Efficient and Robust Spectrum Management for Cognitive Radio Networks

by

Alexander W. Min

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2011

Doctoral Committee:

Professor Kang G. Shin, Chair
Professor Edwin Romeijn
Associate Professor Achilleas Anastasopoulos
Associate Professor Mingyan Liu

To Joung Won and Eric

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor, Professor Kang G. Shin, for his constant support and encouragement and invaluable advice and guidance during my doctoral study. I have benefited tremendously from his vision, technical insights, and insightful comments, and more importantly, his value of hard work and consistency in pursuit of excellence. I also thank Professors Edwin Romeijn, Achilleas Anastasopoulos, and Mingyan Liu for serving on my dissertation committee, and for their valuable feedback and suggestions.

I am also grateful to my many talented and friendly colleagues. First, I thank the RTCL members for their friendship and support, especially Eugene Chai, Katharine Chang, Karen Hou, Xin Hu, Fangjian Jin, Xiaoen Ju, Matt Knysz, Buyoung Yun, Yuanyuan Zeng, Caoxie Zhang, and others. I also thank my collaborators and former mentors, Dr. Kyu-Han Kim, Dr. Dhruv Gupta, Professor Prasant Mohapatra, Lingjie Duan, Professor Jianwei Huang, and Dr. Elizabeth Hildinger.

I thank the RTCL/SSL alumni, visitors and staff at the University of Michigan, including Zhigang Chen, Min-gyu Cho, Hyoil Kim, Ashwini Kumar, Jai-Jin Lim, Pradeep Padala, Wei Sun, Howard Tsai, Jisoo Yang, Jian Wu, Jaehyuk Choi, Young-June Choi, Hahnsang Kim, Ji-Hoon Yun, and Stephen Reger.

This thesis research has been supported in part by the NSF grants Grant No. CNS-0519498 and CNS-0721529, and by the AFOSR grant FA9550-10-1-0393.

Finally, I would like to express special gratitude to my wife, Joung Won Yoon, for each and every moment of her life with me in Ann Arbor, and my dear son, Eric Min, for all the great happiness and unexpected inspirations that he brought to my family. I am indebted to my sister and parents and parents-in-law for their love, encouragement, and support.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

APPENDIX

CHAPTER 1

Introduction

1.1 Cognitive Radio Networks

Due to the ever-increasing demand for new wireless services and applications for anytime and anywhere connectivity in our daily businesses and lives, we expect to face a shortage of available wireless spectrum in the near future. However, this spectrum-shortage problem is known to be rooted in the conventional *static* spectrum-allocation policy where only licensed devices can operate on a designated spectrum band. For example, according to a recent measurement report, an average of only $5.2\,\%$ of wireless spectrum under $3\,$GHz was being actively used, indicating that a large fraction of spectrum bands are un-/under-utilized at any given location and time [98].

Cognitive radios (CRs) have recently been proposed as an attractive means to mitigate the imminent spectrum-scarcity problem [72]. In cognitive radio networks (CRNs), secondary (unlicensed) users (SUs) can opportunistically access temporarily available licensed spectrum bands, i.e., spectrum bands not being used by the primary (licensed) users (PUs). Such opportunistic spectrum access is also referred to as *dynamic spectrum access* (DSA). As an initial step towards realizing this new concept of DSA, the Federal Communications Commission (FCC) has recently approved the operation of unlicensed CR devices in TV spectrum bands (a.k.a. TV white spaces) [49, 50]. This decision creates new opportunities for improving the efficiency of wireless spectrum resources without the restrictions of the conventional static spectrum access policy, thus facilitating the accommodation of

new wireless services and increased user demands.

Meanwhile, various standardization efforts are under development to utilize spectrum opportunities, such as IEEE 802.22 Wireless Regional Area Networks (WRANs) [1], 802.11af [2] and Ecma 392 [77]. For instance, the IEEE 802.22 WRANs [1] is the first standard using TV white space, and it aims to provide a last-mile wireless broadband access to rural areas.

Although DSA has been recognized as a promising solution to the spectrum-scarcity problem by significantly improving the wireless spectrum efficiency, its realization entails several technical and political challenges. In this thesis, we identify the two main challenges that hinder the realization of DSA. The first and most fundamental challenge is PUs' concern about potential interference from SUs' communications. If PUs allow unlicensed users (or SUs) to opportunistically access their licensed spectrum bands, they risk an increased level of interference caused by SUs' communications. Such interference can degrade the quality of PUs' service, which can easily lead to the loss of PUs' business. Therefore, SUs must provide efficient mechanisms to protect PUs' communications, such as spectrum sensing.

The second challenge is the lack of economic incentives to PUs for spectrum sharing. Even if SUs can almost perfectly avoid interfering with PUs' communication, PUs may still not be fully convinced to open up their spectrum resources to unlicensed users since there is no clear incentive for them to share their spectrum. Therefore, we need to provide efficient economic mechanisms, such as spectrum pricing strategies, that can provide incentives to both PUs and SUs by facilitating their interactions. Without resolving the above two main challenges, DSA will not widely accepted.

To address these challenges, we propose an efficient and robust spectrum-management framework that addresses several key CR-unique challenges, such as spectrum sensing, security, mobility and the dynamic spectrum market.

1.1.1 Efficient Detection of Primary Users

Accurate sensing of spectrum condition is key to the realization of DSA, so it can help mitigate the spectrum-scarcity problem. The main goal of spectrum sensing is to accurately detect, in real time, the presence/absence of primary signals on a spectrum band. However, achieving this goal is not easy since the FCC imposes very strict detection requirements to protect primary communications from potential interference from SUs. For example, in 802.22 WRANs, SUs must be able to detect a primary signal as weak as -20 dB within 2 seconds with high accuracy, i.e., both mis-detection and false-alarm probability less than 0.1 [38]. Unfortunately, this stringent performance requirement cannot be met with one-time sensing with a single sensor regardless of the underlying sensing technique, e.g., energy/feature detection [127, 131, 133].

In order to improve the detection performance, distributed (or cooperative) spectrum-sensing [55, 106] has recently emerged as a viable means to enhance the detection performance by exploiting sensor-location diversity. In cooperative sensing, the base station (BS) directs multiple cooperative sensors to perform spectrum sensing simultaneously and collects the sensing results (i.e., measured received primary signal strengths) to make a final decision as to the existence of a primary signal. Another approach to improve detection performance in the temporal domain is sensing scheduling. In sensing scheduling, the BS schedules spectrum sensing multiple times to exploit the temporal variations in received primary signal strengths. The BS can schedule sensings until it accumulates enough information to make a decision with high accuracy.

In CRNs, spectrum sensing must be able to detect heterogeneous types of PUs. For example, in 802.22, there are two types of PUs in TV white space: *large-scale*, e.g., TV, and *small-scale*, e.g., wireless microphone (WM) signals. In this thesis, we chose the terms *large-* and *small-scale* PUs based on the size of the spatial signal footprint, and different approaches are required for their detection. For the detection of large-scale PUs, cooperative sensors need to be chosen carefully by the BS since sensors may exhibit different detection performance based on their location and wireless conditions. Moreover, the BS needs to schedule spectrum sensing optimally to minimize sensing overhead and detection

latency. The detection of small-scale PUs is even more challenging due to their unpredictable spatial and temporal usage patterns and their small signal footprint. Therefore, knowing the PUs' characteristics, e.g., location and transmit-power level, is important for efficient opportunistic spectrum reuse.

1.1.2 Robust Detection of Primary Users

Spectrum sensing is vulnerable to attacks and device failures, such as primary user emulation [11, 26] or sensing report manipulation attacks [28]. These sensing-targeted attacks or malfunctioning sensors can severely undermine the detection of primary signals and spectrum white spaces because the fusion rule for a final decision on a PU's presence/absence relies solely on measurement results reported by the sensors, i.e., received signal strengths (RSSs). Therefore, sensing-targeted attacks can disable accurate spectrum sensing, the basic premise of DSA. In this thesis, we refer to these unique sensing-targeted attacks in CRNs as *sensing-disorder attacks*.

A sensing-disorder attack aims to obscure the existence of a primary signal or white space by manipulating the spectrum-sensing information (e.g., measured RSSs) either by raising or lowering the signal strengths that they report to the BS. When no primary signal exists, attackers can raise RSSs to generate an illusion of a primary signal. Otherwise, attackers can lower RSS to veil the presence of a primary signal. In both cases, attackers mislead the fusion center (i.e., base station, BS) to make an incorrect decision on the presence/absence of a primary signal, causing either waste of spectrum resources or unacceptable interference to the primary communications. For the detection of a small-scale primary signal, a sensing-disorder attack can significantly increase the error for the localization of a primary transmitter, resulting in the waste of spectrum opportunities in the space domain.

While sensing-disorder attacks can be easily launched with the aid of programmable software-defined radio (SDR) devices, their detection is difficult. Unlike the ordinary Denial-of-Service (DoS) attacks that exhaust all the network resources, they can be easily mounted by using SDR devices, such as USRP [3] and Sora [84]. These open-source

SDR platforms can be an attractive target for attackers because of their accessibility to low-layer protocol stacks like PHY and MAC. Detecting these attacks, however, is very difficult. While security mechanisms such as MAC-layer or crypto-based authentication work well in traditional wireless networks, lack of primary-secondary communications precludes their usage in CRNs. Moreover, detection difficulty is exacerbated by the volatile nature of the wireless medium itself, which makes it hard to differentiate between legitimate and manipulated sensing reports. Despite the grave consequence of these threats, they have been overlooked in the design of existing distributed spectrum-sensing schemes.

1.1.3 Mobile Cognitive Radios

The main goal of DSA is to allow CR-equipped SUs to safely coexist with PUs without disrupting PU communications. To achieve this goal, various aspects of DSA, such as spectrum sensing [88, 100, 104], spectrum sharing [102, 164], and security [103], have been studied extensively. Most existing efforts, however, focus on stationary CRNs, in which the location of both PUs and SUs are known to the BS in secondary systems, and thus, they may not be suitable when SUs are *mobile*. We envision that future mobile devices will incorporate CR-functionality and will be capable of dynamic and flexible spectrum access. Various standardization efforts for mobile CRs are being developed to utilize spectrum white spaces, such as 802.11af [2] and Ecma 392 [77].

Enabling DSA for mobile SUs entails new practical challenges. First, existing spectrum-availability models are derived based solely on PUs' temporal traffic statistics and might thus be unsuitable for CRNs with mobile CRs/SUs. Unlike in stationary CRNs (e.g., [1]), in which spectrum opportunity (or availability) is mostly affected by PUs' temporal channel usage patterns, in mobile CRNs, availability can also change as SUs move towards or away from PUs that are actively transmitting data. Second, protecting PUs from the SU mobility-induced interference is a challenging problem that calls for an efficient spectrum-sensing strategy tailored to mobile CRNs. Mobile SUs may need to sense spectrum more frequently to avoid interfering with PU communications. However, frequent spectrum sensing may not only incur significant time overhead [88], but also quickly drain the battery of

mobile CR devices due to the power-intensive nature of spectrum sensing [7, 68]. Third, mobile SUs will experience heterogeneous spectrum opportunities across space and time domains based on the geographical distribution of PUs and SUs' mobility patterns. The three challenges mentioned above are interrelated. Hence, to fully realize the benefits of DSA for mobile SUs, they must be considered jointly.

1.1.4 Dynamic Spectrum Market

The dynamic spectrum market (DSM) will play a key role in realizing DSA by facilitating spectrum trading between legacy spectrum owners and secondary consumers. This spectrum trading can be encouraged by a suitable pricing model through which DSM provides attractive economic incentives to legacy spectrum owners, and cost-effective spectrum access to secondary consumers. This will, in turn, enable more efficient and flexible usage of spectrum resources. Such a DSM already exists in various forms, such as mobile virtual network operators (MVNO) [4] and online spectrum markets (e.g., `specex.com` [5]).

A wide range of heterogeneous frequency bands will be available in the DSM considering the current trend of deregulating wireless resources. For example, the TV white space recently opened for unlicensed usage spans a wide range of frequencies over the VHF/UHF bands. Given this availability, it is natural for WSPs to want heterogeneous spectrum bands so as to avoid interference between them. Due to the difference in propagation profile (i.e., frequency-dependent attenuation rate), heterogeneous channels have different transmission and interference ranges, even with the same transmit power. Rational secondary consumers would be able to evaluate the value/utility of different channels and exploit the capability of their SDRs to access the different ranges of spectrum bands available in the market.

Another important but largely overlooked feature of DSM is the necessity of sharing leased spectrum bands with other SUs, which is a common feature of wireless communications. This feature has some implications in establishing the way market participants interact with each other. In a DSM, WSPs sublease their spectrum resources to multiple SUs in the same geographical area to maximize their revenue, exploiting the spatial reusability

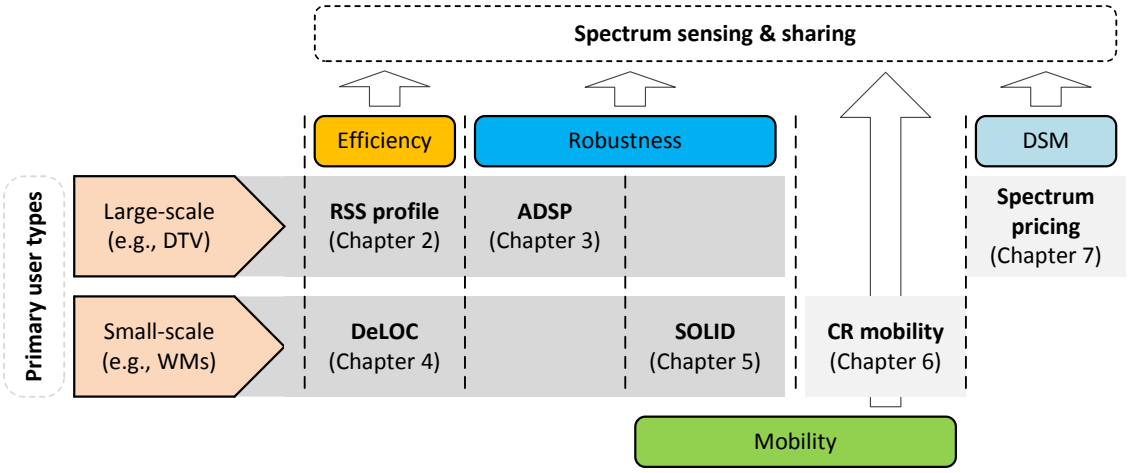


Figure 1.1: An overview of the thesis proposal.

of wireless spectrum resources. Such spectrum sharing complicates the spectrum price-demand relationship, making the DSM different from the traditional market where goods are exclusively owned by buyers [97]. For example, when SUs share a leased channel, quoting a low spectrum price would lead to paradoxical results: A low price may attract more users, but it will also increase the level of interference among SUs, thus discouraging SUs from accessing it even at a low price. Therefore, understanding this price-demand relationship is of great importance to the design of WSPs' optimal spectrum pricing and SUs' WSP selection strategies.

1.2 Research Contributions

The main objective of this research is to improve spectrum efficiency in CRNs. Specifically, this thesis presents new techniques and frameworks that allow SUs to accurately monitor spectrum conditions with high efficiency and robustness, and facilitate PU-SU interactions in DSM by providing economic incentives to PUs for spectrum sharing (see Fig. 1.1). The main research contributions are summarized as follows.

- **Efficient Detection of Large-Scale Primary Users**: In CRNs, regulatory bodies, such as the FCC, enforce extremely demanding detectability requirements to protect PUs' communications, which can hardly be achieved with one-time sensing using only a

single sensor. To overcome this challenge, we present a spectrum sensing framework that minimizes the sensing overhead by jointly optimizing cooperative sensing and sensing scheduling, while meeting the desired detection performance.

Specifically, we design an optimal sensing framework for large-Scale PU detection in CRNs that directs the base station (BS) to manage spectrum sensing by (i) constructing each primary signal's spatial profile of received signal strengths (RSSs) as a detection criterion, (ii) selecting an optimal set of sensors for cooperative sensing, and (iii) finding an optimal time to stop sensing. The evaluation results show that the proposed sensing framework significantly reduces sensing overhead while meeting the detection requirements set by the FCC.

- **Robust Detection of Large-Scale Primary Users**: In CRNs, making distributed sensing secure is challenging because of two unique CR-features—openness of a low-layer protocol stack in SDR devices and nonexistence of communications between primary and secondary devices. Moreover, cooperative sensors can be faulty or erroneous due to hardware/software defects. As a result, the sensing reports that they produce may have non-zero offsets.

  As a first step towards addressing this challenge, we propose an attack-tolerant cooperative sensing scheme for large-scale PU detection. In the proposed scheme, the fusion center cross-checks sensors' measurement results with neighboring sensors to prevent compromised (or faulty) sensors from affecting the final decision at the fusion center. The key idea is to pre-filter abnormal sensing reports by exploiting shadow fading correlation in RSSs among neighboring sensors by means of a correlation filter. Our evaluation results show that the proposed sensing framework can still meet the detection requirements even in the presence of attackers.

- **Efficient Detection of Small-Scale Primary Users**: In CRNs, detecting small-scale primary signals such as WMs is a challenging problem, due to their small signal footprint and the unpredictability of their spatial and temporal usage patterns. To overcome these challenges, we propose a small-scale PU detection framework based on the following two key observations: (i) we identify the data-fusion range for cooperative

sensing as a key factor in effective small-scale primary detection, and (ii) we observe that sensing performance is sensitive to the accuracy of location and transmit-power level information available to the secondary network.

Based on these observations, we propose an efficient sensing framework that iteratively performs location/transmit-power estimation and cooperative sensing with adaptive sensor selection based on the estimates, to achieve near-optimal detection performance. Our in-depth evaluation results show that our proposed small-scale primary detection framework provides high detection accuracy while maintaining a low false-triggering rate.

- **Robust Tracking of Mobile Small-Scale Primary Users**: In CRNs, in order to enhance utilization of spatial spectrum opportunities, SUs must be able to accurately and reliably track the location of small-scale mobile PUs. To accomplish this, we propose a framework, for accurate, attack/fault-tolerant tracking of small-scale mobile PUs.

  The key idea is that it exploits the temporal shadow fading correlation in the primary signal strengths measured at cooperative sensors induced by the primary's mobility. To realize this idea, we augment the conventional Sequential Monte Carlo (SMC)-based target tracking with shadow-fading estimation. By estimating shadow-fading gain between the primary transmitter and sensors, the proposed framework will not only significantly improve the accuracy of primary tracking in the absence of attack, but will also successfully tolerate sophisticated attacks, such as "slow-poisoning," preserving localization accuracy and improving spatial spectrum efficiency.

- **Mobile Cognitive Radio Networks**: We envision that future mobile devices will incorporate CR-functionality and be capable of dynamic and flexible spectrum access. To enable DSA for mobile CRs, we identify and address fundamental challenges posed by mobile SUs that do not exist in the case of stationary CRNs where the locations of PUs and SUs are known *a priori* to the secondary BS.

