the true REM. In Figure 4.5a and 4.5b, RMSE values of the REM construction methods are depicted under log-normal shadowing and log-normal shadowing plus Rayleigh fading channel conditions for $N_m = 100$, respectively. Under a shadowing/fading channel, Kriging and LIvE REM construction methods are affected adversely. However, increasing N_m (averaging sample size) decreases the adverse results by eliminating the fading effects as shown in Figure 4.5c.

More useful performance metrics such as CDZR_1 and FAZR provide a better un-

derstanding of the network. In Figure 4.3, zones for the true REM and the estimated REM are depicted. Better REM construction methods fit better on the true REM contour with a predetermined threshold value (-120 dB in our case). In Figure 4.6, FAZR versus $CDZR_1$ is depicted for understanding the achievable $CDZR_1$ for a given FAZR. Having higher $CDZR_1$ means disturbing the primary user is less probable. Having less FAZR means that the white spaces can be utilized better.

In Figure 4.6, LIvE REM outperforms other methods with the help of the channel knowledge which is a reasonable assumption. Knowing the channel parameters of the operating area such as the path loss exponent and the correction enables us to estimate the primary user location and the transmit power. Therefore, constructing the REM becomes easier with the help of the estimated data.

In Figure 4.6a and 4.6b, ROC curves of the REM construction methods are depicted under log-normal shadowing and log-normal shadowing plus Rayleigh fading channel conditions, respectively. Under a shadowing/fading channel, Kriging and LIvE REM construction methods are affected adversely as in the RMSE case. In Figure 4.6c for FAZR=0.002, we have $CDZR_1$ values equal to 0.9, 0.026 and below for LIvE, Kriging, IDW1 and IDW2 REM construction methods, respectively. The LIvE REM construction technique clearly outperforms the other methods. If we have $CDZR_1$=0.9 constraint, then achievable FAZR values are 0.002, 0.021, 0.062, and 0.065 for LIvE, Kriging, IDW1 and IDW2, respectively. Having less FAZR provides us better utilization of the white spaces which is in the case of LIvE REM construction technique.

4.5.2. Effect of N_m on $CDZR_1$

Increasing the sample size for averaging RSS values mitigates the adverse effects of the fading channel. Hence, it provides better statistics for estimating the PU location and the transmit power. If the channel conditions are harsher, then increasing the N_m value improves the performance of the LIvE REM construction method.

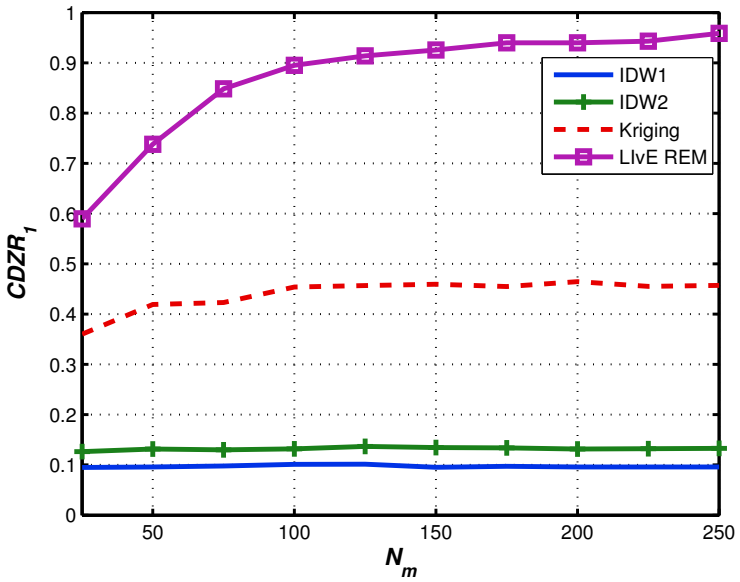


Figure 4.7. N_m versus CDZR_1 (FAZR=0.005, shadowing/fading).

In Figure 4.7, N_m versus CDZR_1 is depicted with FAZR= 0.005 constraint and under shadowing/fading channel conditions. Under these conditions, the LIvE REM construction method outperforms the others. Only the LIvE REM construction method can achieve acceptable CDZR_1 values. Increasing N_m (up to $N_m = 100$) improves the results significantly for the Kriging and LIvE REM construction methods. Under shadowing/fading, for LIvE REM method having $N_m > 200$ does not improve the system performance significantly. Hence, the appropriate averaging sample size for RSS can be close to 200 in this setting.

4.5.3. Effect of Distance to Closest MCD on CDZR_1

In Figure 4.8, d_{min} versus CDZR_1 is depicted where d_{min} is the distance between PU and the closest MCD. Some interpolation techniques heavily depend on the location of PU. If PU is close to any one of the MCDs, the resulting REM is very close to the actual REM. However, this condition is not necessarily satisfied in real life. The LIvE REM construction method does not depend on the location of the PU or closeness to any one of the MCDs as in the case of other interpolation techniques. The LIvE REM construction technique first estimates the PU location and the transmit power

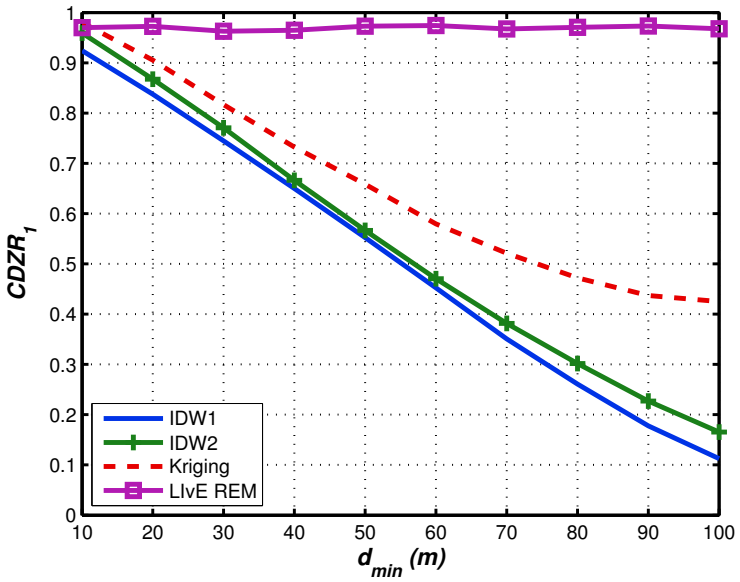


Figure 4.8. d_{min} versus CDZR_1 ($N_m = 300$, FAZR=0.005, shadowing/fading).

to remove the dependency.

Increasing d_{min} severely affects the performance of the Kriging, IDW1, and IDW2. The LIvE REM construction technique is not affected by the d_{min} parameter. Therefore, LIvE REM is a robust and efficient REM construction technique that outperforms Kriging, IDW1, and IDW2.

4.6. Chapter Summary

REMs act as cognition engines by building long-term knowledge via processing spectrum measurements collected from sensors to estimate the spectral state of locations which do not have any measurement data. The LIvE REM construction technique is proposed and compared with the well-known REM construction techniques in the literature such as Kriging and inverse distance weighted interpolation under shadowing and multipath fading channels. The simulation results suggest that the location estimation based REM construction outperforms the compared methods in terms of RMSE and CDZR_1 by utilizing additional information of channel parameters which is a reasonable assumption. We analyze the effect of N_m and d_{min} on CDZR_1 and conclude

that increasing N_m mitigates the fading effects and improves the performance of LIvE REM significantly. Analysis on the effect of d_{min} focuses on how the algorithms are affected by the location or closeness of PU to any MCD. If PU is not close enough to any of the MCDs, the achievable CDZR_1 for Kriging for a given FAZR constraint is not reasonable ($d_{min} = 50$ implies $\text{CDZR}_1 \approx 0.65$). However, LIvE REM is not affected by d_{min}. It always achieves CDZR_1 higher than 0.9 since it first estimates the location and the transmit power of PU.

For the future directions, we plan to analyze the LIvE REM performance in heterogeneous environments by utilizing mixed path loss formulations.

5. CONTRIBUTIONS AND FUTURE DIRECTIONS

This chapter summarizes the contributions of this thesis to the research on CRNs and presents several promising research directions that can utilize our findings and proposed solutions.