  Specifically, we model spectrum availability from the mobile CR devices' perspective. Based on the spectrum availability model, we design an efficient spectrum-sensing strategy to protect PUs' communications from SUs' mobility-induced interference. In

addition, to better utilize spatio-temporal spectrum opportunities, we design an optimal distributed channel-access strategy for mobile SUs. We demonstrate the accuracy of our SU mobility-aware spectrum availability model via in-depth simulation study. Moreover, our evaluation results show that our proposed spectrum sensing and access mechanisms significantly improve SUs' throughput performance and reduce energy consumption due to spectrum sensing, while protecting PUs' communications.

- **Optimal Spectrum Pricing in DSM**: In future wireless environments, a wide range of spectrum resources will be available in the market as a result of the current trend of deregulation of wireless spectrum. To obtain useful insights on the impact of spectrum heterogeneity, we introduce a new DSM model where WSPs with heterogeneous spectrum resources compete for a higher market share.

  In particular, we propose a new spectrum price-demand model based on the desire of SUs to maximize their own utility, by evaluating the key contributing factors, such as the impact of spectrum heterogeneity, spatial spectrum sharing, and total spectrum demand. We then derive SUs' optimal WSP selection strategy based on a mean-field approach to study how spectrum heterogeneity affects market equilibrium. Finally, we model the pricing strategies among WSPs as a non-cooperative game and identify the key factors that influence the Nash Equilibrium (NE) points, taking into account the price-demand relation caused by the utility maximizing behavior of SUs.

1.3 Organization of the Thesis

The remainder of this thesis is organized as follows. Chapter 2 presents an efficient cooperative sensing framework based on spatial-RSS profile of a primary signal. The proposed framework minimizes sensing overhead while meeting the detectability requirements. Chapter 3 proposes a secure cooperative spectrum sensing framework, called ADSP, for the detection of large-scale PUs. ADSP accurately detects and filters out compromised or erroneous sensing reports by exploiting shadow fading correlation in sensing reports among nearby sensors. Chapter 4 presents a small-scale primary detection framework, called DeLOC. DeLOC allows SUs to accurately detect small-scale PUs by jointly per-

forming cooperative sensing and location/transmit-power estimation. Chapter 5 proposes a robust mobile small-scale PU tracking system, called `SOLID`. `SOLID` enables accurate and robust location tracking of mobile primary transmitter, thus achieving high spatial spectrum efficiency. Chapter 6 develops optimal spectrum sensing and access strategies for mobile cognitive radios. The proposed schemes protect primary communications from mobile cognitive radios while minimizing energy consumption by spectrum sensing. Chapter 7 studies optimal spectrum pricing and WSP selection strategies for a duopoly DSM in which WSPs compete with heterogeneous spectrum resources. Finally, the thesis concludes with Chapter 8.

CHAPTER 2

Efficient Detection of Large-Scale Primary Users

2.1 Introduction

In DSA networks, spectrum opportunities can be detected in various ways, such as spectrum sensing [57, 81], beacon protocol [18, 86], or geo-location database [21, 63, 110]. Spectrum sensing directly detects the presence or absence of a primary signal in a target spectrum band. To protect PUs' communications, spectrum sensing must meet strict requirements set by regulatory bodies, such as the Federal Communications Commission (FCC). For example, in the IEEE 802.22 Wireless Regional Area Networks (WRANs) [1, 138], a primary signal as weak as -20 dBm must be detected with high accuracy, i.e., both false-alarm and mis-detection probabilities must be less than $10\,\%$ [38]. Unfortunately, this stringent requirement cannot be met using one-time sensing with a single sensor, regardless of the underlying PHY-layer sensing schemes, e.g., energy/feature detection [127, 131, 133]. A second method for detecting a primary signal, or the beacon protocol, has also been proposed to detect the existence of a primary signal more efficiently. However, the beacon protocol requires legacy devices to be equipped with an external beacon device. Therefore, it may not be a feasible solution for already widely-deployed legacy systems because of the high cost involved. Moreover, such modifications violate the basic premise of DSA—opportunistic spectrum access should require no modification to legacy systems. Thus, the beacon protocol cannot obviate the need for spectrum sensing. Alternatively, a geo-location database can be used to identify spectrum availability at a given

time and location [110]. Such a database can be constructed and maintained by regulatory bodies, such as the FCC in the US or Ofcom in the UK, or by a trusted third party. A database can be used for spectrum bands on which PU activities are relatively predictable, e.g., DTV signals with long ON/OFF periods. In fact, the FCC recently mandated the use of a geo-location database for accessing TV band white-spaces [50]. Algorithms and frameworks for implementing geo-location databases for TV spectrum bands are currently under development [63, 110]. However, it may require a considerable amount of time to construct a database with a complete spectrum map. Looking up the database to check spectrum availability also consumes system resources.

In this chapter, we focus on improving spectrum-sensing performance via joint design of two MAC-layer sensing methods, *cooperative sensing* and *sensing scheduling*. This joint design allows SUs not only to overcome the performance deficiency of PHY-layer sensing, but also to make the tradeoff between performance gain and sensing overhead. In cooperative sensing, a fusion center (a base station) directs multiple sensors at different locations to perform spectrum sensing simultaneously during each sensing (quiet) period, thus exploiting sensor location diversity [20, 35, 55, 57, 106, 122, 143, 144]. Scheduling sensing also aims to improve detection performance by having SUs perform spectrum sensing at various time intervals, thereby exploiting temporal variations in received signal strengths (RSSs) at each sensor [39, 81]. However, all SUs (sensors)[1] must remain silent during sensing periods so that SU signals are not misinterpreted as a primary signal [112]. These periods of silence waste precious resources, such as energy and time, and ultimately degrade the quality-of-service (QoS) of SU communications. Therefore, the fusion center must carefully select a set of cooperative sensors and optimally schedule sensing periods so as to minimize sensing-induced interruptions, while guaranteeing the required detection performance, even for weak primary signals.

To address this practical challenge, we propose an efficient spectrum-sensing framework that jointly exploits spatial and temporal RSS variations to minimize sensing overhead subject to the detectability requirements. In particular, we address the following three

[1] We use the terms *SUs* and *sensors* interchangeably since we focus on the sensing functionality of SUs in this thesis.

key issues in MAC-layer spectrum sensing: (i) which sensors to use for cooperative sensing, (ii) how to incorporate their heterogeneous sensitivities in data fusion, and (iii) how to minimally schedule spectrum sensing.

2.1.1 Contributions

This chapter makes the following main contributions.

- Introduction of a new concept of spatial RSS-signature-based cooperative sensing that exploits the spatial variations in RSSs among cooperating sensors by *learning* the RSS distributions at sensor locations. This is a feasible and useful approach in CRNs where sensor locations are *stationary*, thus making their RSS distributions unique and (pseudo) time-invariant.

- Development of a simple and near-optimal *linear* data-fusion rule for detection of a primary signal based on a one-time sensing via a linear discriminant analysis (LDA). This is based on the observation that, when energy detection is employed in a low SNR environment such as IEEE 802.22 WRANs, spatial RSS distributions can be approximated as multi-dimensional Gaussian with a common covariance matrix. The theoretical performance of the LDA-based decision rule under shadow fading is also presented.

- Proposal of an optimization framework for minimizing the sensing overhead of cooperative sensing, which consists of: (i) an algorithm for selecting an optimal set of sensors for cooperative sensing , and (ii) an online sensing-period scheduling algorithm that finds an *optimal stopping time* via a sequential probability ratio test (SPRT) based on the sensing results.

- In-depth simulation to demonstrate the benefits of the proposed spectrum-sensing algorithms. Our simulation results show that the proposed RSS-profile-based detection schemes for both one-time (i.e., LDA-based) and sequential (i.e., SPRT-based) sensing significantly improve detection performance over the conventional decision

fusion rules, such as the OR-rule.[2] The results also show that our algorithms for sensor selection reduce the average sensing overhead significantly.

2.1.2 Organization

The remainder of this chapter is organized as follows. Section 2.2 briefly reviews the IEEE 802.22 WRAN and the energy-detection technique, followed by our approach to exploiting the spatio-temporal variations in RSSs for spectrum sensing. Section 2.3 presents our RSS-signature-based detection scheme for one-time sensing and its theoretical performance. Section 2.4 introduces our cooperative sensing algorithms designed to (i) select an optimal set of sensors and (ii) find an optimal time to stop sensing. Section 2.5 evaluates the performance of the proposed algorithms, Section 2.6 reviews related work on spectrum sensing. Section 2.7 concludes the chapter.

2.2 Preliminaries

In this section, we first introduce the DSA network model, spectrum sensing for incumbent detection, and then outline our proposed spectrum-sensing framework.

2.2.1 DSA Network Model

We consider a DSA network in which primary and secondary systems coexist in the same geographical area, as shown in Fig. 2.1. We assume a large-scale stationary primary system, such as DTV users in TV spectrum bands. A secondary system is an infrastructure-based network and each cell consists of a single base station (or fusion center) and multiple sensors. The fusion center selects sensors and directs them to perform sensing by scheduling sensing (quiet) periods multiple times. At the end of each sensing period, sensors report their measurement results to the fusion center sequentially. Based on the sensing reports, the fusion center will make a final decision on the presence or absence of a primary signal

[2]The OR-rule is the most common decision-fusion rule in the absence of prior knowledge of RSS distributions [35, 55, 81, 127, 129].

and will announce the decision to the SUs in the cell. We assume the existence of a separate control channel which provides reliable communication between the fusion center and sensors.

2.2.2 Spectrum-Sensing Model

Spectrum sensing consists of PHY- and MAC-layer mechanisms. For PHY-layer sensing, we assume energy detection instead of other sensing techniques, such as matched-filter detection [87], cyclostationary feature detection [94] and compressed sensing [160]. The energy detector is one of the most widely-used because of its simple design and low complexity; it simply measures signal power on a target frequency band and does not require *a priori* knowledge of primary-signal-specific features.

Regarding the existence of a primary signal on a given channel, there are two hypotheses, i.e.,

$$y_i(n) = \begin{cases} w_i(n) & \mathcal{H}_0 \text{ (no primary signal)} \\ s_i(n) + w_i(n) & \mathcal{H}_1 \text{ (primary signal exists)}, \end{cases}$$

where $y_i(n)$ is the signal received by a sensor, $s_i(n)$ is the primary signal, and $w_i(n)$ is an independent and identically distributed (i.i.d.) additive white Gaussian noise (AWGN) at sensor i in the n-th time slot within the sensing duration. The test statistic of the energy detector is an estimate of average RSS [127]:

$$T_i = \frac{B}{M} \sum_{n=1}^{M} y_i(n) * y_i(n), \tag{2.1}$$

where B is the channel bandwidth, and M is the number of signal samples during a sensing period. The test statistic can be approximated as a Gaussian distribution using the central limit theorem (CLT) because the signal sample size, M, is sufficiently large, even with a short sensing duration (e.g., 1 ms). For example, assuming that the signal is sampled at the Nyquist rate, a sensor can obtain $M = 6 \times 10^3$ samples for a 6 MHz TV channel within 1 ms [129].

Then, the probability density function (p.d.f.) of the test statistic T_i at sensor i is given as [127]:

$$T_i \sim \begin{cases} \mathcal{N}\big(NB, \frac{(NB)^2}{M}\big) & \mathcal{H}_0 \\ \mathcal{N}\big(P_i + NB, \frac{(P_i+NB)^2}{M}\big) & \mathcal{H}_1, \end{cases} \tag{2.2}$$

where P_i is the received primary signal strength and N is the noise spectral density.

We make two additional assumptions as follows.

A1. The separation between the primary transmitter and sensors is relatively larger than the separation between sensors, which is reasonable in large-scale networks such as the IEEE 802.22 WRANs [137].

A2. The impact of multi-path fading on spectrum sensing is negligible due to a wide channel bandwidth (e.g., larger than the coherent bandwidth) as in DTV signal detection in IEEE 802.22 WRANs [127].[3]

Based on the above assumptions, the received primary signal strength at sensor i can be expressed as $P_i = P_R \cdot e^{Y_i}$, where P_R is the average RSS within a cell, and e^{Y_i} is the shadowing gain between the primary transmitter and sensor i. Shadow fading can be characterized by the shadowing dB-spread, σ_{dB}, and it has the relationship $\sigma = 0.1 \ln(10)\, \sigma_{dB}$.

Remark: It is important to note that in *stationary* secondary systems where SUs do not move, the shadowing gain Y_i is a specific realization of a normal random variable $Y \sim \mathcal{N}(0, \sigma^2)$. Thus, the channel gain is also (pseudo) time-invariant and determined based on sensor locations. We will show that there is no correlation among stationary sensors in Section 2.4.1.

2.2.3 Outline of the Proposed Approach

The performance of energy detection is highly susceptible to signal-to-noise ratio (SNR), thus limiting its ability to meet the FCC's requirement of accurately detecting a primary signal as weak as -20 dB SNR. Moreover, *noise uncertainty* [142] at sensors also pre-

[3]The detection performance of the energy detector under multi-path fading can be found in [41].

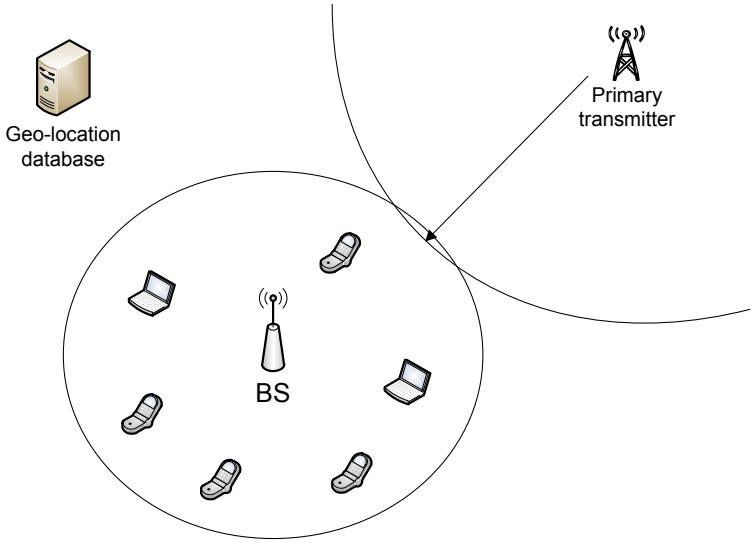


Figure 2.1: DSA network model: SUs perform spectrum sensing simultaneously during each sensing period and report their measurement results to the fusion center.

vents the applicability of energy detection to such a weak primary signal. To overcome these challenges and design an efficient spectrum sensing algorithm, we exploit both *spatial* and *temporal* variations of RSSs among collaborating sensors. As mentioned earlier, in DSA networks with static sensor deployment, the measured RSS at each sensor is (pseudo) time-invariant. This allows the fusion center to *learn* the RSS distributions at sensors and construct their spatial RSS profile. Upon collecting the sensing results during each scheduled sensing period, the fusion center compares the observed RSS values with the RSS profile. A similarity between the RSS distribution and the primary signal can be interpreted as an indication of the presence of a primary signal, and vice versa. Using the RSS profile, the fusion center can adopt the sequential hypothesis testing framework for sensing scheduling, and can minimally schedule sensing periods only until it accumulates sufficient observations to determine whether or not a primary signal exists within a certain performance bound.

Fig. 2.2 illustrates our proposed spectrum-sensing framework, where the fusion center directs an optimally-chosen set of sensors to perform sensing until a decision is made on the existence of a primary signal, using a sequence of reported RSS values.

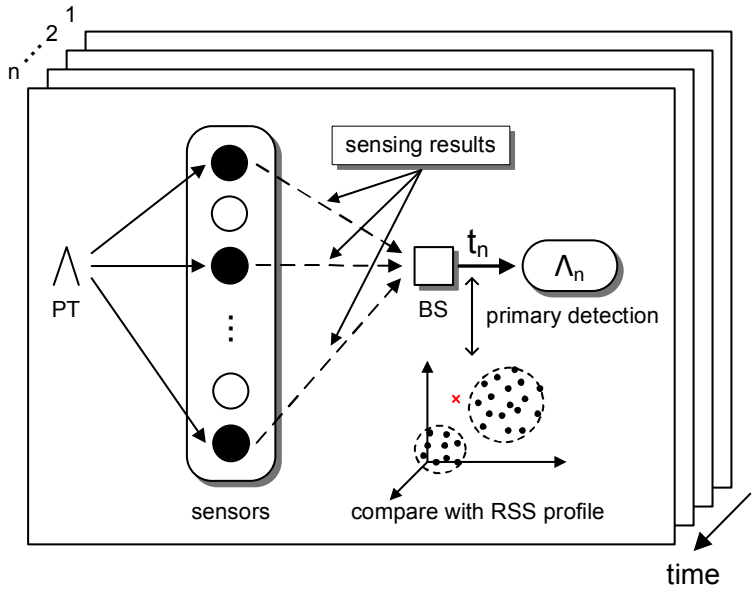


Figure 2.2: An illustration of the proposed spectrum-sensing framework: The base station (or fusion center) selects an optimal set of sensors (black nodes) for cooperative sensing and schedule sensing periods until it accumulates sufficient information to make a final decision.

2.3 RSS-Profile-Based Cooperative Sensing

In this section, we present the construction of a spatial RSS profile, formulate the problem of incumbent detection based on *one-time* sensing as a binary classification problem via *linear discriminant analysis* (LDA), and analyze its detection performance. We then characterize wireless network conditions, under which sensing scheduling is needed to meet desired detection requirements.

2.3.1 Construction of a Spatial RSS Profile

We propose to build a spatial profile of RSS distributions at multiple sensor locations, which will be used as a main reference for incumbent detection. Fig. 2.3 shows an example of spatio-temporal variations of the test statistics of the energy detector (i.e., T in Eq. (2.1)) at 15 sensors in various locations within a secondary cell. The *spatial* RSS diversity is due mainly to the different sensors' locations (thus different channel gains from the primary transmitter), whereas the *temporal* RSS variations are due mainly to the measurement error of the energy detector. The intensity of temporal variations depends on the sensing time

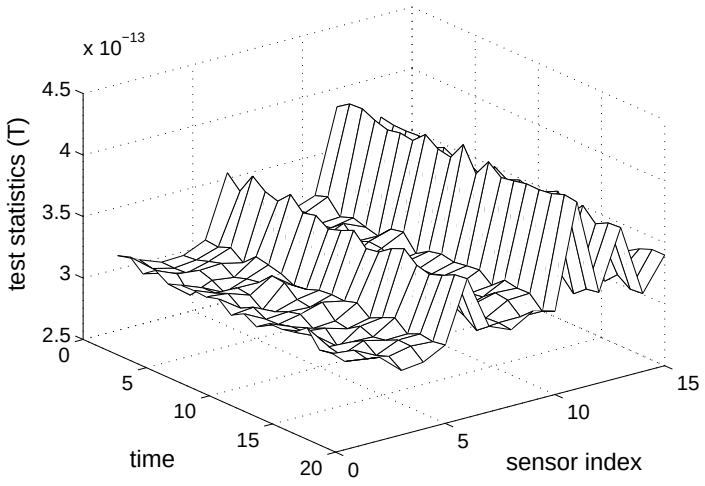


Figure 2.3: Example time sequences of test statistics $\{T_i\}$ for $1 \leq i \leq 15$ where $P_R = -110\,\text{dBm}$, $NB = -95.2\,\text{dBm}$, $M = 6 \times 10^3$, and $\sigma_{dB} = 5.5\,\text{dB}$. They clearly show spatial and temporal variations in the test statistics.

(i.e., M in Eq. (2.1)); the shorter the sensing time, the larger the temporal RSS variation due to the increase in measurement error. The fusion center can construct a unique profile of RSS distributions for a given set of sensors and sensing times. For RSS profiling, we assume a large enough training period (including both ON/OFF periods of the primary transmitter) for accurate estimation of RSS distributions.

Recall that the distribution of test statistic, T (equivalent to the estimated RSS), of the energy detector can be approximated as Gaussian (see Eq. (2.2)) using the CLT in both ON/OFF periods. Thus, the RSS profile of n_s stationary sensors is an n_s-dimensional Gaussian distribution, the parameters of which can be easily estimated using well-known techniques, such as maximum-likelihood estimation (MLE) as:

$$\hat{\pi}_k = N_k / N_{obs}, \tag{2.3}$$

$$\hat{\boldsymbol{\mu}}_k = \sum_{n=1}^{N_k} \mathbf{x}_n / N_k, \tag{2.4}$$

$$\hat{\boldsymbol{\Sigma}}_k = \sum_{n=1}^{N_k} (\mathbf{x}_n - \hat{\boldsymbol{\mu}}_k)(\mathbf{x}_n - \hat{\boldsymbol{\mu}}_k)^T / N_k, \tag{2.5}$$

where N_k is the number of observations under $\mathcal{H}_k$ where $k \in \{0, 1\}$, N_{obs} the number of

total observations in the training period, and $\hat{\pi}_k$ is *a priori* probability of class (hypothesis) k. $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ represent the mean vector and the covariance matrix during ON(1)/OFF(0) periods. Since the RSS distributions at each sensor location are (pseudo) time-invariant due to the static deployment of SUs, the RSS profile can be used reliably without frequent updating.

2.3.2 Detection with One-Time Sensing Based on Linear Discriminant Analysis (LDA)

We now present a detection rule using an RSS profile, given a one-time sensing measurement. Let $\mathbf{x} = [\, T_1, \ldots, T_{n_s}\,]^T$ denote the vector of test statistics of the energy detector measured by n_s cooperating sensors. Then, the incumbent detection problem can be cast into a binary Gaussian classification problem where the observed test statistic $\mathbf{x} \in \mathbb{R}^{n_s \times 1}$ belongs to one of two classes, $\mathcal{H}_0$ or $\mathcal{H}_1$, where

$$\begin{aligned} \mathcal{H}_0 &: \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \quad \text{(no primary signal)} \\ \mathcal{H}_1 &: \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \quad \text{(primary signal exists)}, \end{aligned}$$

where $\boldsymbol{\mu}_k \in \mathbb{R}^{n_s \times 1}$ and $\boldsymbol{\Sigma}_k \in \mathbb{R}^{n_s \times n_s}$ are the estimated mean vector and the covariance matrix of RSS distributions under $\mathcal{H}_k$, respectively. Note that $\boldsymbol{\Sigma}_0 = \sigma_n^2 \mathbf{I}$ where $\mathbf{I}$ is an $n_s \times n_s$ identity matrix and $\sigma_n^2 = (NB)^2/M$.

Under the general assumption of unequal covariance matrices, i.e., $\boldsymbol{\Sigma}_0 \neq \boldsymbol{\Sigma}_1$, the optimal decision rule for our detection problem can be found via *quadratic discriminant analysis* (QDA) [65]. Although QDA provides optimal detection performance for a general multivariate Gaussian with unequal covariance matrices, its quadratic decision boundaries do not yield a closed-form expression for detection performance [12].

In our problem, the quadratic decision rule can actually be linearized using *linear discriminant analysis* (LDA) on the basis of the following two important observations, i.e., the covariance matrix under $\mathcal{H}_1$, $\boldsymbol{\Sigma}_1$, can be:

1. *assumed as an identity matrix with fixed sensor locations,* and then,

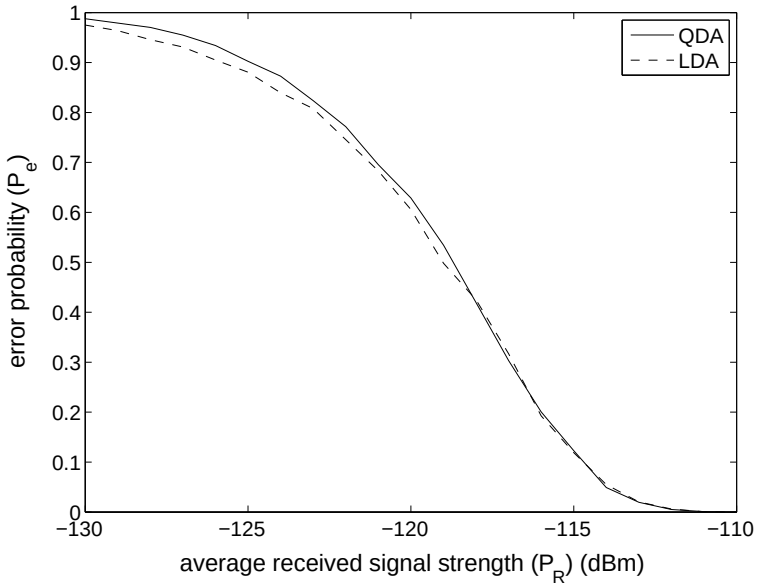


Figure 2.4: Error performances of QDA vs. LDA. The performance difference is insignificant even in a very low SNR environment. These are the results of a Monte Carlo simulation with 10^7 runs.

2. *approximated as* $\mathbf{\Sigma}_1 \approx \mathbf{\Sigma}_0 = \sigma_n^2 \mathbf{I}$ *in a very low SNR environment.*

Regarding the first observation, the covariance matrix $\mathbf{\Sigma}_1$ may not appear to be an identity matrix because of the existence of shadow correlation in primary signal strengths [60]. However, as mentioned earlier, when sensor locations are fixed, their RSSs are also (pseudo) time-invariant and the randomness in the test statistics comes only from the noise processes (i.e., measurement errors), which are independent of each other. Thus, the correlation of RSSs between any pair of sensors does not exist, so we can assume that $\mathbf{\Sigma}_1$ is also an identity matrix as $\mathbf{\Sigma}_0$ (see Section. 2.4.1).

Regarding the second observation, the received primary signal strength may be significantly lower than that of the noise power for a very weak primary signal. For example, the FCC requires the detection of a DTV signal as weak as $-20\,$dB, assuming the typical noise level $NB = -95.2\,$dBm [133]. Therefore, it is reasonable to assume that $P_i + NB \approx NB \ \forall i$, and thus, $\mathbf{\Sigma}_1 \approx \mathbf{\Sigma}_0 = \sigma_n^2 \mathbf{I}$. Fig. 2.4 justifies these assumptions by showing that the error performances of QDA and LDA are almost the same in very low SNR environments.

Based on the above observations, we compute the log-likelihood of the two hypotheses

$\mathcal{H}_0$ and $\mathcal{H}_1$ as:

$$\log\left(\frac{Pr(\mathcal{H}_1|\mathbf{x})}{Pr(\mathcal{H}_0|\mathbf{x})}\right) = \log\frac{g_1(\mathbf{x})}{g_0(\mathbf{x})} + \log\frac{\pi_1}{\pi_0}$$
$$= \log\frac{\pi_1}{\pi_0} - \frac{1}{2}(\boldsymbol{\mu}_1+\boldsymbol{\mu}_0)^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0) + \mathbf{x}^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1-\boldsymbol{\mu}_0), \quad (2.6)$$

where $g_k(\mathbf{x})$ is the estimated Gaussian distribution of the sensing reports, $\mathbf{x}$, under $\mathcal{H}_k$ and $\boldsymbol{\Sigma}$ is the common covariance matrix. π_k is *a priori* probabilities of hypotheses $\mathcal{H}_k$ as defined in Eq. (2.3).

From Eq. (2.6), we have the following linear discriminant function, $\delta_k(\mathbf{x})$, as:

$$\delta_k(\mathbf{x}) = \mathbf{x}^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_k - \frac{1}{2}\boldsymbol{\mu}_k^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_k + \log(\pi_k). \quad (2.7)$$

That is, the fusion center will assume hypothesis $\mathcal{H}_k$ where δ_k is maximized, i.e., $\mathcal{H}_k = \arg\max_k \delta_k(\mathbf{x})$.

In our two-class problem, the fusion center will assume $\mathcal{H}_1$ if the following condition holds:

$$\mathbf{x}^T\boldsymbol{\Sigma}^{-1}(\hat{\mu}_1-\hat{\mu}_0) > \frac{1}{2}\hat{\mu}_1^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\mu}_1 - \frac{1}{2}\hat{\mu}_0^T\hat{\boldsymbol{\Sigma}}^{-1}\hat{\mu}_0 + \log\frac{\pi_0}{\pi_1}. \quad (2.8)$$

Otherwise, the fusion center will assume $\mathcal{H}_0$.

Note that the fusion center may not have an accurate estimation of *a priori* probability of a primary signal. In such a case, the fusion center can set $\pi_0 = \pi_1 = 0.5$. Then, we have a simple distance-based decision rule for incumbent detection:

$$\|\mathbf{x}-\boldsymbol{\mu}_0\| \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \|\mathbf{x}-\boldsymbol{\mu}_1\|. \quad (2.9)$$

Eq. (2.9) indicates that, under both hypotheses, the decision is made based solely on the distance between the observed RSS vector, $\mathbf{x}$, and the mean vectors of the RSS profile, $\boldsymbol{\mu}_k$. Although Eq. (2.9) is optimal in minimizing detection error performance (i.e., the sum of false-alarm and mis-detection probabilities), the detection requirements are often expressed in terms of mis-detection probability for a fixed false-alarm probability. In what follows,

we analyze the performance of the proposed RSS-profile based spectrum sensing.

2.3.3 Theoretical Performance

Let $\mathcal{T}(\mathbf{x}) \triangleq \mathbf{w}^T\mathbf{x}$ denote the test statistic for incumbent detection, which is calculated based on the observed RSS vector $\mathbf{x}$, where $\mathbf{w} \triangleq (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) \in \mathbb{R}^{n_s \times 1}$. Note that $\|\mathbf{w}\|$ is the Euclidean distance between the centroids of two Gaussian distributions under both hypotheses, where the centroids are the vectors of average RSSs at sensor locations. It can be easily shown that the test statistic $\mathcal{T}(\mathbf{x})$ follows a Gaussian distribution, i.e., $\mathcal{T}(\mathbf{x}) \sim \mathcal{N}\big(\mathbf{w}^T\boldsymbol{\mu}_k, \sigma_n^2\|\mathbf{w}\|^2\big)$ under $\mathcal{H}_k$.

Then, the probability of false alarm under our LDA-based decision rule with the decision threshold $\eta \in \mathbb{R}$ is given as:

$$P_{FA}^{LDA} \triangleq Pr(\mathcal{T}(\mathbf{x}) > \eta \,|\, \mathcal{H}_0) = Q\bigg(\frac{\eta - \mathbf{w}^T\boldsymbol{\mu}_0}{\sigma_n\|\mathbf{w}\|}\bigg), \tag{2.10}$$

where $Q(\cdot)$ is the Q-function. Using Eq. (2.10), the decision threshold η can be derived for the desired P_{FA}^{LDA} as:

$$\eta = \sigma_n \cdot \|\mathbf{w}\| \cdot Q^{-1}\big(P_{FA}^{LDA}\big) + \mathbf{w}^T\boldsymbol{\mu}_0. \tag{2.11}$$

Then, based on Eqs. (2.10) and (2.11), the probability of mis-detection, P_{MD}^{LDA}, is given as:

$$P_{MD}^{LDA} \triangleq Pr(\mathcal{T}(\mathbf{x}) < \eta \,|\, \mathcal{H}_1) = 1 - Q\bigg(Q^{-1}\big(P_{FA}^{LDA}\big) - \frac{\|\mathbf{w}\|}{\sigma_n}\bigg). \tag{2.12}$$

Eq. (2.12) indicates that, when the desired false-alarm probability, P_{FA}^{LDA}, is given, the achievable mis-detection, P_{MD}^{LDA}, depends on the noise variance $\sigma_n^2 = \frac{(NB)^2}{M}$ in energy detection and the distance $\|\mathbf{w}\|$. That is, P_{MD}^{LDA} decreases as the sensing duration (thus the number M of sensing samples) increases, since a large number of samples would make the decision more accurate due to the reduced noise variance (measurement error).

Recall that $\mathbf{w}$ is defined as the difference in RSSs under both hypotheses, i.e., $\mathbf{w} \triangleq (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) = [P_1, \ldots, P_{n_s}]^T$. Therefore, $\|\mathbf{w}\|$ under shadow fading is given as:

$$\|\mathbf{w}\| = \bigg[\sum_{i=1}^{n_s} P_i^2\bigg]^{1/2} = P_R \cdot \bigg[\sum_{i=1}^{n_s} \big(e^{Y_i}\big)^2\bigg]^{1/2}, \tag{2.13}$$

where P_R is the average RSS in the secondary cell due to path loss and Y_i is a location-dependent realization of a random variable $Y \sim \mathcal{N}(0, \sigma^2)$ where $\sigma = 0.1 \ln(10)\, \sigma_{dB}$.

To understand the impact of shadow fading on detection performance (in terms of P_{MD}^{LDA} given a fixed P_{FA}^{LDA}), we study the distribution of $\|\mathbf{w}\|$ in Eq. (2.13). Although there is no closed-form expression available for the power sum of log-normal random variables in Eq. (2.13) [123], the power sum can be approximated accurately by rendering the sum itself as another log-normal random variable [52].

Let $e^{Z'} \sim e^{2Y_1} + e^{2Y_2} + \cdots + e^{2Y_{n_s}}$. Then, by following the result in [52], the sum can be approximated by matching its mean and variance with $e^{Z'}$. The first two moments of $e^{Z'}$ are $\mathbb{E}[e^{Z'}] = e^{\mu_{Z'} + \sigma_{Z'}^2/2}$ and $\mathbb{E}[e^{2Z'}] = e^{2\mu_{Z'} + 2\sigma_{Z'}^2}$. Our final goal is to approximate the square root of the power sum, i.e., $e^{Z} = (e^{Z'})^{1/2}$, which is still a log-normal random variable. Thus, by equating the first two moments of $e^{Z'}$ and the power sum, $\sum_{i=1}^{n_s} (e^{Y_i})^2$, and then taking $e^{Z} = (e^{Z'})^{1/2}$, we have $\|\mathbf{w}\| \approx P_R \cdot e^{Z}$ with the random variable $e^{Z} \sim$ Log-$\mathcal{N}(\mu_Z, \sigma_Z^2)$ where:

$$\sigma_Z^2 = \frac{1}{4} \log \left[\frac{(e^{4\sigma^2} - 1)}{n_s} + 1 \right] \tag{2.14}$$

and

$$\mu_Z = \frac{1}{2} \log(n_s) + \sigma^2 - \frac{\sigma_Z^2}{4}. \tag{2.15}$$

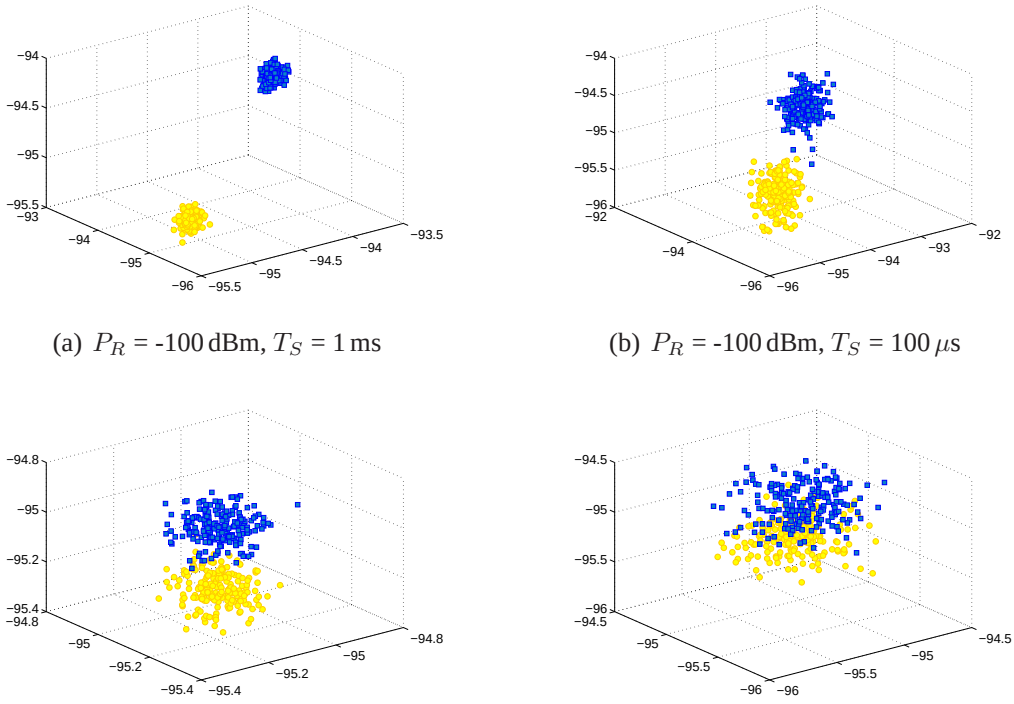
Assuming that the sensors experience independent log-normal shadow fading, we can derive the average mis-detection probability as:

$$\overline{P}_{MD}^{LDA} = \int_{-\infty}^{\infty} \left[1 - Q\left(Q^{-1}\left(P_{FA}^{LDA}\right) - \frac{P_R \cdot e^z}{\sigma_n} \right) \right] \cdot f_z \cdot \mathrm{d}z,$$

where $f_z = \frac{1}{\sigma_Z \sqrt{2\pi}} \exp\left[- \frac{(z-\mu_Z)^2}{2\sigma_Z^2} \right],\ -\infty < z < \infty$.

Fig 2.5 shows examples of RSS distributions under various wireless environments. The figure shows that the centroids of the two Gaussian distributions will have a larger separation, $\|\mathbf{w}\|$, with a higher average primary signal strength P_R, or a longer sensing time T_S. As indicated in Eq. (2.12), a larger separation between the two RSS distributions increases the incumbent detection performance.

Fig. 2.6 plots mis-detection probabilities, P_{MD}, for a given false-alarm probability.



(a) P_R = -100 dBm, T_S = 1 ms

(b) P_R = -100 dBm, T_S = 100 μs

(c) P_R = -110 dBm, T_S = 1 ms

(d) P_R = -110 dBm, T_S = 100 μs

Figure 2.5: Distribution of the measured RSS vector from three cooperating sensors under various wireless environments. The yellow (blue) dots indicate the vector of measured test statistics under $\mathcal{H}_0$ ($\mathcal{H}_1$). The simulation parameters are set $NB = -95.2$ dBm and $\sigma_{dB} = 5.5$ dB.