5.1. Summary of Contributions

Main contributions of this thesis can be summarized as follows:

(i) *Cooperative detection strategy via uniform quantization:* This thesis proposes a cooperative sensing fusion strategy that uses uniform quantization for local sensing. Proposed method generalizes the fusion strategy for methods that are using quantization.

Most of the works on quantization focus on just the local quantization process. On the contrary, we optimize the overall performance in terms of cooperative detection probability by optimizing thresholds and the fusion function weights. Numerical results demonstrate that under imperfect reporting channel and false reports, the proposed method performs better than hard decision algorithms such as Majority and M-of-N in terms of probability of detection and false alarm at the cost of a marginal increase in overhead bits. We also compare performance of our proposal to that of EGC scheme, which performs significantly better than hard decision algorithms and is an upper bound when the nodes are not differentiated. Our algorithm performs close to EGC in terms of probability of detection.

(ii) *Radio environment map construction technique:* Being aware of the surrounding RF environment and state is a challenging problem in CR. Due to the challenges of internal sensing, external sensing recently has gained noticeable interest. Our work highlights this phenomenon, we propose new useful quality metrics, and propose a novel REM construction technique that outperforms the conventional techniques in shadow and multipath fading channels. Simulation results suggest

that the LIvE REM construction outperforms the compared methods in terms of RMSE and $CDZR_1$. LIvE REM construction technique's time complexity is better than Kriging and inverse distance weighted interpolation, which are commonly used techniques in the literature.

5.2. Future Directions

As the future work in internal sensing, we plan to address the multi-hop performance analysis for reporting the quantized measurements (multi-hop $\mathbf{P_D}$ and $\mathbf{P_F}$ formulations are important in modeling and evaluating the performance of detection in ad-hoc CRNs) and develop a hybrid sensing method that utilizes internal and external sensing information. We expect to achieve better sensing performance by utilizing both internal and external sensing and merge the accumulated cognition.

Regarding new quality metrics and the REM construction technique, we plan to analyze the performance of the LIvE REM construction technique in heterogeneous environments by utilizing mixed path loss formulations. Since the CRN operation area may not be homogeneous in terms of channel properties, we need to analyze the performance under more realistic heterogeneous environments. We also plan to work on new metrics related to the capacity (maximum detectable opportunity) and analyze the capacity achievable by REM construction techniques.

REFERENCES

1. Patil, K., R. Prasad and K. Skouby, "A Survey of Worldwide Spectrum Occupancy Measurement Campaigns for Cognitive Radio", *Proceeding of the International Conference on Devices and Communications (ICDeCom)*, pp. 1–5, February 2011.

2. McHenry, M. A., *NSF Spectrum Occupancy Measurements Project Summary*, Technical Report, Shared Spectrum Company, 2005.

3. Mitola III, J. and G. Maguire Jr, "Cognitive Radio: Making Software Radios More Personal", *IEEE Personal Communications*, Vol. 6, No. 4, pp. 13–18, 1999.

4. Haykin, S., "Cognitive Radio: Brain-empowered Wireless Communications", *IEEE Journal on Selected Areas in Communications*, Vol. 23, No. 2, pp. 201–220, IEEE, 2005.

5. Mitola III, J., "Software Radios: Survey, Critical Evaluation and Future Directions", *IEEE Aerospace and Electronic Systems Magazine*, Vol. 8, No. 4, pp. 25–36, IEEE, 1993.

6. FCC, *Spectrum Poicy Take Force (SPTF) Report ET Docket No: 02-135*, Technical Report, November 2002.

7. Mitola, J., *Cognitive Radio: An Integrated Agent Architecture for Software Defined Radio*, Ph.D. Thesis, Royal Institute of Technology (KTH), 2000.

8. Akyildiz, I., W. Lee, M. Vuran and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey", *Elsevier Computer Networks*, Vol. 50, No. 13, pp. 2127–2159, Elsevier, 2006.

9. Arslan, H., *Cognitive Radio, Software Defined Radio, and Adaptive Wireless Systems*, Springer, Dordrecht, 2007.

10. Herath, S., N. Rajatheva and C. Tellambura, "Energy Detection of Unknown Signals in Fading and Diversity Reception", *IEEE Transactions on Communications*, Vol. 59, No. 9, pp. 2443–2453, 2011.