The figure shows that the mis-detection performance increases as the average primary signal strength decreases. It shows that spectrum sensing with a single sensor may not be sufficient to protect primary communications. To further improve detection performance while introducing minimal sensing overhead, in what follows, we jointly optimize sensor selection, sensing time, and scheduling sensing.

2.3.4 The Necessity of Sensing Scheduling

We now characterize the network conditions under which the fusion center must schedule spectrum sensing multiple times to meet a given detectability requirement. As we observed in previous sections, the performance of one-time spectrum sensing depends on various network parameters, such as average primary signal strength, P_R, the number of sensors for sensing collaboration n_s, and detection requirements, P_{FA} and P_{MD}. In prac-

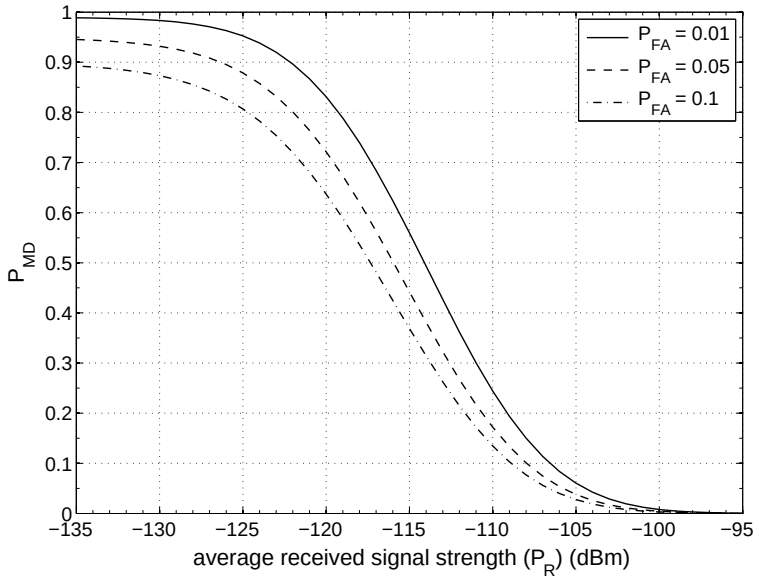


Figure 2.6: Detection performance of RSS-profile-based cooperative spectrum sensing. The incumbent mis-detection performance for a given false-alarm probability increases as the average primary received signal strengths, P_R, decreases. The simulation parameters are set $n_s = 1$, $NB = -95.2$ dBm, $M = 6 \times 10^3$, and $\sigma_{dB} = 5.5$ dB.

tice, the fusion center may avoid scheduling sensing multiple times by employing a larger number of sensors for cooperative sensing. We define the minimum number of sensors, n_s^*, to meet a given detection requirement as:

$$n_s^* = \arg\min_{n_s} \Big\{ n_s : \overline{P}_{FA}, \overline{P}_{MD} \leq \xi \Big\}. \tag{2.16}$$

Fig. 2.7 plots the minimum number of sensors required to achieve the desired level of detection probability, ξ. This figure shows that the required number of sensors increases exponentially as the average primary signal strength, P_R, decreases. The number of sensors also increases as the detection requirement, ξ, becomes stricter. For example, when $P_R = -116$ dBm, more than 70 sensors are required to achieve $P_{FA}, P_{MD} \leq 0.01$. In practice, it may be difficult to find such a large number of sensors within a cell, and more importantly, having such a large number of sensors may incur significant overhead. Therefore, in a low SNR environment with strict detection requirements, the fusion center must schedule sensing multiple times to make the best tradeoff between performance and overhead.

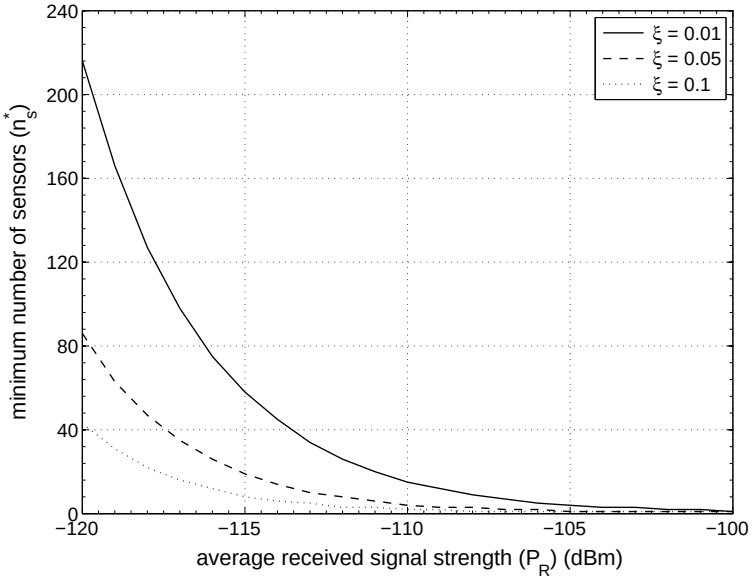


Figure 2.7: Minimum number of sensors required to achieve a desired incumbent detection performance with one-time sensing. The fusion center must employ more sensors as the average primary signal strength decreases or as the detection requirement, ξ, becomes more strict. The simulation parameters are set $NB = -95.2\,\text{dBm}$, $M = 6 \times 10^3$, and $\sigma_{dB} = 5.5\,\text{dB}$.

2.4 Optimal Cooperative Sensing Framework for Sensing Overhead Minimization

In this section, we first propose an adaptive online algorithm that finds an optimal stopping time for scheduling sensing periods, subject to the detection requirements, given a set of collaborative sensors. We then present an algorithm to find an optimal set of sensors and an optimal sensing duration that minimizes the average sensing-time overhead.

2.4.1 Correlation Analysis

Before presenting an optimal sensing scheduling algorithm, we would like to show that there is no correlation in RSSs among stationary sensors. As we discussed in Section 2.3.2, the randomness in sensing results (i.e., the output of the energy detector) comes only from measurement noise, which is independent for each sensor. To illustrate this, we measure the Pearson's correlation coefficient between a pair of sensors. The Pearson's correlation

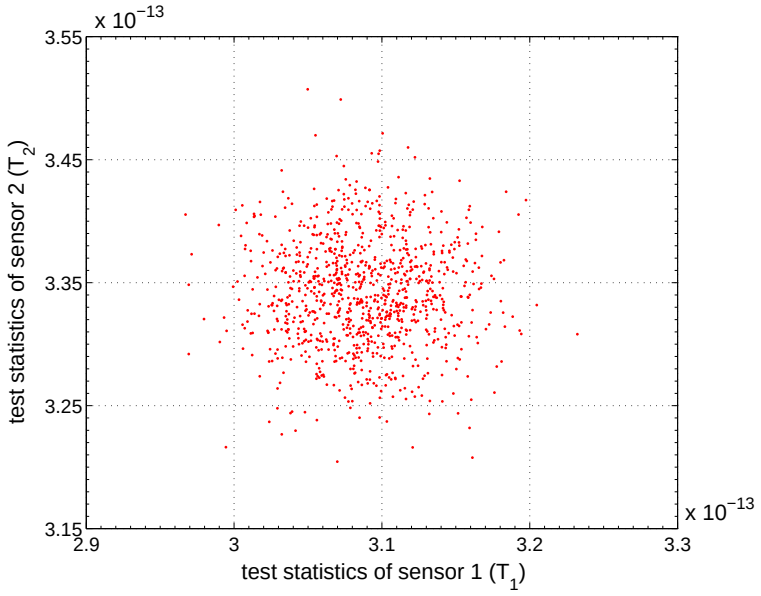


Figure 2.8: Distribution of test statistics measured as two sensors over 10^3 sensing periods. The correlation coefficient of the measurements is $\hat{\rho}_{1,2} \approx 0.0053$. The simulation parameters are set as $NB = -95.2$ dBm, $M = 6 \times 10^3$, and $\sigma_{dB} = 5.5$ dB.

coefficient of sensing reports from the sensors i and j can be estimated as:

$$\hat{\rho}_{i,j} = Corr(i,j) = \frac{\sum_{i=1}^{n}(T_i - \mathbb{E}[T_i])(T_j - \mathbb{E}[T_j])}{(n-1)\hat{\sigma}_i\hat{\sigma}_j}, \tag{2.17}$$

where $\hat{\sigma}_i$ and $\hat{\sigma}_j$ are the sample standard deviations of test statistics T_i and T_j, i.e.,

$$\hat{\sigma}_i = \left(\frac{1}{n-1}\sum_{i=1}^{n}\left(T_i - \mathbb{E}[T_i]\right)^2\right)^{\frac{1}{2}}. \tag{2.18}$$

Fig. 2.8 plots 10^3 sets of test statistics (T_1, T_2) measured at two sensors. Interestingly, the corresponding Pearson coefficient is only about 0.0053, indicating that the sensing results from the two sensors are not correlated. This is in sharp contrast to most previous work, in which avoiding shadow fading correlation is one of the most important criteria for the selection of sensors [81, 125].

This implies that when the fusion center can obtain the average primary signal strength at each sensor (via estimation or learning), then the physical separation between cooperating sensors may not critically affect detection performance. Note that, in stationary DSA network environments, such as the IEEE 802.22 WRANs, where all the sensors (called

CPEs) in the network do not move, the fusion center can easily obtain the RSS at each sensor, e.g., by simply observing the reports from the sensors.

2.4.2 Optimal Stopping Rule for Sensing Scheduling

The fusion center can schedule sensing multiple times before making a final decision, and thus, it receives a sequence of observations (i.e., measured RSSs) from the sensors. This makes sequential detection suitable for our problem. In particular, among the various sequential detection schemes, we adopt Wald's *Sequential Probability Ratio Test* (SPRT) [147] since it is optimal in the sense of minimizing the average number of observations, given bounded probabilities of false-alarm and mis-detection.

Let $t_n \triangleq \sigma_n^{-1} \cdot \|\mathbf{w}\|^{-1} \cdot \mathcal{T}(\mathbf{x}_n)$ denote the *normalized* test statistic based on the observed RSS vector $\mathbf{x}_n$ in the n-th sensing period. The decision statistic Λ_N is the log-likelihood ratio based on N sequential observations (i.e., test statistics) $t_1, \ldots, t_N$ as:

$$\Lambda_N \triangleq \lambda(t_1, \ldots, t_N) = \ln \frac{f_1(t_1, \ldots, t_N)}{f_0(t_1, \ldots, t_N)}, \tag{2.19}$$

where $f_k(t_1, \ldots, t_N)$ is the joint p.d.f. of the sequence of observations under hypotheses $\mathcal{H}_k \ \forall k$. Recall that $\{t_n\}_{n=1}^N$ are Gaussian, and without loss of generality, we assume that they are i.i.d. Then, Eq. (2.19) becomes:

$$\Lambda_N = \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} \ln \frac{f_1(t_n)}{f_0(t_n)}, \tag{2.20}$$

where $f_k(t_n)$ is $\mathcal{N}(\theta_k, 1)$ with $\theta_k \triangleq \mathbb{E}[t_n \,|\, \mathcal{H}_k] = \frac{\mathbf{w}^T \mu_k}{\sigma_n \|\mathbf{w}\|} \ \forall k$. Then, we have:

$$\lambda_n = \ln \frac{f_1(t_n)}{f_0(t_n)} = (\theta_1 - \theta_0)\, t_n + \frac{1}{2}(\theta_0^2 - \theta_1^2). \tag{2.21}$$

Based on Eqs. (2.20) and (2.21), the decision statistic Λ_N can be expressed as:

$$\Lambda_N = (\theta_1 - \theta_0) \sum_{n=1}^{N} t_n + \frac{N}{2}(\theta_0^2 - \theta_1^2). \tag{2.22}$$

Then, in SPRT, a decision is made based on the observed sequence of test statistics, $\{t_n\}_{n=1}^{N}$, using the following rule:

$$\begin{aligned} \Lambda_N \geq B &\Rightarrow \text{accept } \mathcal{H}_1 \text{ (primary signal exists)} \\ \Lambda_N < A &\Rightarrow \text{accept } \mathcal{H}_0 \text{ (no primary signal)} \\ A \leq \Lambda_N < B &\Rightarrow \text{take another observation}, \end{aligned}$$

where A and B ($0 < A < B < \infty$) are the detection thresholds that depend on the desired values of P_{FA} and P_{MD}.

Let α^* and β^* denote the desired values of false-alarm and mis-detection probabilities, respectively. Then, the decision boundaries are given by [147]:

$$A = \ln \frac{\beta^*}{1-\alpha^*} \quad \text{and} \quad B = \ln \frac{1-\beta^*}{\alpha^*}, \tag{2.23}$$

and the actual achievable error probabilities, denoted as α and β, have the following relationships:

$$\alpha \leq \frac{\alpha^*}{1-\beta^*}, \quad \beta \leq \frac{\beta^*}{1-\alpha^*}, \text{ and } \alpha + \beta \leq \alpha^* + \beta^*. \tag{2.24}$$

Eq. (2.24) indicates that the actual achievable error probabilities, i.e., α and β, can only be slightly larger than the desired values α^* and β^*. For example, with the desired values of $\alpha^* = \beta^* = 0.1$, the actual values α and β will be no larger than 0.111.

2.4.3 Sensing Delay Analysis

Recall that our goal is to minimize the number of times the spectrum needs to be sensed, with the decision thresholds derived from target detection probabilities as shown in Eq. (2.23). We therefore consider the number of sensing periods scheduled until a decision is made (i.e., either the boundary A or B is reached) as our main performance metric. The average number of sensing periods, $\mathbb{E}[N]$, required for decision-making can be computed

as:

$$\mathbb{E}[\Lambda_N] = \mathbb{E}[N] \times \mathbb{E}[\lambda \,|\, \mathcal{H}_k]. \tag{2.25}$$

First, using Eq. (2.21), the average value of λ under $\mathcal{H}_k$ can be derived as:

$$\mathbb{E}[\lambda \,|\, \mathcal{H}_k] = (\theta_1 - \theta_0)\,\theta_k + \frac{1}{2}(\theta_0^2 - \theta_1^2). \tag{2.26}$$

The average of Λ_N can then be found as follows. Suppose $\mathcal{H}_0$ holds, then Λ_N will reach B (i.e., false alarm) with the desired false alarm probability α^*; otherwise, it will reach A. Thus, using Eq. (2.23), we have:

$$\mathbb{E}[\Lambda_N \,|\, \mathcal{H}_0] = \alpha^* \,\ln \frac{1-\beta^*}{\alpha^*} + (1-\alpha^*)\,\ln \frac{\beta^*}{1-\alpha^*}. \tag{2.27}$$

Based on Eqs. (2.25), (2.26) and (2.27), we can derive the average required number of sensing periods for decision-making as:

$$\mathbb{E}[N \,|\, \mathcal{H}_0] = \frac{\alpha^* \,\ln \frac{1-\beta^*}{\alpha^*} + (1-\alpha^*)\,\ln \frac{\beta^*}{1-\alpha^*}}{(\theta_1 - \theta_0)\,\theta_0 + \frac{1}{2}(\theta_0^2 - \theta_1^2)}. \tag{2.28}$$

Similarly, we can derive:

$$\mathbb{E}[N \,|\, \mathcal{H}_1] = \frac{(1-\beta^*)\,\ln \frac{1-\beta^*}{\alpha^*} + \beta^* \,\ln \frac{\beta^*}{1-\alpha^*}}{(\theta_1 - \theta_0)\,\theta_1 + \frac{1}{2}(\theta_0^2 - \theta_1^2)}. \tag{2.29}$$

Based on Eqs. (2.19)–(2.29), Algorithm 1 describes our online algorithm for scheduling sensing periods that finds the optimal stopping time for sensing.

In practice, the number of sensing periods that can be scheduled before the fusion center makes a final decision can be upper-bounded by N_{max} due to several factors, such as the detection delay requirement, inter-sensing interval, initial sensing delay, and sensing time [132]. For example, in the 802.22 WRAN standard draft, secondary users must be able to detect the return of a primary user within 2 seconds [38]. Therefore, we set a threshold P_{th}—a design parameter—such that the fusion center must reach a conclusion within N_{max} sensing periods with a probability greater than or equal to P_{th}.

Let N_{opt} denote the optimal stopping time of sensing under Algorithm 1. Then, we

want to derive the probability of satisfying $N_{opt} \leq N_{max}$, which should be no less than P_{th}. Although an approximate expression for the distribution of N_{opt} can be derived, we instead derive a lower bound of the probability for computational efficiency [147].

Suppose $\Lambda_{N_{max}} \geq B$. Then, we have $N_{opt} \leq N_{max}$, so the following inequality holds:

$$Prob\big(N_{opt} \leq N_{max}\big) \;\geq\; Prob\big(\Lambda_{N_{max}} \geq B\big). \tag{2.30}$$

Since N_{max} is sufficiently large in practice, we can use the central limit theorem (CLT), and the inequality $\Lambda_{N_{max}} \geq B$ can then be written as:

$$\frac{\Lambda_{N_{max}} - N_{max}\,\mathbb{E}[\lambda \,|\, \mathcal{H}_1]}{\sqrt{N_{max}}\,\sigma_1(\lambda)} \geq \frac{B - N_{max}\,\mathbb{E}[\lambda \,|\, \mathcal{H}_1]}{\sqrt{N_{max}}\,\sigma_1(\lambda)}, \tag{2.31}$$

where $\sigma_1(\lambda)$ is the standard deviation of λ under $\mathcal{H}_1$, which can be derived as $\sigma_k(\lambda) = (\theta_1 - \theta_0)\ \forall k$ from Eq. (2.21). Then, the left-hand side of Eq. (2.31) is normally distributed with zero mean and unit variance when $\mathcal{H}_1$ is true.

Therefore, based on Eqs. (2.30) and (2.31), we have the following lower bound of the probability that the BS makes a decision within N_{max} observations:

$$Prob\big(N_{opt} \leq N_{max}\big) \geq Q\left(\frac{B - N_{max}\,\mathbb{E}[\lambda \,|\, \mathcal{H}_1]}{\sqrt{N_{max}}\,\sigma_1(\lambda)}\right). \tag{2.32}$$

This lower bound will be considered in our algorithm for selecting an optimal set of sensors as described next.