11. Digham, F. F., M. S. Alouini and M. K. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels", *Proceedings of the International Conference on Communications (ICC)*, Vol. 5, pp. 3575–3579, 2003.

12. Cabric, D., S. Mishra and R. Brodersen, "Implementation Issues in Spectrum Sensing for Cognitive Radios", *Conference Record of the 38th Asilomar Conference on Signals, Systems and Computers*, Vol. 1, pp. 772–776, 2004.

13. Ganesan, G. and Y. Li, "Cooperative Spectrum Sensing in Cognitive Radio Networks", *Proceedings of the First International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 137–143, 2005.

14. Ganesan, G. and Y. Li, "Agility Improvement Through Cooperative Diversity in Cognitive Radio", *Proceedings of the Global Telecommunications Conference (GLOBECOM)*, Vol. 5, pp. 2505–2509, 2005.

15. Tandra, R. and A. Sahai, "Fundamental Limits on Detection in Low SNR Under Noise Uncertainty", *Proceedings of the International Conference on Wireless Networks, Communications and Mobile Computing (WirelessCom)*, Vol. 1, pp. 464–469, 2005.

16. Shankar, N., C. Cordeiro and K. Challapali, "Spectrum Agile Radios: Utilization and Sensing Architectures", *Proceedings of the First International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 160–169, 2005.

17. Cabric, D., A. Tkachenko and R. Brodersen, "Spectrum Sensing Measurements of Pilot, Energy, and Collaborative Detection", *Proceedings of the Military Communications Conference (MILCOM)*, pp. 1–7, 2006.

18. Qihang, P., Z. Kun, W. Jun and L. Shaoqian, "A Distributed Spectrum Sensing Scheme Based on Credibility and Evidence Theory in Cognitive Radio Context", *Proceedings of the 17th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–5, 2006.

19. Geirhofer, S., L. Tong and B. Sadler, "A Measurement-based Model for Dynamic Spectrum Access in WLAN Channels", *Proceedings of the Military Communications Conference (MILCOM)*, pp. 1–7, 2006.

20. Pawełczak, P., C. Guo, R. V. Prasad and R. Hekmat, *IRCTR-S-004-07: Cluster-based Spectrum Sensing Architecture for Opportunistic Spectrum Access Networks*, Technical Report, International Research Centre for Telecommunications and Radar, 2006.

21. Yuan, Y., P. Bahl, R. Chandra, P. A. Chou, J. I. Ferrell, T. Moscibroda, S. Narlanka and Y. Wu, "KNOWS: Cognitive Radio Networks Over White Spaces", *Proceedings of the Second International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 416–427, 2007.

22. Digham, F. F., M. S. Alouini and M. K. Simon, "On the Energy Detection of Unknown Signals Over Fading Channels", *IEEE Transactions on Communications*, Vol. 55, No. 1, pp. 21–24, January 2007.

23. Urkowitz, H., "Energy Detection of Unknown Deterministic Signals", *Proceedings of the IEEE*, Vol. 55, No. 4, pp. 523–531, 1967.

24. Hashemi, H., "The Indoor Radio Propagation Channel", *Proceedings of the IEEE*, Vol. 81, No. 7, pp. 943–968, IEEE, 1993.

25. Atapattu, S., C. Tellambura and H. Jiang, "Energy Detection Based Cooperative Spectrum Sensing in Cognitive Radio Networks", *IEEE Transactions on Wireless Communications*, Vol. 10, No. 4, pp. 1232–1241, IEEE, 2011.

26. Unnikrishnan, J. and V. Veeravalli, "Cooperative Sensing for Primary Detection in Cognitive Radio", *IEEE Journal of Selected Topics in Signal Processing*, Vol. 2, No. 1, pp. 18–27, IEEE, 2008.

27. Zhang, W., R. Mallik and K. Letaief, "Optimization of Cooperative Spectrum Sensing with Energy Detection in Cognitive Radio Networks", *IEEE Transactions on Wireless Communications*, Vol. 8, No. 12, pp. 5761–5766, IEEE, 2009.

28. Lee, S. H., D. C. Oh and Y. H. Lee, "Hard Decision Combining-based Cooperative Spectrum Sensing in Cognitive Radio Systems", *Proceedings of the International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, pp. 906–910, 2009.

29. Leu, A. E., M. McHenry and B. L. Mark, "Modeling and Analysis of Interference in Listen-Before-Talk Spectrum Access Schemes", *Wiley International Journal of Network Management*, Vol. 16, No. 2, pp. 131–147, 2006.

30. Visotsky, E., S. Kuffner and R. Peterson, "On Collaborative Detection of TV Transmissions in Support of Dynamic Spectrum Sharing", *Proceedings of the First International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 338–345, 2005.

31. Weiss, T. A., "A Diversity Approach for the Detection of Idle Spectral Resources in Spectrum Pooling Systems", *Proceedings of the 48th International Scientific Colloquium*, pp. 37–38, September 2003.

32. Mishra, S. M., A. Sahai and R. W. Brodersen, "Cooperative Sensing Among Cognitive Radios", *Proceedings of the International Conference on Communications (ICC)*, Vol. 4, pp. 1658–1663, 2006.

33. Chair, Z. and P. K. Varshney, "Optimal Data Fusion in Multiple Sensor Detection Systems", *IEEE Transactions on Aerospace and Electronic Systems*, Vol. AES-22, No. 1, pp. 98–101, IEEE, 1986.

34. Gandetto, M. and C. S. Regazzoni, "Spectrum Sensing: A Distributed Approach for Cognitive Terminals", *IEEE Journal on Selected Areas in Communications*, Vol. 25, No. 3, pp. 546–557, IEEE, 2007.

35. Gandetto, M., A. F. Cattoni and C. S. Regazzoni, "A Distributed Approach to Mode Identification and Spectrum Monitoring for Cognitive Radios", *Proceedings of the SDR Forum Technical Conference*, November 2005.

36. Cattoni, A. F., I. Minetti, M. Gandetto, R. Niu, P. K. Varshney and C. S. Regazzoni, "A Spectrum Sensing Algorithm Based on Distributed Cognitive Models", *Proceedings of the SDR Forum Technical Conference*, 2006.

37. Peh, E. and Y. C. Liang, "Optimization for Cooperative Sensing in Cognitive Radio Networks", *Proceedings of the Wireless Communications and Networking Conference (WCNC)*, pp. 27–32, 2007.

38. Akyildiz, I. F., B. F. Lo and R. Balakrishnan, "Cooperative Spectrum Sensing in Cognitive Radio Networks: A Survey", *Elsevier Physical Communication*, Vol. 4, pp. 40–62, Elsevier, 2011.

39. Zou, Q., S. Zheng and A. H. Sayed, "Cooperative Spectrum Sensing via Coherence Detection", *Proceedings of the 15th Workshop on Statistical Signal Processing (SSP)*, pp. 610–613, 2009.

40. Shen, J., T. Jiang, S. Liu and Z. Zhang, "Maximum Channel Throughput via Cooperative Spectrum Sensing in Cognitive Radio Networks", *IEEE Transactions on Wireless Communications*, Vol. 8, No. 10, pp. 5166–5175, IEEE, 2009.

41. Huang, H., Z. Zhang, P. Cheng, A. Huang and P. Qiu, "Cooperative Spectrum Sensing in Cognitive Radio Systems with Limited Sensing Ability", *Springer Journal of Zhejiang University-Science C*, Vol. 11, No. 3, pp. 175–186, Springer, 2010.

42. Oh, D. C. and Y. H. Lee, "Cooperative Spectrum Sensing with Imperfect Feed-

back Channel in the Cognitive Radio Systems", *Wiley International Journal of Communication Systems*, Vol. 23, No. 6-7, pp. 763–779, Wiley Online Library, 2010.

43. Ghasemi, A. and E. S. Sousa, "Opportunistic Spectrum Access in Fading Channels Through Collaborative Sensing", *Wiley Journal of Communications*, Vol. 2, No. 2, pp. 71–82, 2007.