2.4.4 Algorithm for Joint Optimization of Sensor Selection and Sensing Time

We now turn to the problem of finding an optimal set of sensors and an optimal sensing time that together minimize average sensing overhead. Let Φ denote the *total set* of sensors in the network available for cooperative sensing with estimated RSS distributions via training. The key idea is to utilize a subset $\Omega \subseteq \Phi$ of sensors with relatively high average RSS values, and also to select the sensing time T_S, thus minimizing both the number of

Algorithm 1 ONLINE SENSING SCHEDULING

The fusion center does the following

while each round $n \in [1, N_{max}]$ of sensing period **do**
  Receive results of energy detector (i.e., RSS) $\mathbf{x}_n$ from sensors
  $t_n \leftarrow \sigma_n^{-1} \cdot \|\mathbf{w}\|^{-1} \cdot \mathcal{T}(\mathbf{x}_n)$ // Calculate test statistic
  $\Lambda_N \leftarrow \Lambda_N + (\theta_1 - \theta_0)\, t_n + \frac{1}{2}\,(\theta_0^2 - \theta_1^2)$
  if $\Lambda_N \geq B$ **then**
    A primary exists and we schedule fine-sensing (or initiate the channel vacation procedure)
  else if $\Lambda_N < A$ **then**
    A primary does not exist
  else if $n == N_{max}$ **then**
    Schedule fine-sensing for in-depth measurement
  else
    Schedule another sensing period and wait for the observation
  end if
end while

cooperating sensors and the number of sensing periods in incumbent detection, while guaranteeing the detectability requirements. Given a subset of sensors, Ω, and sensing time T_S, the total expected sensing overhead before the fusion center accumulates enough sensing samples can be expressed as:

$$\mathcal{O}(\Omega, T_S) = \min\Big\{ \max\Big\{\mathbb{E}[N(\Omega)], 1\Big\}, N_{max}\Big\} \times T_D(\Omega, T_S), \tag{2.33}$$

where $T_D(\Omega, T_S)$ is the total time duration for a single sensing, which consists of a sensing period and a measurement reporting period:

$$T_D(\Omega, T_S) = T_S + |\Omega| \times T_R, \tag{2.34}$$

where T_S is the sensing duration and T_R is the duration of a time-slot for reporting the sensing result to the fusion center.

Then, based on Eqs. (2.32), (2.33), and (2.34), our problem of finding an optimal set of

Algorithm 2 JOINT OPTIMIZATION OF SENSOR SELECTION AND SENSING TIME

1: Initialize the desired detection parameters P_{FA},P_{MD},P_{th}
2: Initialize the set of available sensors $\Phi = \{\chi_1, \ldots, \chi_{n_s}\}$
3: Initialize the optimal set of sensors $\Omega^* \leftarrow \emptyset$
4: Initialize the set of sensing time $\mathbb{T}_S \in [1, 2, 3, 4, 5]$ ms
5: Initialize the sensing overhead $\mathcal{O}^* \leftarrow \infty$
6: **while** $\Phi \neq \emptyset$ **do**
7: $\quad \chi^* \leftarrow \arg\max_{\chi_i \in \Phi}\{P_i\}$ // $P_i = P_R \cdot e^{Y_i}$
8: $\quad \Phi \leftarrow \Phi \setminus \{\chi^*\}$
9: $\quad \Omega \leftarrow \Omega^* \cup \{\chi^*\}$
10: $\quad$ **for** each $T_S \in \mathbb{T}_S$ **do**
11: $\quad\quad T_D(\Omega, T_S) \leftarrow T_S + |\Omega| \times T_R$
12: $\quad\quad N^* \leftarrow \min\{\max\{\mathbb{E}[N(\Omega, P_{FA}, P_{MD}, T_S)], 1\}, N_{max}\}$
13: $\quad\quad \mathcal{O}(T_S) \leftarrow N^* \times T_D(\Omega, T_S)$
14: $\quad$ **end for**
15: $\quad T_S^* \leftarrow \arg\min_{T_S \in \mathbb{T}_S}\{\mathcal{O}(T_S)\}$
16: $\quad \mathcal{O}_{min} \leftarrow \mathcal{O}(T_S^*)$
17: $\quad$ **if** $\mathcal{O}_{min} > \mathcal{O}^*$ and $Pr(N_{opt} \leq N_{max}) \geq P_{th}$ **then**
18: $\quad\quad$ **return** (Ω^*, T_S^*)
19: $\quad$ **else**
20: $\quad\quad \Omega^* \leftarrow \Omega$
21: $\quad\quad \mathcal{O}^* \leftarrow \mathcal{O}_{min}$
22: $\quad$ **end if**
23: **end while**

sensors can be formally stated as:

$$\begin{aligned} \text{Find} \quad & (\Omega^*, T_S^*) = \arg \min_{\Omega \subseteq \Phi, T_S \in \mathbb{T}_S} \mathcal{O}(\Omega, T_S) \\ \text{subject to} \quad & Pr(N_{opt} \leq N_{max}) \geq P_{th}, \\ & P_{FA} \leq \alpha, \\ & P_{MD} \leq \beta, \end{aligned}$$

where α and β are the desired false-alarm and mis-detection probabilities.

For this, we propose a simple algorithm as described in Algorithm 2. The idea is that we sort the sensors in descending order of average RSS (i.e., P_i) and then add sensors to Ω from the top of the list until the total sensing overhead increases by adding another sensor, and the detection constraint (i.e., P_{th}) is satisfied (line 17). The algorithm provides an optimal solution with a low computational overhead, i.e., $O(|\Phi| \cdot |\mathbb{T}_S|)$, where $|\Phi|$ and

$|\mathbb{T}_S|$ are the total number of available sensors and sensing times, respectively. On the other hand, the exhaustive search requires $O(2^{|\Phi|\cdot|\mathbb{T}_S|}-1)$. The algorithm is shown to reduce sensing overhead significantly (see Section 2.5.3), while guaranteeing the desired level of detection performance (in terms of false-alarm and mis-detection probabilities).

2.5 Performance Evaluation

This section comparatively evaluates the proposed algorithms using MATLAB-based simulation under realistic wireless environments.

2.5.1 Simulation Setup

We consider a DSA network with a large-scale primary transmitter (e.g., a TV transmitter) and multiple secondary users (or sensors). To demonstrate the efficacy of the proposed schemes in realistic wireless environments, we consider the network parameters, which are used widely in IEEE 802.22 WRANs. We assume that noise power is $NB_i = -95.2 \pm \Delta_i$ dBm, which is commonly used in IEEE 802.22 WRANs [133], where Δ_i is the noise uncertainty (in dB) at sensor i and B is channel bandwidth. The channel bandwidth is set to $B = 6$ MHz as in the TV channels. We consider shadow fading with dB-spread $\sigma_{dB} = 5.5$ (dB), which is also typical in rural area networks such as 802.22 WRANs [137]. Throughout the simulation, we assume that the time-slot duration for reporting a RSS measurement (T_R) is fixed at 0.2 ms, and that $n_s = 10$ cooperating sensors, unless specified otherwise. For RSS profiling, 10^4 samples were used for estimating the RSS distributions. This consumes only a total sensing time of 10 seconds, assuming the sensing time of $T_S = 1$ ms. We fix the desired false-alarm probability at $P_{MD} = 0.01$ and $P_{th} = 0.95$, throughout the evaluation.

To demonstrate the benefits of the proposed sensing algorithms, we evaluate the performance of the following decision- and data-fusion rules: (i) the OR-rule, (ii) Equal Gain Combining (EGC), (iii) Maximal Ratio Combining (MRC), and (iv) RSS-profile-based sensing. The OR-rule is a simple decision fusion rule, in which the fusion center con-

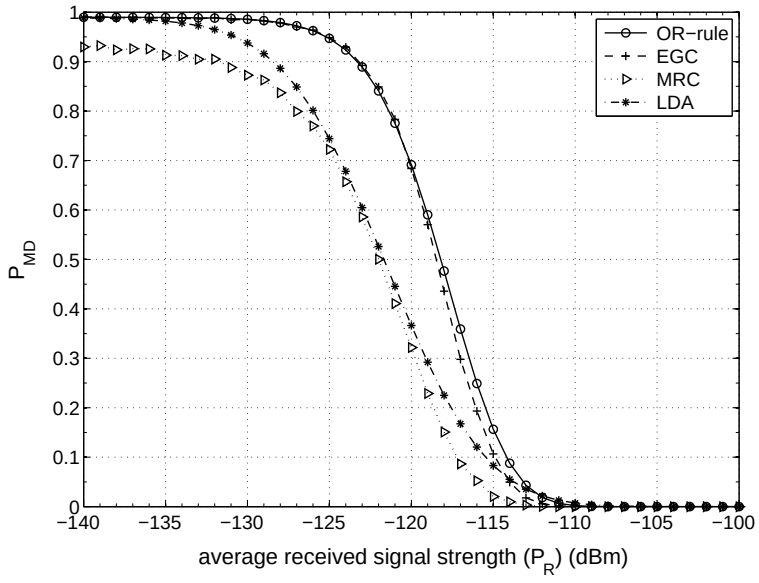


Figure 2.9: Performance of detection schemes for one-time sensing with $n_s = 10$ sensors. RSS-profile-based cooperative sensing (denoted as LDA) shows near-optimal performance even in a very low SNR environment.

cludes that there exists a primary signal if at least one sensor reports the existence of a primary signal. The other two data-fusion rules, EGC and MRC, are different in that EGC does not require any channel state information at the sensors, whereas MRC requires channel state information. EGC and MRC are known to be near-optimal in high and low SNR regions, respectively [95].

2.5.2 Performance of RSS-Profile-Based Detection for One-Time Sensing

We first evaluate the performance of the proposed LDA-based detection scheme for *one-time* sensing, assuming that the sensors are randomly selected.

2.5.2.1 Performance Comparison

Fig. 2.9 compares the performance of the testing schemes for different average RSS values. The figure shows that mis-detection probability, P_{MD}, increases as average primary signal strength, P_R, decreases for all the tested schemes. It also indicates that the OR-rule performs the worst because it does not fully utilize the sensing results; in the OR-rule,

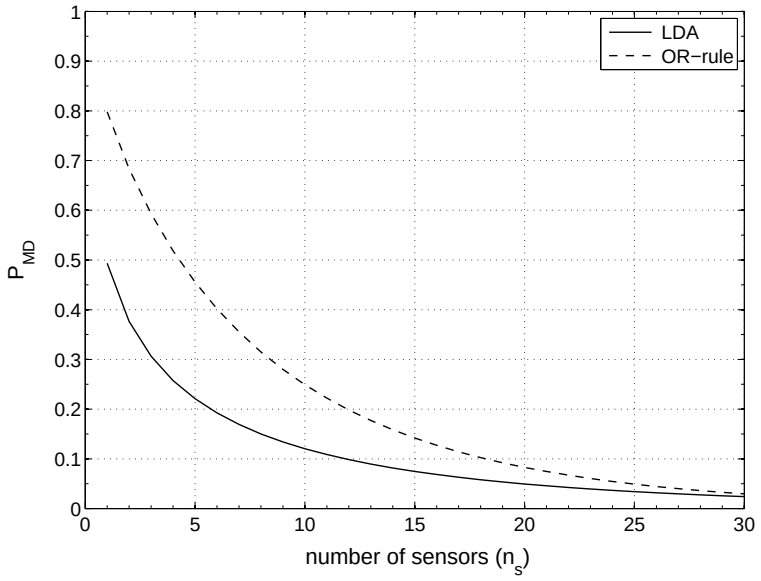


Figure 2.10: Effect of number of cooperating sensors for one-time sensing. RSS-profile-based sensing outperforms the OR-rule since it better exploits the diversity of the sensors.

sensors make a binary decision on the presence or absence of a primary signal locally and report it to the fusion center. Among the data-fusion rules, MRC outperforms EGC because MRC exploits the SNR information at the sensors, whereas EGC does not require channel estimation. It also shows that our RSS-profile-based detection (denoted as LDA) significantly outperforms the OR-rule and EGC, thanks to its ability to set the near-optimal detection threshold (i.e., an $(n_s - 1)$-dimensional hyperplane) based on the spatial RSS profile. Moreover, the performance of the RSS-profile-based scheme is close to MRC, thanks to its ability to exploit sensor heterogeneity.

2.5.2.2 Effects of Number of Sensors

Fig. 2.10 shows the impact of the number of cooperative sensors on incumbent detection performance. In our simulations, the average signal strength (P_R) is set to $P_R = -116\,\text{dBm}$, which is the DTV signal detection threshold in IEEE 802.22 WRANs. The figure shows that RSS-profile-based detection performs well even with a small number of sensors compared to the OR-rule. This is because, in RSS-profile-based detection, all the sensors contribute to the enhancement of detection performance via RSS-profile, thus fully exploiting spatial RSS diversity. On the other hand, in the OR-rule, only a few sensors with

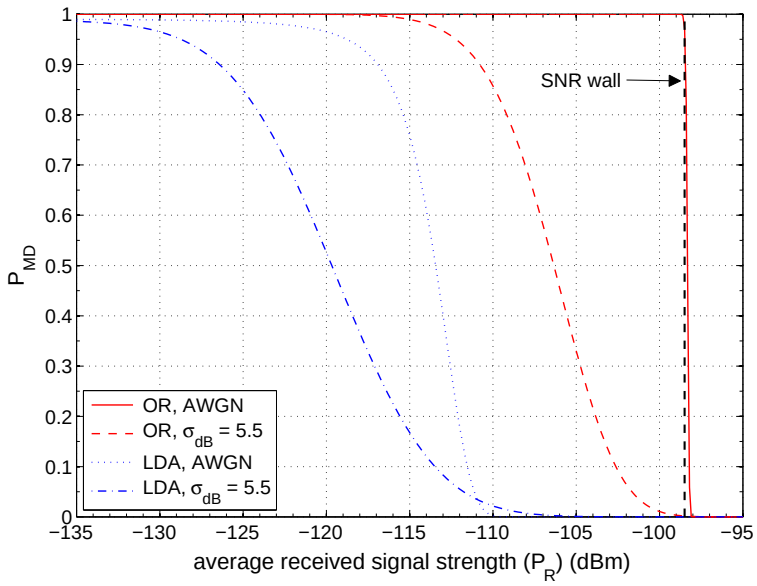


Figure 2.11: Effect of noise uncertainty for AWGN and shadow fading environments. RSS-profile-based spectrum sensing works well even when the average RSS is below SNR_{wall}, thus overcoming the noise uncertainty.

high RSSs (e.g., above the detection threshold) contribute to the detection of a primary signal.

2.5.2.3 Effects of Noise Uncertainty

Noise uncertainty is one of the main obstacles in using energy detection in a very low SNR environment such as 802.22 WRANs [142]. Noise uncertainty creates a performance barrier called SNR_{wall}, below which signal detection is infeasible irrespective of sensing time or the number of cooperative sensors. SNR_{wall}, in fact, depends solely on noise uncertainty as [142]:

$$SNR_{wall} = \frac{\rho^2 - 1}{\rho}, \tag{2.35}$$

where $\rho = 10^{\Delta/10}$ and Δ (in dB) is the noise uncertainty. We assume that noise uncertainty is bounded by 1 dB for all sensors,[4] with a corresponding SNR_{wall} of -98.5 dBm.

Fig. 2.11 shows that, when the OR-rule is employed, the detector completely fails to detect signals below SNR_{wall} under the AWGN channel. However, under the practical

[4]This is a reasonable assumption since noise uncertainty can be bounded by ± 1 (dB), considering several contributing factors such as calibration error, thermal noise variation, changes in LNA amplifier gain, etc. [130].

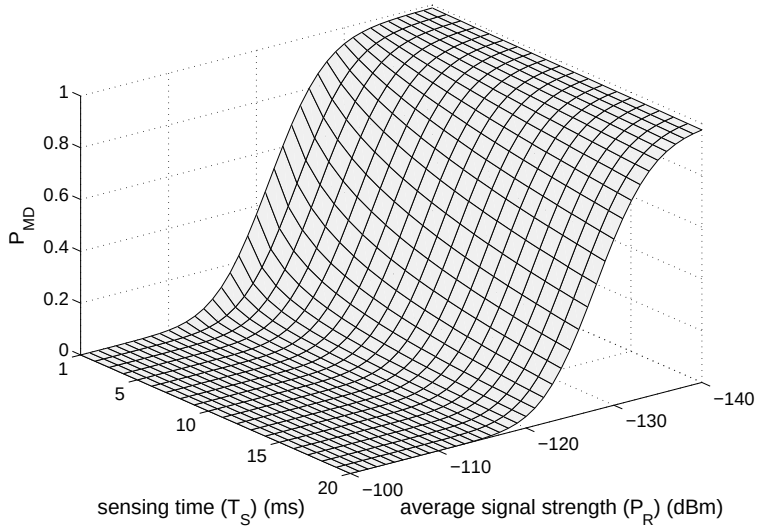


Figure 2.12: Impact of sensing time on detection performance. The amount of time a sensor spends on sensing for one-time sensing affects the detection performance.

assumption that noise uncertainty is independent at the sensors, the RSS-profile-based detection works well for signals below SNR_{wall}, even with a one-time sensing. Detection performance further improves under shadow fading, i.e., $\sigma_{dB} = 5.5$ dB, thanks to its ability to exploit spatial RSS diversity.

2.5.2.4 Effects of Sensing Time

Fig. 2.12 plots the mis-detection performance for various sensing time durations, $T_S \in [1, 20]$ ms. As we observed in Eq. (2.12), the detection performance not only depends on the average primary signal strength, P_R, but also on the sensing time. The figure shows that, as sensing time increases, the mis-detection rate decreases for all tested values of P_R. This is because the more samples the detector is provided with, the more accurate the sensing results, thus eliminating ambiguity on the existence of a primary signal. As described in Algorithm 2, the fusion center finds a combination of an optimal set of sensors and an optimal sensing time T_S^* that minimizes the average sensing overhead.

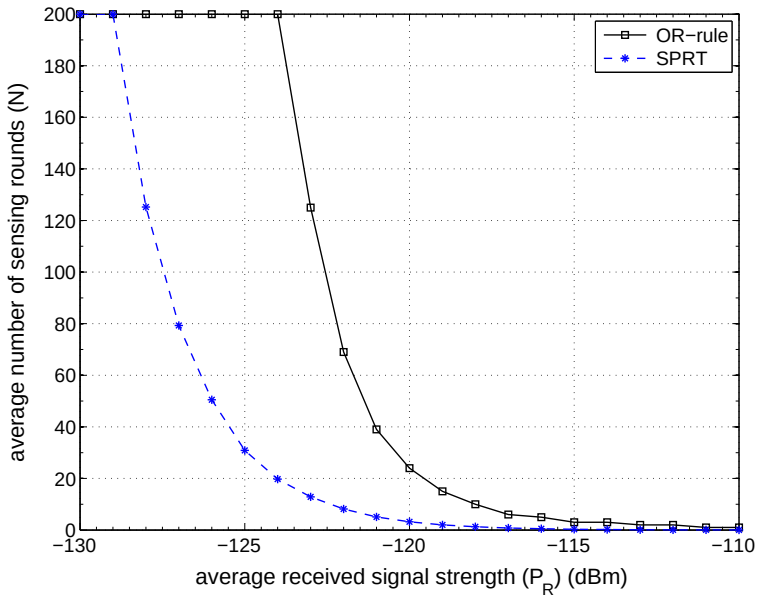


Figure 2.13: Average number of sensing periods scheduled to meet the detectability requirement of $P_{MD}, P_{FA} \leq 0.01$.

2.5.3 Performance of Online Sensing Scheduling with Optimal Sensor Selection

We now evaluate the performance of the proposed online sensing scheduling by jointly optimizing the selection of sensors and sensing time.

2.5.3.1 Impact on Incumbent Detection Delay

Fig. 2.13 shows the average number of sensing periods that must be scheduled to meet the detection requirements. The figure shows that our SPRT-based online sensing scheduling algorithm significantly reduces the average number of sensing periods compared to the OR-rule-based scheduling scheme, thanks to its ability to fully utilize the sensing results via RSS-profiles. As a result, our algorithm expands the feasible region of the energy detector significantly. On the other hand, the OR-rule benefits relatively less from scheduling sensing periods because RSSs do not change over time (except the measurement errors) at fixed sensor locations.

For the OR-rule, the false-alarm probability at each sensor, $\widetilde{P}_{FA}$, to achieve the global

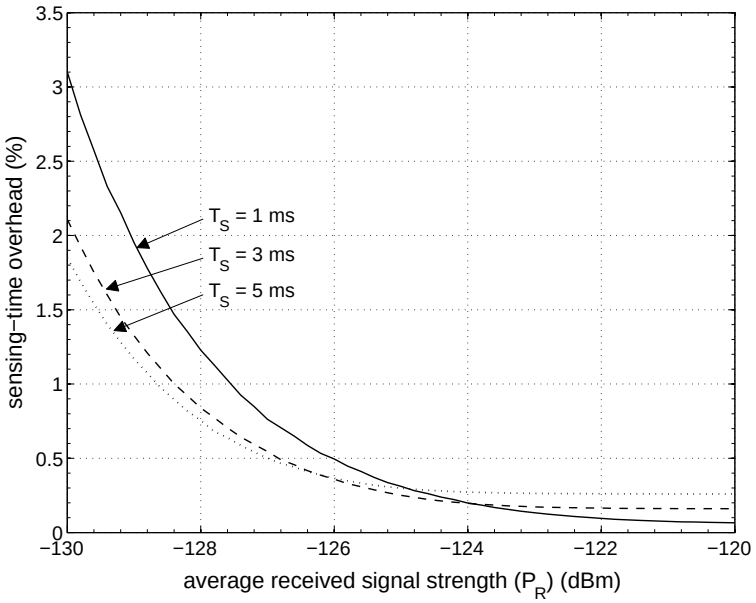


Figure 2.14: Impact of sensing time on sensing overhead under optimal sensor selection and sensing scheduling algorithms. A longer sensing time becomes more desirable as the average primary signal strength decreases, and vice versa.

false-alarm probability, P_{FA}, can be calculated as:

$$\widetilde{P}_{FA} = 1 - (1 - P_{FA})^{1/(n_s \times n_{iter})}, \tag{2.36}$$

where n_s is the number of cooperative sensors and n_{iter} is the number of sensing periods for sensing scheduling.

The detection threshold for a local decision to achieve the desired false-alarm probability $\widetilde{P}_{FA}$ can be derived from Eq. (2.36) as:

$$\eta = NB\left(1 + \frac{1}{\sqrt{M}} Q^{-1}(\widetilde{P}_{FA})\right), \tag{2.37}$$

where M is the number of sensing samples.

From Eqs. (2.36) and (2.37), the mis-detection probability for individual sensors can be derived as:

$$\widetilde{P}_{MD} = Q\left(\frac{\sqrt{M}}{P_R + NB}\left(P_R + NB - \eta\right)\right), \tag{2.38}$$

where P_R is the average primary received signal strength.

Finally, the global mis-detection probability of the OR-rule is given as $P_{MD} = \left(\widetilde{P}_{MD}\right)^{n_s \times n_{iter}}$.

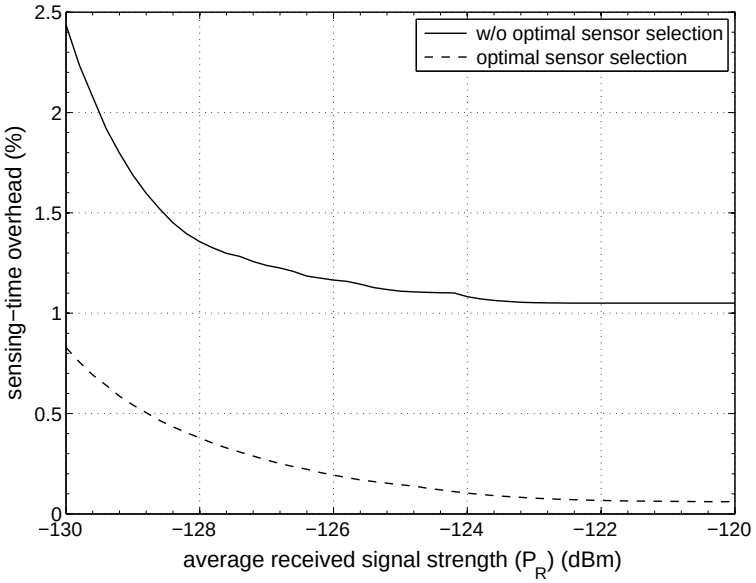


Figure 2.15: Performance of the optimal sensor selection algorithm. Our proposed sensor selection algorithm reduces sensing overhead significantly over the algorithm without sensor selection.

2.5.3.2 Impact on Sensing Overhead

We now demonstrate the performance of our optimal sensor-selection algorithm in terms of the reduction in sensing overhead. Sensing overhead is defined as the average fraction of time (in $\%$) spent on sensing within a 2 second interval, which is the *channel detection time* (CDT) period (see Eq. (2.33)). Fig. 2.14 shows the sensing time overhead for various sensing times $T_S \in [1, 3, 5]$ ms. The figure shows that a larger sensing time, i.e., $T_S = 5$ ms, is favored in low SNR environments, whereas a smaller sensing time, i.e., $T_S = 1$ ms is desirable in relatively high SNR environments, in which sensing scheduling may not be needed to achieve detection requirements.

Fig. 2.15 compares the average sensing-time overhead, i.e., the fraction of time spent on spectrum sensing and reporting the sensing results within a sensing interval, with the optimal selection of sensors and sensing time against the average sensing overhead without sensor selection. A sensing interval is assumed to be 2 s, which is equivalent to *channel detection time* (CDT) in IEEE 802.22 WRANs. The figure shows that our algorithm minimizes the average sensing overhead by up to $94\,\%$ because it selects only a subset of sensors with high average RSSs, thus minimizing both the number of sensing rounds and

the sensing result reporting time.

2.6 Related Work

Various aspects of cooperative sensing have been studied, such as cooperation gain [55, 57], sensor selection [125], and performance-overhead tradeoffs [71, 88, 102]. The benefits of sensor collaboration have been reported to diminish as the degree of shadowing correlation among sensors increases [55, 56, 57, 106]. To minimize the detrimental effects of shadowing correlation on cooperative sensing, several sensor-selection algorithms have been introduced. For example, Selén *et al.* [125] proposed heuristic algorithms for selecting an uncorrelated set of sensors, given different levels of information about sensor locations. In a similar vein, Kim and Shin [81] suggested selecting sensors based on their geographical separation, so as to make the sensors uncorrelated from each other. While these approaches seek to avoid shadow fading correlation among them, we show that, when sensors are *stationary*, there is virtually no shadowing correlation among sensors. Based on this observation, we show that shadowing correlation (or separation between sensors) is not a determining factor in stationary DSA networks.

Sensing scheduling has also been studied as an efficient way of improving incumbent detection performance [31, 67, 85, 152, 166]. For example, Hoang *et al.* [67] developed an adaptive sensing scheduling mechanism that takes into account both time-varying channel and traffic conditions. In the IEEE 802.22 WRAN standard draft, a two-stage sensing mechanism has been proposed to provide flexible scheduling of quiet periods [78]. Recently, a sequential hypothesis testing framework has been proposed as an attractive way to minimize the sensing delay for given detection requirements (in terms of false-alarm and mis-detection probabilities). Lai *et al.* [85] presented sequential detection of primary signals using the cumulative sum (CUSUM) algorithm. Similar to our work, Zou *et al.* [166] proposed a sensing scheduling scheme based on the framework of sequential probability ratio testing (SPRT) under the assumption of unknown primary signal characteristics. However, the interactions between cooperative sensing and sensing scheduling have not been considered. We on the other hand, jointly optimize the sensor selection to minimize

overall sensing overhead.

In essence, our work seeks to fill this important gap by jointly optimizing sensor selection and sensing scheduling, thus synergetically improving spectrum sensing performance. An extensive survey of cooperative sensing in DSA networks can be found in [8].

2.7 Conclusion

In this chapter, we proposed to jointly optimize cooperative sensing and sensing scheduling, in order to minimize average sensing overhead, while guaranteeing the desired level of detection requirements. Our spectrum sensing framework exploits the spatio-temporal variations in received primary signal strength by constructing a spatial RSS-profile for an incumbent signal. We showed that the RSS distribution of a primary signal can be accurately approximated as a Gaussian distribution in low SNR environments and we analyzed the detection performance of the RSS-profile-based detection scheme. We also showed that there is virtually no correlation among stationary sensors. Based on these observations, we formulated the problem of sensing scheduling as a sequential hypothesis test, which finds an optimal time to stop scheduling sensing subject to given detection requirements. We also proposed an optimal algorithm that makes the tradeoff between detection performance and sensing overhead. Our evaluation results have shown that the proposed sensing algorithms reduce sensing overhead by up to $94\,\%$ in practical scenarios.

CHAPTER 3

Secure Detection of Large-Scale Primary Users

3.1 Introduction

Accurate sensing of spectrum condition is key to the opportunistic use of licensed spectrum bands in CRNs. However, reports from sensors can be manipulated by attackers in various ways, such as primary signal emulation [11, 26] and sensing results falsification [28]. These sensing-targeted attacks can severely undermine incumbent detection performance because the fusion rule for a final detection decision relies solely on reported RSSs. Sensing-targeted attacks pose a significant threat as they can disrupt opportunistic spectrum access, the basic premise of DSA. We call these unique sensing-targeted attacks in DSA networks *sensing-disorder attacks*.

A sensing-disorder attack aims to obscure the existence/absence of a primary signal by manipulating spectrum sensing information (e.g., measured RSSs) either by raising or lowering the signal strength. When no primary signal exists, attackers or compromised sensors can manipulate their reports (i.e., RSSs) to generate the illusion of a primary signal. For example, in the IEEE 802.22 WRANs [38], an attacker can report a fake sensing report to force all users in the entire cell (of radius up to 100 km) to immediately vacate the channel [120]. Once users in the cell vacate the channel, the attacker can freely use the channel without interruption. When there is a primary signal, on the other hand, attackers can lower the RSSs to veil the presence of a primary signal, leading to an unacceptable level of interference to the PUs. In both cases, attackers mislead the fusion center, i.e., base station

(BS), to make an incorrect decision on the presence/absence of a primary signal, wasting spectrum resources or causing unacceptable interference to the primary communications. Therefore, there is a clear incentive for attackers to launch sensing-disorder attacks.

While sensing-disorder attacks can be easily launched with the aid of programmable SDR devices, their detection is difficult. Unlike ordinary Denial-of-Service (DoS) attacks that exhaust all the network resources, they can be easily mounted by using SDR devices, such as USRP [3] and Sora [84]. These open-source SDR platforms can be attractive targets for attackers because of their accessibility to low-layer protocol stacks like PHY and MAC [155]. Detecting these attacks, however, is not an easy task. While secure mechanisms such as MAC-layer or crypto-based authentication work well in traditional wireless networks, lack of primary-secondary communications precludes their usage. Moreover, the detection of attacks is exacerbated by the volatile nature of the wireless medium itself, which makes it hard to differentiate between legitimate and deliberately-manipulated sensing reports. We thus need to devise a mechanism that can protect cooperative sensing from the above-mentioned attacks.

In this chapter, we propose an attack-tolerant distributed sensing protocol (ADSP) for the IEEE 802.22 WRANs that filters out abnormal sensing reports (caused by either adversaries or malfunctioning sensors) by exploiting shadow-fading correlation in RSSs. This RSS-based filtering is motivated by the fact that attackers cannot control the physical-layer signal propagation.

3.1.1 Contributions

This chapter makes several main contributions as follows.

- Proposal of a novel *correlation filter* for detection of abnormal sensing reports that (i) exploits *shadow-fading correlation* in RSSs without any additional communication, (ii) safeguards spectrum sensing against attacks that increase either the incumbent false-alarm (type-1) or mis-detection (type-2) rates, and (iii) minimizes processing and sensing overheads. Despite their importance, type-2 attacks have not been considered before.

- Introduction of cluster-based cooperative sensing to exploit shadowing correlation. Correlation between sensors, which is entailed by sensor clustering, is known to have a detrimental impact on incumbent detection performance [55, 56, 106]. Our evaluation study, however, shows that the proposed clustering does not incur any perceivable performance degradation even in a very low SNR environment. Therefore, sensor clustering is an efficient and useful approach to sensing-disorder attacks.

- Development of a new data fusion rule tailored to attack-tolerance. Specifically, we propose *weighted gain combining* (WGC) that adaptively assigns different weights to sensing reports according to their statistical significance based on the normal shadowing profile. As a result, it minimizes the influence of unfiltered attacks (due to their small deviations) on a final decision, further improving attack-tolerance.

- Design of a sensing scheduling scheme that guarantees satisfaction of the detection requirements of 802.22 even in the presence of attacks, while minimizing the number of sensing rounds. Although ADSP significantly improves attack-tolerance, our simulation results indicate that the detection requirements of 802.22 may not be satisfied with one-time sensing. To solve this problem, we propose an optimal stopping time for sensing scheduling using sequential hypothesis testing so as to meet detectability requirements.

- In-depth evaluation of ADSP in a realistic two-dimensional shadow fading environments in IEEE 802.22 WRANs. Most previous work uses a simple but inaccurate one-dimensional model. Our simulation results show that the proposed filtering scheme successfully withstands attacks by reducing the false-alarm rate up to $99.2\,\%$ and achieving up to $97.4\,\%$ of maximum achievable detection rate.

3.1.2 Organization

The remainder of this chapter is organized as follows. Section 3.2 describes the system and attack models used in this chapter. Section 3.3 presents our proposed approach for attack detection, and the generation of a realistic two-dimensional shadowing field. Sec-

tion 3.4 details our approaches to filter design and data-fusion, and Section 3.5 proposes a sensing scheduling algorithm. Section 3.6 evaluates the performance of ADSP and Section 3.7 reviews related work. Section 3.8 concludes the chapter.

3.2 System and Attack Model

We first describe the IEEE 802.22 WRANs and the signal propagation and sensing models to be used throughout the chapter. We then introduce the data-fusion model, and finally, present the attack model.

3.2.1 IEEE 802.22 WRANs

We consider an IEEE 802.22 WRAN, an infrastructure-based cellular system where each cell consists of a BS and the associated end-users called *consumer premise equipments* (CPEs). The CPEs represent households in a rural area, and are thus stationary. The typical coverage of each 802.22 cell is 33 km (up to 100 km). The main goal of IEEE 802.22 WRANs is to provide broadband wireless access in rural areas by allowing opportunistic access of TV white spaces recently opened up by the FCC [48]. The BS, which we assume adversaries cannot compromise, schedules the sensing of channels and decides on the presence/absence of a primary signal in each channel, based on sensing reports from a set $\mathcal{C}$ of collaborating sensors. Among different types of PUs in TV bands, we focus on detecting DTV signals with 6 MHz channel bandwidth in the US. We consider an 802.22 cell located at the edge of the keep-out-radius (i.e., 150.3 km) of a TV transmitter, and the entire secondary network (or cell) lies within the detection range of the DTV signal.

3.2.2 Signal Propagation and Sensing Models

The received primary (DTV) signal strength at sensor (CPE) i can be expressed as the propagation model [59]:

$$P_i = P_o\Big(\frac{d_o}{d_i}\Big)^{\alpha} e^{X_i}, \quad \text{(Watt)} \tag{3.1}$$

where P_o is the signal strength at the primary transmitter, α the path-loss exponent, d_o the reference distance, and d_i the distance from the primary transmitter to the sensor i. Shadow fading is accounted for in e^{X_i} where $X_i \sim \mathcal{N}(0, \sigma^2)\ \forall i$. Log-normal shadow fading is often characterized by its dB-spread, σ_{dB}, which has the relationship $\sigma = 0.1 \log_e(10)\sigma_{dB}$. We assume that in the energy detector for PHY-layer sensing which measures the power level over the wide 6 MHz-wide DTV channel, the effect of multi-path fading can be ignored [127, 131] as is commonly assumed in the literature [91, 100].[1]

The energy detector is widely used for its simple design and efficiency [40, 127]. Although the feature detector is more reliable, it takes much longer (e.g., 24 ms for the field-sync detector for ATSC) [131] because it looks for a specific signature of the primary signal that appears infrequently. The test statistic of the energy detector is an estimate of average RSS (including the noise power), and can be approximated as a Gaussian using the Central Limit Theorem (CLT) as [38]:

$$T_i \sim \begin{cases} \mathcal{N}\left(N_o, \frac{N_o^2}{M}\right) & \mathcal{H}_0 \text{ (no primary signal)} \\ \mathcal{N}\left(P_i + N_o, \frac{(P_i+N_o)^2}{M}\right) & \mathcal{H}_1 \text{ (primary signal exists)}, \end{cases} \tag{3.2}$$

where P_i is the received power of a primary signal, N_o the noise power, and M the number of signal samples. We assume that sensors measure the entire 6 MHz DTV channel at the Nyquist rate for 1 ms, i.e., $M = 6 \times 10^3$.

3.2.3 Data-Fusion Model

We consider data fusion as the rule for incumbent detection. While decision fusion reduces the overhead in reporting sensing results, it is difficult to thwart sensing-disorder attacks, since it only provides a binary value based on a local decision.

In fading channels, equal gain combining (EGC) is known to have near-optimal performance without requiring estimation of the channel gains. EGC has the following decision

[1] For signal-specific sensing techniques, e.g., FFT-based pilot sensing [128], the effect of multipath fading may not be ignored.

statistic:

$$T_\Sigma \triangleq \sum_{i=1}^{N_s} w_i T_i, \tag{3.3}$$

where T_i is the test statistic of the energy detector at sensor i, N_s is the number of collaborating sensors, and the sensors have an identical weight, i.e., $w_i = 1 \, \forall i$. The decision threshold η to achieve the desired level of false-alarm probability Q^*_{FA} can be derived as [40]:

$$\eta = Q^{-1}(Q^*_{FA}) \, \frac{\sqrt{N_s} N_o}{\sqrt{M}} + N_s N_o, \tag{3.4}$$

where $Q(\cdot)$ is the well-known Q-function. The performance of EGC will be used as a baseline in evaluating the efficacy of the proposed scheme.

In order to achieve better attack-tolerance, we propose *weighted gain combining* (WGC) in ADSP that adjusts the weights $\{w_i\}_{i \in \mathcal{C}}$ so as to minimize the impact of attack misdetection on the final decision.

3.2.4 Attack Model

3.2.4.1 Attack Scenarios and Types

Sensing can be disrupted as follows.

- A sensor is compromised, and then manipulates its sensing reports, i.e., raises or lowers RSSs.
- A sensor is malfunctioning or faulty, yielding readings that differ from the actual RSS.

A common consequence of the above two cases is that sensing reports to the fusion center are distorted, thus increasing the probability that the fusion center will make a wrong decision. To solve this problem efficiently, we focus on the detection of any abnormal sensing reports instead of pinpointing the actual cause of abnormality.

Note that another possible attack scenario is a primary user emulation attack (PUEA), as studied in [11, 26, 91]. However, PUEA is relatively easy to detect mainly because the

attacker has only a coarse-grained control of RSSs at sensors since signals are broadcast. In the above two scenarios, however, the attacker has a fine-grained control of RSSs at individual sensors, making their detection more difficult. Therefore, we will focus on the above two attack scenarios.