44. Kaligineedi, P. and V. K. Bhargava, "Sensor Allocation and Quantization Schemes for Multi-band Cognitive Radio Cooperative Sensing System", *IEEE Transactions on Wireless Communications*, Vol. 10, No. 1, pp. 284–293, IEEE, 2011.

45. Zhang, X., Q. Wu and J. Wang, "Optimization of Sensing Time in Multichannel Sequential Sensing for Cognitive Radio", *Wiley International Journal of Communication Systems*, Wiley Online Library, `doi:10.1002/dac.1341`, 2011.

46. Chen, H. and H. H. Chen, "Spectrum Sensing Scheduling for Group Spectrum Sharing in Cognitive Radio Networks", *Wiley International Journal of Communication Systems*, Vol. 24, No. 1, pp. 62–74, Wiley Online Library, 2011.

47. Sun, C., W. Zhang and K. B. Letaief, "Cooperative Spectrum Sensing for Cognitive Radios Under Bandwidth Constraints", *Proceedings of the Wireless Communications and Networking Conference (WCNC)*, pp. 1–5, 2007.

48. Picinbono, B. and P. Duvaut, "Optimum Quantization for Detection", *IEEE Transactions on Communications*, Vol. 36, No. 11, pp. 1254–1258, IEEE, 1988.

49. Nguyen-Thanh, N. and I. Koo, "Evidence Theory Based Cooperative Spectrum Sensing with Efficient Quantization Method in Cognitive Radio", *IEEE Transactions on Vehicular Technology*, Vol. 60, No. 1, pp. 185–195, IEEE, 2011.

50. Sakran, H., M. Shokair, E. S. El-Rabaie and A. A. El-Azm, "Three Bits Softened Decision Scheme in Cooperative Spectrum Sensing Among Cognitive Radio Net-

works", *Proceedings of the 28th National Radio Science Conference (NRSC)*, pp. 1–9, 2011.

51. Sakran, H. and M. Shokair, "Hard and Softened Combination for Cooperative Spectrum Sensing Over Imperfect Channels in Cognitive Radio Networks", *Springer Telecommunication Systems*, pp. 1–11, Springer, 2011.

52. Chaudhari, S., J. Lundén and V. Koivunen, "BEP Walls for Collaborative Spectrum Sensing", *Proceedings of the International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2984–2987, 2011.

53. Nekovee, M., "Cognitive Radio Access to TV White Spaces: Spectrum Opportunities, Commercial Applications and Remaining Technology Challenges", *Proceedings of the International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 1–10, 2010.

54. Morgado, A. and N. B. Carvalho, "White Spaces Communications in Europe", *Proceedings of the 30th General Assembly and Scientific Symposium (URSI)*, pp. 1–4, 2011.

55. IEEE, *Singapore TVWS Trial*, `https://mentor.ieee.org/802.22/dcn/11/22-11-0138-00-rasg-singapore-tvws-trial-publication.pdf`, accessed at January 2012.

56. UBAK, *Türkiye'de DVB-T*, `http://www.ubak.gov.tr/BLSM_WIYS/HGB/tr/Sag_Menu/20100816_162441_10472_1_64.html`, accessed at May 2011.

57. FCC, *Public Notice, ET Docket No:04-186. Office of Engineering and Technology Announces the Approval of Spectrum Bridge, Incorporated's TV Bands Database System for Operation*, Technical Report, December 2011.

58. Harrison, K., S. M. Mishra and A. Sahai, "How Much White-space Capacity is There?", *Proceedings of the International Symposium on New Frontiers in Dy-*

namic Spectrum Access Networks (DySPAN), pp. 1–10, 2010.

59. Contreras, S., G. Villardi, R. Funada and H. Harada, "An Investigation into the Spectrum Occupancy in Japan in the Context of TV White Space Systems", *Proceedings of the 6th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pp. 341–345, 2011.

60. CEPT, *Electronic Communications Committee (ECC), SE 43 Working Group on Cognitive Radio Systems*, `http://www.cept.org/ecc/groups/ecc/wg-se/se-43`, accessed at February 2012.

61. *IEEE 802.22 WG on WRANs*, `http://www.ieee802.org/22/`, accessed at December 2011.

62. *IEEE 802.22-2011 Standard for Wireless Regional Area Networks in TV Whitespaces Completed*, `http://standards.ieee.org/news/2011/802.22.html`, accessed at July 2011.

63. *Cognitive Networking Alliance (CogNeA)*, `http://www.cognea.org`, accessed at May 2010.

64. ECMA, *Ecma 392: MAC and PHY for Operation in TV White Space, Ecma International Standard, 1st Edition*, `http://www.ecma-international.org/publications/standards/Ecma-392.htm`, accessed at December 2009.

65. Bahl, P., R. Chandra, T. Moscibroda, R. Murty and M. Welsh, "White Space Networking with Wi-Fi Like Connectivity", *Newsletter of the SIGCOMM Computer Communication Review*, Vol. 39, pp. 27–38, 2009.

66. Rahman, M. A., C. Song and H. Harada, "Development of a TV White Space Cognitive Radio Prototype and its Spectrum Sensing Performance", *Proceedings of the 6th International ICST Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, pp. 231–235, 2011.

67. Keybridge, *Proposal to Administer a TV Bands Database*, Technical Report, Key Bridge Global LLC, January 2010.

68. IETF, *Protocol to Access White Space Database PAWS, Problem Statement*, `https://datatracker.ietf.org/doc/draft-patil-paws-problem-stmt/`, accessed at July 2011.

69. TVfool, *TVfool Web Site*, `http://www.tvfool.com/`, accessed at March 2011.

70. SpectrumBridge, *Spectrum Bridge Whitespaces*, accessed at May 2011.

71. UKFreeTV, *Independent Free Digital TV Advice*, `http://ukfree.tv`, accessed at May 2011.

72. Microsoft, *Microsoft Research WhiteFiService*, `http://whitespaces.msresearch.us`, accessed at May 2011.

73. FCC, *TVWS Database Tests*, `http://transition.fcc.gov/oet/whitespace/guides/TVWS_Database_Tests.pdf`, accessed at January 2012.

74. Sum, C. S., H. Harada, F. Kojima, Z. Lan and R. Funada, "Smart Utility Networks in TV White Space", *IEEE Communications Magazine*, Vol. 49, No. 7, pp. 132–139, IEEE, 2011.

75. SpectrumBridge, *The Future is Now – Nation's First Smart Grid TV White Space Network Trial*, `http://www.spectrumbridge.com/WhiteSpacesSolutions/success-stories/plumas.aspx`, accessed at February 2012.

76. Cordeiro, C., K. Challapali, D. Birru and N. Sai Shankar, "IEEE 802.22: The First Worldwide Wireless Standard Based on Cognitive Radios", *Proceedings of the First International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 328–337, 2005.

77. Kawade, S. and M. Nekovee, "Can Cognitive Radio Access to TV White Spaces Support Future Home Networks?", *Proceedings of the International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, pp. 1–8, 2010.

78. Zhao, Y., B. Le and J. H. Reed, "Network Support–The Radio Environment Map", *Elsevier Cognitive Radio Technology*, pp. 337–363, Elsevier, 2006.

79. Subramani, S., J. Riihijarvi, B. Sayrac, L. Gavrilovska, M. Sooriyabandara, T. Farnham and P. Mahonen, "Towards Practical REM-based Radio Resource Management", *Proceedings of the Future Network and Mobile Summit (FutureNetw)*, pp. 1–8, 2011.

80. Dagres, I., A. Polydoros, J. Riihijärvi, J. Nasreddine, P. Mähönen, L. Gavrilovska, V. Atanasovski, J. van de Beek, B. Sayrac, S. Grimoud, M. L. Benitez, J. P. Romero, R. Agusti and F. Casadevall, *Flexible and Spectrum-Aware Radio Access Through Measurements and Modelling in Cognitive Radio Systems FARAMIR, D4.1 Radio Environmental Maps: Information Models and Reference Model*, Technical Report, April 2011.

81. Riihijärvi, J., P. Mähönen and S. Sajjad, "Influence of Transmitter Configurations on Spatial Statistics of Radio Environment Maps", *Proceedings of the 20th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 853–857, 2009.

82. Bolea, L., J. Pérez-Romero, R. Agustí and O. Sallent, "Context Discovery Mechanisms for Cognitive Radio", *Proceedings of the 73rd Vehicular Technology Conference (VTC)*, pp. 1–5, 2011.