3.2.4.2 Attack Types

We consider two types of attacks that can be mounted (or caused) by attackers (or faulty nodes):

- **Type-1 Attacks** increase the *false-positive* rate (classifying a non-primary signal or no signal as a primary signal) by raising RSSs, and
- **Type-2 Attacks** increase the *false-negative* rate (causing failure to detect a primary signal) by lowering RSSs.

We assume that the attackers know the presence/absence of a primary signal regardless of the decision made by the fusion center, and launch type-1 (type-2) attacks under $\mathcal{H}_0$ ($\mathcal{H}_1$); otherwise, attacks only serve to improve incumbent detection performance.

3.2.4.3 Sensing Reports in the Presence of Attacks

Under the above model, a final sensing report to the fusion center can be expressed (in Watt) as:

$$R_i = \underbrace{P_i \cdot \mathbf{1}_{\{\mathcal{H}_1\}} + N_o + E_i}_{\text{energy detector output } (T_i)} + D_i \quad \forall i \in \mathcal{C}, \tag{3.5}$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function, T_i is the test statistic of the energy detector (in Eq. (3.2)) including the measurement error E_i, and $D_i \in \mathbb{R}$ is the deviation or *attack strength*, tampered with by a compromised (or faulty) sensor; $D_i = 0$ for normal sensors. Note that no loss of reporting packets is assumed, so we can focus on the detection of abnormal sensing reports.

3.3 The Proposed Approach

We now present the design rationale behind ADSP, its framework, and the methodology to generate a spatially-correlated shadow fading field.

3.3.1 Design Rationale

To maximize attack-tolerance and preserve the detection accuracy of data fusion, ADSP employs anomaly detection based on statistics. Specifically, ADSP exploits physical-layer signal propagation characteristics, or the spatial correlation in RSSs among neighboring sensors. The key insight behind ADSP is that, in shadow fading environments, RSSs at nearby sensors are likely to be highly correlated, which can be used to identify manipulated sensing reports. The adversaries must be aggressive in raising or lowering the RSSs reported to the fusion center in order to influence the outcome of the final decision. However, any sensing report that significantly deviates from what is expected is deemed suspicious of being compromised or erroneous, and will hence be discarded or penalized by the fusion center in making a final decision. Adversaries must, therefore, lower their attack strength, reducing the chance for the fusion center to make a wrong decision; otherwise, they must risk getting caught by the detector. This way, the fusion center can achieve a high level of attack-tolerance, provided the majority of its neighbors behave well.

3.3.2 ADSP Framework

ADSP resides at the fusion center (i.e., BS) and consists of the following three building blocks:

- **sensing manager** that manages sensor clusters and directs the sensors to report their readings at the end of each scheduled sensing period,
- **attack detector** that detects and discards (or penalizes) abnormal sensing reports based on the pre-established shadowing correlation profile, and
- **decision maker** that determines the presence or absence of a primary signal based

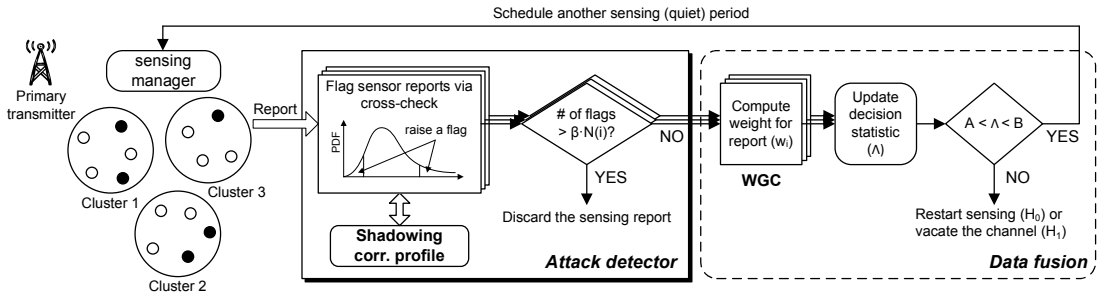


Figure 3.1: The ADSP framework: Compromised (or malfunctioning) sensors might contaminate their sensing reports $\{R_i\}$. The attack detector filters out these contaminated sensing reports based on the shadowing correlation profile and then feeds the remaining ones to the fusion center. This process is repeated until the decision statistic at the fusion-center reaches one of the predefined thresholds, i.e., A and B, in order to guarantee satisfaction of the detection requirements of 802.22.

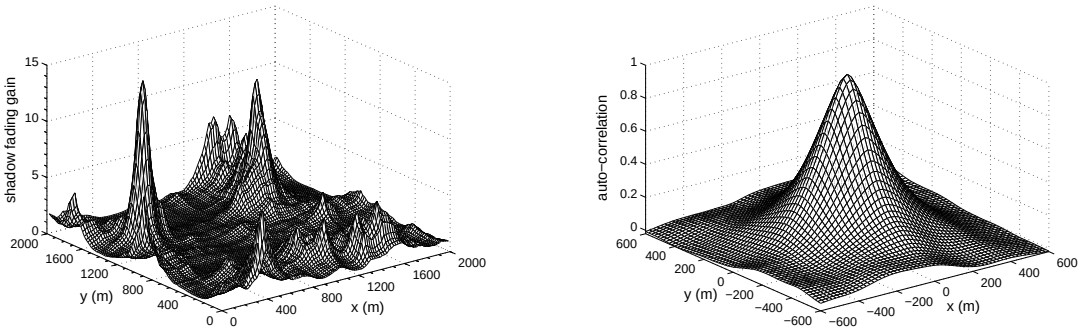
on filtered sensing results using sequential hypothesis testing.

These three components closely interact with each other and form a robust distributed sensing system. Fig. 3.1 depicts the ADSP framework, which can be implemented at the 802.22 BS without requiring any modification to sensors (i.e., CPEs).

One important and unique feature of the attack detector is the ability to tolerate *both* type-1 and type-2 attacks. This feature is attributed to the fact that the detector *cross-checks* sensing reports and the assumption that the majority of the sensors behave well. As a result, under type-1(2) attacks, their sensing reports with relatively high (low) values are likely to be flagged by more of their neighboring sensors, thus making our scheme applicable regardless of the existence of a primary signal. This makes the system design simple and efficient, while achieving high attack-tolerance.

3.3.3 Generation of Spatially-Correlated Shadow Fading

To incorporate spatially-correlated shadow fading in our analysis and simulation, we need a shadowing correlation model in which the statistics accurately reflect the real-world wireless shadowing environment. Note that one must rely on a model-based approach since measurement data for shadow fading is very scarce, and conducting a field test is too expensive. Gudmundson's model [60] is one of the most widely-used models in accounting for shadowing correlation. However, it cannot capture spatial shadowing correlation, and



(a) Shadowing random field (b) Two-dimensional auto-correlation function

Figure 3.2: Spatially-correlated shadowing random field $\mathbf{p}(\cdot,\cdot)$: (a) An example of $\mathbf{p}(\cdot,\cdot)$ with exponentially-decaying spatial correlation, where the dB-spread and decorrelation distance are assumed to be $\sigma_{dB} = 4.5\,\text{dB}$ and $D_{corr} = 150\,\text{m}$, respectively, and (b) Illustration of the two-dimensional auto-correlation function of shadow fading.

hence, analyses based on this model might yield results that are significantly different from those in real-world wireless environments, as evidenced in both the theoretical study in [109] and empirical measurements in [117]. Recently, the authors of [119] proposed a statistical modeling approach to characterization of spatial spectrum behavior of primary signals in the context of DSA networks.

Along the same line as in [119], we generate spatially-correlated shadow fading in a two-dimensional area by applying the convolution method proposed in [53]. We refer to the thus-generated data set as a *shadowing random field* $\mathbf{p}$ where $\mathbf{p}(x,y)$ represents the shadowing gain at a unit grid area, i.e., $\Delta\,\text{m} \times \Delta\,\text{m}$, centered at the coordinate $(x,y) \in \mathbb{R}^2$.

The shadowing random field $\mathbf{p}(\cdot,\cdot)$ is assumed to be an isotropic,[2] wide-sense stationary, and log-normally distributed random field with zero mean and exponentially-decaying spatial correlation. Then, the covariance between the two points $\boldsymbol{\theta}_i = (x_i, y_i)$ and $\boldsymbol{\theta}_j = (x_j, y_j)$ in $\mathbf{p}$ is given as:

$$\mathbb{E}\big[\mathbf{p}(\boldsymbol{\theta}_i), \mathbf{p}(\boldsymbol{\theta}_j)\big] = R_{\mathbf{p}}(d_{ij}) = \sigma^2 \cdot e^{-d_{ij}/D_{corr}}, \tag{3.6}$$

where $d_{ij} = \|\mathbf{p}(\boldsymbol{\theta}_i) - \mathbf{p}(\boldsymbol{\theta}_j)\|$ is the Euclidean distance between the locations $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_j$, σ is the standard deviation of shadow fading, and D_{corr} is the decorrelation distance, which

[2]Note that we do not consider angular dependency in shadowing correlation for analytical tractability.

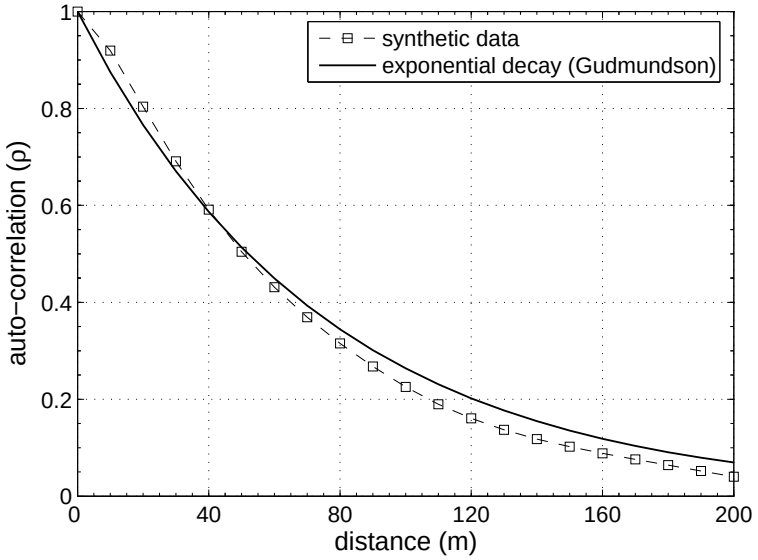


Figure 3.3: Comparison of auto-correlation function: Theoretical model (solid line) vs. synthetic data from a random field $\mathbf{p}(\cdot,\cdot)$ (dotted line).

depends on local wireless environments (e.g., urban or suburban).[3]

Fig. 3.2(a) shows an example of a *shadowing random field* in a $2\,\text{km} \times 2\,\text{km}$ region, which clearly exhibits a strong spatial correlation in shadow fading. This is clearly shown in Fig. 3.2(b), which depicts the two-dimensional auto-correlation of shadow fading. To demonstrate the accuracy of this method, Fig. 3.3 compares the one-dimensional auto-correlation function (ρ) of a random field against the Gudmundson's empirical model with $\sigma_{dB} = 4.5\,\text{dB}$ and $D_{corr} = 150\,\text{m}$. The figure indicates that the synthetic data in the shadowing random field accurately emulates real-world shadowing correlations. Note that our attack detection scheme in ADSP only requires the one-dimensional auto-correlation function of the shadowing field, which can be estimated by the service provider at the time of system deployment.

[3]The measurement study in [10] indicates that a typical decorrelation distance is in the range of $120 - 200\,\text{m}$ in suburban areas.

3.4 Detection of Abnormal Sensor Reports via Correlation Analysis

In this section, we formulate the anomaly-detection problem as a hypothesis testing, and present the design of a correlation-based filter. To further improve the attack-tolerance of ADSP, we propose a new data-fusion rule, called the *weighted gain combining* (WGC).

For cooperative sensing, the designated sensors (grouped in clusters) report their energy-detector's output along with their location information to the fusion center, at the end of each sensing period.[4] The location information is required to exploit the shadowing correlation in RSSs; it may be available at the fusion center since the sensors (i.e., CPEs) in 802.22 are stationary and 802.22 standard draft mandates the BS to have sensors' location information. Sensors can employ existing secure localization protocols (e.g., [27, 116]) to obtain accurate sensor location information.

3.4.1 Characterization of the Correlation in Sensing Reports

We first study the correlation structure of the sensing reports. A key observation is that the correlation structure of shadowing components $\{e^{X_i}\}$ is preserved in the sensing reports $\{R_i\}$ when there is no attack (or misbehavior), i.e., $D_i = 0$. To simplify the analysis, we further assume that the variance of the measurement error can be approximated as $\sigma_E^2 \approx \frac{N_o^2}{M}$ regardless of the presence/absence of a primary signal.[5]

Under the above conditions, and treating all the other terms in Eq. (3.1) (except e^{X_i} and E_i) as constants, we can express sensor i's report in Eq. (3.5) as:

$$R_i = C_1\, e^{X_i} + C_2 + E_i \quad \text{(Watt)}, \tag{3.7}$$

where $C_1 = P_o\big(d_o/d_i\big)^{\alpha}$, $C_2 = N_o$, and $E_i \sim \mathcal{N}(0, \frac{N_o^2}{M})$ is the measurement error of the energy detector. The correlation in shadowing component e^{X_i} does not change when we

[4] We consider two-dimensional sensor coordinates for simplicity, while the actual terrain profile is three-dimensional.

[5] This assumption is reasonable in a very low SNR environment, e.g., -20 dB, where the average primary signal power is only about 1 % of the noise power, i.e., $\mathbb{E}[P_i] = 0.01 \times \mathbb{E}[N_o]$.

add/multiply the same number to all of the shadowing components.

Moreover, the variance of measurement error is much smaller than that of a shadowing component, i.e., $\sigma_E^2 < \sigma_X^2$, since the number of samples M is sufficiently large even with a short sensing time, e.g., $M = 6 \times 10^3$ for the duration of 1 ms. So, the correlation in the received sensing reports $\{R_i\}$ almost preserves the correlation of the shadow fading e^{X_i}, i.e., $Corr(R_i, R_j) \approx Corr(e^{X_i}, e^{X_j})$.

3.4.2 Cluster-based Hypothesis Testing

While we exploit shadowing correlation for attack detection, the degree of correlation decreases exponentially with the distance between sensors. Therefore, we form *sensor clusters* among the sensors in close proximity, such that sensors within the same cluster are highly correlated. A measurement study in [124] indicates that households in rural areas tend to be clustered, and thus, it is reasonable to assume that a BS can identify several sensor (i.e., CPE) clusters within its own cell of typical radius of 33 km. If such sensor clusters exist, the BS can easily identify them based on their location information. If such sensor clusters do not exist, additional sensors can be deployed to form such sensor clusters.

Therefore, for each collaborative sensor $i \in \mathcal{C}$, the correlation-filter checks if the sensor exhibits proper correlation behavior based on the following hypothesis testing for each neighbor within its cluster:

$$\mathcal{H}_0^a : Corr(R_i, R_j) = \rho(d_{ij}) \quad \forall j \in N(i), \tag{3.8}$$

where the neighbor set $N(i)$ is defined as the sensors belonging to the same cluster of sensor i. As a result of this cross-checking, the number of flags raised by neighboring sensors will be used as a filtering criterion (see Section 3.5.3 for details). We will henceforth focus on the analysis of shadowing correlation in sensing reports.

3.4.3 Correlation Analysis for Filter Design

Although the shadowing correlation coefficient (ρ) is an obvious metric for the above hypothesis testing (i.e., Eq. (3.8)), it is not suitable for direct use in our problem because estimation of the correlation coefficient would require a sequence of samples; this can incur significant time and energy overhead for sensing, and can also deter the detection of returning PUs. Therefore, we detect per-sample abnormal behavior by examining *similarity* between the sensing reports using their conditional probability distributions. This is an alternative, but efficient approach since higher correlation entails greater similarity, which can be measured via a conditional distribution of sensor reports, as we will describe next.

In order to capture the similarity between sensing reports, we first derive the probability distribution of R_i, which is the sum of non-zero mean normal (i.e., E_i) and log-normal (i.e., e^{X_i}) random variables, as indicated in Eq. (3.7). To the best of our knowledge, there is no closed-form expression for such a distribution. However, a close examination of Eq. (3.7) implies that R_i can be approximated as a *shifted log-normal random variable*, i.e., the sum of a log-normal random variable and a constant.

Let us denote the sensing reports by a shifted log-normal random variable, i.e., $R_i = e^{Z_i} + N_o + C$ where $Z_i \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$. From Eq. (3.7), we have the following approximation after simple manipulation:

$$e^{Z_i} + N_o + C \approx e^{X_i + \ln C_1} + N_o + E_i, \tag{3.9}$$

where $Z_i \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$ and $X_i \sim \mathcal{N}(0, \sigma_X^2)$ with $\sigma_X = \sigma$. We set the constant $C = 4\,\sigma_E$ where $\sigma_E = \frac{N_o}{\sqrt{M}}$ so that the probability of the right-hand side of Eq. (3.9) becomes less than C is close to zero (i.e., $\approx 3 \times 10^{-5}$). This is important to preserve the non-negativeness of the log-normal random variable e^{Z_i}.

Then, we estimate the mean and variance of e^{Z_i} using a moment-matching method. By matching the mean and variance of both sides of Eq. (3.9), we have:

$$\hat{\sigma}_Z^2 = \log\left[\frac{C_1^2\,(e^{\sigma_X^2} - 1)\,e^{2\mu_X + \sigma_X^2} + \sigma_E^2}{(C_1\,e^{\mu_X + \sigma_X^2/2} + \mu_E + C)^2} + 1\right], \tag{3.10}$$

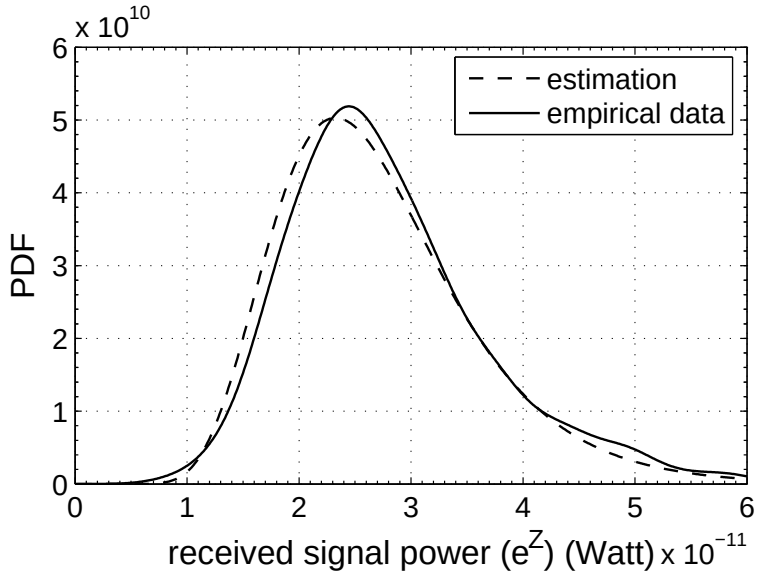


Figure 3.4: Estimation of the distribution of sensing reports as a shifted log-normal distribution: The empirical data for sensing reports (solid line) obtained from the shadowing field can be accurately approximated as a log-normal distribution (dashed line).

and

$$\hat{\mu}_Z = \log\left[\frac{C_1\, e^{\mu_X+\sigma_X^2/2} + \mu_E + C}{e^{\hat{\sigma}_Z^2/2}}\right]. \tag{3.11}$$

The derivations of Eqs. (3.10) and (3.11) are straightforward, and thus omitted due to space limitation.