83. Grimoud, S., S. Ben Jemaa, B. Sayrac and E. Moulines, "A REM Enabled Soft Frequency Reuse Scheme", *Proceedings of the GLOBECOM Workshops (GC Wkshps)*, pp. 819–823, 2010.

84. Krige, D. G., *A Statistical Approach to Some Mine Valuations and Allied Problems at the Witwatersrand*, M.S. Thesis, University of Witwatersrand, 1951.

85. Phillips, C., M. Ton, D. Sicker and D. Grunwald, "Practical Radio Environment Mapping with Geostatistics", *Proceedings of the Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN)*, 2012.

86. Cressie, N., "Statistics for Spatial Data", *Wiley Terra Nova*, Vol. 4, No. 5, pp. 613–617, Wiley Online Library, 1992.

87. Wackernagel, H., *Multivariate Geostatistics: An Introduction with Applications*, Springer Verlag, Berlin, 2003.

88. Chiles, J. P. and P. Delfiner, *Geostatistics: Modeling Spatial Uncertainty*, Wiley, New York, 2012.

89. Tonkin, M. J. and S. P. Larson, "Kriging Water Levels with a Regional-Linear and Point-Logarithmic Drift", *Wiley Ground Water*, Vol. 40, No. 2, pp. 185–193, Wiley Online Library, 2002.

90. Bertsekas, D. P., *Nonlinear Programming*, Athena Scientific, Massachusetts, 1999.

91. Boyd, S. P. and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, New York, 2004.

92. Goldberg, D. E., *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley, New York, 1989.

93. Yilmaz, H. B., T. Tugcu and F. Alagoz, "Uniform Quantizer for Cooperative Sensing in Cognitive Radio Networks", *Proceedings of the 21st International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 548–553, 2010.

94. Chaudhari, S., J. Lundén, V. Koivunen and H. Poor, "Cooperative Sensing with Imperfect Reporting Channels: Hard Decisions or Soft Decisions?", *IEEE Transactions on Signal Processing*, Vol. 60, No. 1, pp. 18–28, IEEE, 2010.

95. Murty, R., R. Chandra, T. Moscibroda and P. V. Bahl, "SenseLess: A Database-Driven White Spaces Network", *IEEE Transactions on Mobile Computing*, Vol. 11, No. 2, pp. 189–203, IEEE Computer Society, 2012.

96. Wirastuti, N. and N. P. Sastra, "Application of the Suzuki Distribution to Simulations of Shadowing/Fading Effects in Mobile Communication", *Proceedings of the Forth International Conference on Information and Communication Technology and System*, 2008.

97. Rappaport, T. S., *Wireless Communications, Principles and Practices*, Prentice Hall, New Jersey, 2002.

98. Guvenc, I., S. Gezici and Z. Sahinoglu, "Fundamental Limits and Improved Algorithms for Linear Least-Squares Wireless Position Estimation", *Wiley Wireless Communications and Mobile Computing*, Wiley Online Library, `doi:10.1002/wcm.1029`, 2010.

99. Wang, J., R. V. Prasad, X. An and I. G. M. M. Niemegeers, "A Study on Wireless Sensor Network based Indoor Positioning Systems for Context-aware Applications", *Wiley Wireless Communications and Mobile Computing*, Vol. 12, No. 1, pp. 53–70, Wiley Online Library, 2012.

100. Kim, S., H. Jeon, H. Lee and J. Ma, "Robust Transmission Power and Position Estimation in Cognitive Radio", *Springer, Information Networking. Towards Ubiquitous Networking and Services*, pp. 719–728, Springer, 2008.

101. Srinivasan, B. V., R. Duraiswami and R. Murtugudde, "Efficient kriging for real-time spatio-temporal interpolation", *Proceedings of the 20th Conference on Probability and Statistics in the Atmospheric Sciences*, pp. 228–235, 2010.

102. Coppersmith, D. and S. Winograd, "Matrix Multiplication via Arithmetic Progressions", *Elsevier Journal of Symbolic Computation*, Vol. 9, No. 3, pp. 251–280, Elsevier, 1990.