Fig. 3.4 shows an example of such approximation. While the figure indicates that the sensing reports can be accurately estimated by such a distribution, it becomes less accurate as the sensing duration T_S increases. Note, however, that we want to capture the correlation among sensors in a tractable form, not necessarily as an accurate approximation only complicates the analysis without yielding a noticeable improvement in detection performance. The impact of the approximation error will be discussed in Section 3.6.

Based on Eqs. (3.9), (3.10), and (3.11), the p.d.f. of a sensor report can be expressed as:

$$f_R(r) = \frac{1}{(r-C)\,\sigma_Z\sqrt{2\pi}} \exp\left[-\frac{(\ln(r-C)-\mu_Z)^2}{2\,\sigma_Z^2}\right], \quad z \geq 0. \tag{3.12}$$

Recall that we are interested in studying the similarity of sensing reports measured at nearby (thus spatially-correlated) sensors. To measure the similarity exists between sensing reports, we derive the conditional p.d.f. of sensor i's report R_i given the neighboring sensor

j's report $R_j = r_j$ using Eq. (3.12) as:

$$f_{R_i|R_j}(r_i|r_j) = \frac{f_{Z_i,Z_j}(z_i,z_j)}{f_{Z_j}(z_j)} = \frac{1}{(r_i - C)\,\sigma_{R_i|R_j}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln(r_i - C) - \mu_{Z_i|Z_j}}{\sigma_{Z_i|Z_j}}\right)^{1/2}\right], \quad (3.13)$$

where

$$\mu_{Z_i|Z_j} = \mu_{Z_i} + \rho_{ij}\frac{\sigma_{Z_i}}{\sigma_{Z_j}}\big[\ln(r_j - C) - \mu_{Z_j}\big] \quad (3.14)$$

and

$$\sigma_{Z_i|Z_j} = \sigma_{Z_i}\sqrt{1 - \rho_{ij}^2(d_{ij})}. \quad (3.15)$$

Eq. (3.15) indicates that standard deviation $\sigma_{Z_i|Z_j}$ decreases as the correlation ρ_{ij} increases, and thus greater similarity between sensing reports.

Eqs. (3.13), (3.14), and (3.15) indicate that the conditional distribution of sensing reports is also log-normally distributed. We thus set the lower and upper thresholds on the sensing reports based on the conditional p.d.f. in Eq. (3.13), and then mark any outlier that resides outside of the thresholds. To set the thresholds, we first derive the cumulative distribution function (c.d.f.) of sensor i's report r_i, given sensor j's report r_j as:

$$F_{R_i|R_j}(x) = Pr(R_i \le x \,|\, R_j = r_j) = \frac{1}{2} + \frac{1}{2}\,\text{erf}\left[\frac{\ln(x - C) - \mu_{Z_i|Z_j}}{\sigma_{Z_i|Z_j}\sqrt{2}}\right], \quad x \ge 0, \quad (3.16)$$

where $\text{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\mathrm{d}t$.

Using Eq. (3.16), the thresholds $TH_{\{L,U\}}$ with a $100\times(1-\epsilon)\,\%$ confidence interval can be derived as:

$$TH_{\{L,U\}}(\epsilon) = \exp\Big[\sqrt{2}\cdot \text{erf}^{-1}\big(g(\epsilon)\big)\cdot\sigma_{Z_i|Z_j} + \mu_{Z_i|Z_j}\Big] + C, \quad (3.17)$$

where

$$g(\epsilon) = \begin{cases} \epsilon - 1 & \text{for } TH_L \\ 1 - \epsilon & \text{for } TH_U \end{cases} \quad 0 \le \epsilon \le 0.5, \quad (3.18)$$

where $\mu_{Z_i|Z_j}$ and $\sigma_{Z_i|Z_j}$ are the conditional mean and standard deviation in Eqs. (3.14) and

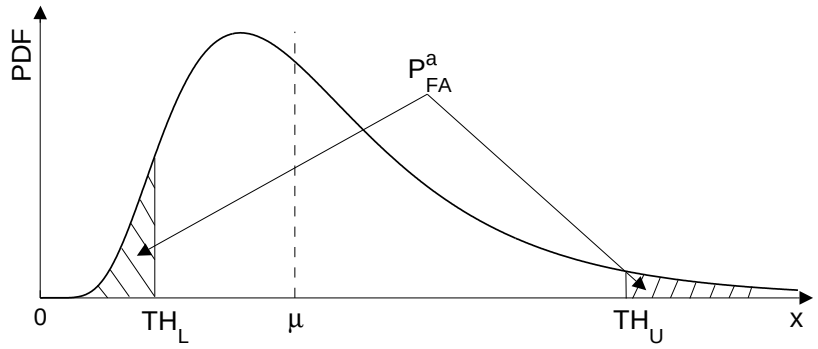


Figure 3.5: The correlation filter for anomaly detection: Sensor i's report r_i will be flagged if it resides outside of the lower and upper thresholds, i.e., TH_L and TH_U.

(3.15), respectively.

Therefore, the null hypothesis $\mathcal{H}_0^a$, i.e., $Corr(R_i, R_j) = \rho(d_{ij})$, cannot be rejected if $r_i \in [TH_L, TH_U]$, as depicted in Fig. 3.5, whereas the attack false-alarm probability can be calculated as $P_{FA}^a = Pr(r_i < TH_L) + Pr(r_i > TH_U)$. Note that the thresholds are set differently for neighboring sensors, depending on their relative distance and measured RSSs.

Clearly, there is a tradeoff in determining the threshold parameter ϵ, i.e., the higher the threshold, the higher (lower) the false-alarm (mis-detection) rate for attack detection. The impact of the thresholds on incumbent detection performance will be studied in Section 3.6.

3.4.4 The Proposed Data-Fusion Rule

While the correlation filter accurately detects RSS deviations in sensing reports, we observed that it often mis-detects small deviations (e.g., ≤ 0.3 dB). These small deviations can still influence the data-fusion results in a very low SNR environment due to the high sensitivity of the fusion decision to RSSs. Therefore, as a second line of defense, we propose a new data-fusion rule, namely *weighted gain combining* (WGC), to provide a better attack-tolerance to such small deviations. The idea is to assign different weights to the sensing reports according to their significance level based on the conditional c.d.f. in Eq. (3.16). This way, the mis-detected (unfiltered) attacks are highly likely to be assigned relatively small weights compared to legitimate sensing reports because of their lack of

significance. Thus, the weights in WGC are defined as:

$$w_i \triangleq \frac{\sum_{j \in N_v(i)} w_{ij}}{|N_v(i)|} \text{ where } w_{ij} = 1 - 2\left|F_{R_i|R_j}(r_i \,|\, r_j) - 0.5\right|, \tag{3.19}$$

where $N_v(i)$ is the set of valid neighbors of sensor i whose reports passed the filter. The thus-obtained weights are used in calculating the decision statistic.

The simulation results (in Section 3.6) show that the WGC for data-fusion significantly reduces the attack false-alarm and mis-detection probabilities. However, the results also indicate that the detectability requirement of 802.22, i.e., $Q_{FA}, Q_{MD} \leq 0.1$, might not be met under weak attack strengths (e.g., ≤ 0.3 dB) as they cannot be easily differentiated from the normal sensing reports. To remedy this and to meet the detectability requirements of 802.22 regardless of attack strengths, we next present a sequential hypothesis testing framework for sensing scheduling.

3.5 The Proposed Data-Fusion Rule via Sequential Hypothesis Testing

In this section, we first formulate the incumbent detection problem as a sequential hypothesis testing, subject to the detection requirements of 802.22. We then provide a description of ADSP.

3.5.1 Attack-Tolerant Sensing Scheduling via SPRT

In ADSP, the BS schedules the sensing periods (stages) until it obtains a sufficient amount of information for making a final decision. Thus, the BS receives a sequence of measured test statistics from the sensors. This makes sequential detection suitable for our problem. In particular, among various sequential detection techniques, we adopt Wald's *Sequential Probability Ratio Test* (SPRT) [147] since it is optimal in the sense of minimizing the average number of observations, given bounded false-alarm probability Q_{FA} and mis-detection probability Q_{MD}. Therefore, by adopting the SPRT along with WGC, the

BS can meet the detection requirement of 802.22 under the existence of malicious sensors by carefully designing the decision statistic as we discuss next.

3.5.1.1 Design of Decision Statistic

For SPRT, the distributions of the weighted test statistics of the sensors that passed the filter should be available to the BS under both hypotheses. In practice, however, it is not feasible to derive a closed-form expression for such distributions. Therefore, instead of relying on the exact distributions of T_Σ, we exploit the threshold property of T_Σ as our main decision criterion.

Let ϑ_n denote a Bernoulli random variable defined as:

$$\vartheta_n \triangleq \begin{cases} 0 & \text{if } T_{\Sigma,n} \leq \eta_n \\ 1 & \text{if } T_{\Sigma,n} > \eta_n, \end{cases} \tag{3.20}$$

where $T_{\Sigma,n}$ is the sum of test statistics from the valid sensors, i.e., those who passed the filter, in sensing stage n, and η_n is the decision threshold, which depends on the number of valid sensing reports and the desired false-alarm probability Q^*_{FA} (see Eq. (3.4) in Section 3.2).

Our detection problem is thus a binary Gaussian classification problem where the observed test statistic $\vartheta_n \ \forall n$ belongs to one of two classes, $\mathcal{H}_0$ or $\mathcal{H}_1$, where:

$$\begin{aligned} &\mathcal{H}_0 : \vartheta \sim Bernoulli(\phi_0) \quad \text{(no primary signal)} \\ &\mathcal{H}_1 : \vartheta \sim Bernoulli(\phi_1) \quad \text{(primary signal exists)}, \end{aligned}$$

When there is no attack, the random variables ϕ_0 and ϕ_1 can be defined as:

$$\phi_0 \triangleq Pr(\vartheta_n = 1 \,|\, \mathcal{H}_0) = Q^*_{FA}, \tag{3.21}$$

$$\phi_1 \triangleq Pr(\vartheta_n = 1 \,|\, \mathcal{H}_1) = Q^*_D = 1 - Q^*_{MD}. \tag{3.22}$$

In this case, there should be a significant difference between ϕ_0 and ϕ_1, i.e., $\phi_1 \gg \phi_0$.[6] However, the actual achievable Q_{FA} and Q_D under attack scenarios can be higher and lower than the desired values, respectively, due to performance deficiency of the filter. For example, Fig. 3.9 in Section 3.6 indicates that $\phi_1 - \phi_0$ can be as low as 0.08 under weak attacks, thus rendering it difficult for the BS to make a correct decision.

Therefore, ϕ_0 and ϕ_1 are the key parameters in our design of SPRT, which must be carefully set so as to meet the detection requirements of 802.22 under various attack scenarios. Thus, we set:

$$\phi_0^{'} = Q_{FA}^{*} + \varepsilon_0 \quad \text{and} \quad \phi_1^{'} = Q_D^{*} - \varepsilon_1, \tag{3.23}$$

where $\varepsilon_0, \varepsilon_1 \in \mathbb{R}$ with the constraint $\phi_1^{'} > \phi_0^{'}$.

We set the values of ε_0 and ε_1 empirically, based on the observations made in our simulation study. Note that inaccurate values of $\phi_0^{'}$ and $\phi_1^{'}$ might introduce additional detection delay. However, as long as $\phi_0^{'}$ used by the BS is closer to the true distribution under $\mathcal{H}_0$ than $\phi_1^{'}$, or vice versa, the SPRT will terminate with the desired level of detection probabilities.

3.5.1.2 Optimal Stopping Rule for Sensing Scheduling

In SPRT, a decision is made based on the observed sequence of test statistics, $\{\vartheta_n\}_{n=1}^{N}$, using the following rule:

$$\begin{aligned} \Lambda_N \geq B &\Rightarrow \text{accept } \mathcal{H}_1 \text{ (primary signal exists)} \\ \Lambda_N < A &\Rightarrow \text{accept } \mathcal{H}_0 \text{ (no primary signal)} \\ A \leq \Lambda_N < B &\Rightarrow \text{take another observation,} \end{aligned}$$

where A and B ($0 < A < B < \infty$) are the detection thresholds that depend on the desired values of Q_{FA} and Q_{MD}. The decision statistic Λ_N is the log-likelihood ratio based on N

[6] For example, the detection requirement of 802.22 is $\phi_1 - \phi_0 = Q_D^{*} - Q_{FA}^{*} = 0.9 - 0.1 = 0.8$.

sequential observations (i.e., test statistics) $\vartheta_1, \ldots, \vartheta_N$ as:

$$\Lambda_N \triangleq \lambda(\vartheta_1, \ldots, \vartheta_N) = \ln \frac{Pr(\vartheta_1, \ldots, \vartheta_N \,|\, \mathcal{H}_1)}{Pr(\vartheta_1, \ldots, \vartheta_N \,|\, \mathcal{H}_0)}. \tag{3.24}$$

Assuming that $\{\vartheta_n\}_{n=1}^N$ are i.i.d., Eq. (3.24) becomes:

$$\Lambda_N = \sum_{n=1}^{N} \lambda_n = \sum_{n=1}^{N} \ln \frac{Pr(\vartheta_n \,|\, \mathcal{H}_1)}{Pr(\vartheta_n \,|\, \mathcal{H}_0)} \tag{3.25}$$

Eq. (3.25) can be rewritten as:

$$\Lambda_N = s_N \ln \frac{\phi_1^{'}}{\phi_0^{'}} + (N - s_N) \ln \frac{1 - \phi_1^{'}}{1 - \phi_0^{'}}, \tag{3.26}$$

where $s_N = \sum_{n=1}^{N} \mathbf{1}_{\{\vartheta_n = 1\}}$ denotes the number of sensing stages n where $\vartheta_n = 1$.

3.5.2 Performance Analysis

We now quantify the performance of our SPRT-based sensing scheduling in terms of (i) detection performance, i.e., Q_{FA} and Q_{MD}, and (ii) average number of sensing rounds needed to meet the detectability requirements.

3.5.2.1 Detection Performance

In SPRT, the desired detection performance can be guaranteed by setting the decision thresholds A and B as follows. Let a^* and b^* denote the desired values of Q_{FA} and Q_{MD}, respectively. Then, the decision boundaries A and B are given by [147]:

$$A = \ln \frac{b^*}{1 - a^*} \quad \text{and} \quad B = \ln \frac{1 - b^*}{a^*}, \tag{3.27}$$

and the actual achievable error probabilities, denoted as a and b can only be slightly larger than the desired values a^* and b^*.

3.5.2.2 Sensing Scheduling Overhead

Recall that our objective is to meet the detection requirements of 802.22 even in the presence of malicious/mal-functioning sensors. Thus, we aim to minimize the number of times the spectrum needs to be sensed, with the decision thresholds derived from the target detection probabilities as shown in Eq. (3.27). Therefore, we are interested in analyzing the number of sensing rounds until a decision is made (i.e., either the boundary A or B is reached).

The average number of sensing rounds (also called quiet periods in 802.22) required to make a decision, denoted by $\mathbb{E}[N]$, can be computed as:

$$\mathbb{E}[N] = \mathbb{E}[\Lambda_N]^{-1} \times \mathbb{E}[\lambda \,|\, \mathcal{H}_k]. \tag{3.28}$$

First, using Eq. (3.26), the average value of λ under both hypotheses can be derived as:

$$\mathbb{E}[\lambda \,|\, \mathcal{H}_0] = \mathbb{E}\Big[\ln \frac{1-\phi_1'}{1-\phi_0'}\Big] \quad \text{and} \quad \mathbb{E}[\lambda \,|\, \mathcal{H}_1] = \mathbb{E}\Big[\ln \frac{\phi_1'}{\phi_0'}\Big] \tag{3.29}$$

Next, the average of Λ_N can be found as follows. Suppose $\mathcal{H}_0$ holds, then Λ_N will reach B (i.e., false alarm) with the desired false-alarm probability a^*; otherwise, it will reach A. Thus, using Eq. (3.27), we get:

$$\mathbb{E}[\Lambda_N \,|\, \mathcal{H}_0] = a^* \ln \frac{1-b^*}{a^*} + (1-a^*) \ln \frac{b^*}{1-a^*}. \tag{3.30}$$

Based on Eqs. (3.28), (3.29) and (3.30), we can derive the average number of required sensing rounds for decision-making as:

$$\mathbb{E}[N \,|\, \mathcal{H}_0] = \frac{a^* \ln \frac{1-b^*}{a^*} + (1-a^*) \ln \frac{b^*}{1-a^*}}{\mathbb{E}\Big[\ln \frac{1-\phi_1'}{1-\phi_0'}\Big]}. \tag{3.31}$$

Similarly, we can derive $\mathbb{E}[N \,|\, \mathcal{H}_1]$.

3.5.3 Protocol Description

We now present the attack-tolerant distributed sensing protocol (ADSP) with the proposed WGC for final fusion. Algorithm 4 describes the overall data-fusion procedure in ADSP. At the end of each sensing period, the fusion center collects sensing reports $\{R_i\}$ from the collaborating sensors, which are co-located in clusters. Then, the fusion center activates the correlation filter to selectively discard abnormal sensing reports and updates the decision statistic Λ_n based on the remaining sensing reports with their weights. Note that the weights are assigned after the filtering process (line 11) so that the filtered abnormal sensing reports have no influence on them. The fusion center repeats this process until the decision statistic reaches one of the predefined thresholds, i.e., A and B.

Algorithm 4 details the filtering procedure. For each sensing report, the filter counts the number of flags raised by neighbors in the cluster. Then, the filter will return $\texttt{Isnormal} = 0$ if more than $\beta \in [0, 1]$ fraction of the neighboring sensors mark it as abnormal, where β is a design parameter; otherwise, it will return $\texttt{Isnormal} = 1$. The filter also returns the weight vector ($\mathbf{w}_i$) for future use in the final data-fusion process (i.e., WGC). The computational complexity of the algorithm is bounded by $\mathcal{O}(m^2)$ where m is the number of sensors in a cluster.

Remark: Although the key assumptions we have made, i.e., negligible multipath fading and presence of sensor clusters, are valid for the DTV signal detection in IEEE 802.22 WRANs, they might not always hold, depending on a given DSA environment, thus limiting the practicality of ADSP. For example, multipath fading in sensing reports may be negligible when sensors are mobile, or a primary signal is sensed with narrow channel bandwidth. However, relaxation of such assumptions may require a major modification to ADSP, and thus, extension of ADSP to such challenging environments is left as our future work.

3.6 Performance Evaluation

The performance of ADSP is evaluated via MATLAB-based simulations. We first describe the simulation setup and then present the simulation results for both types of attacks

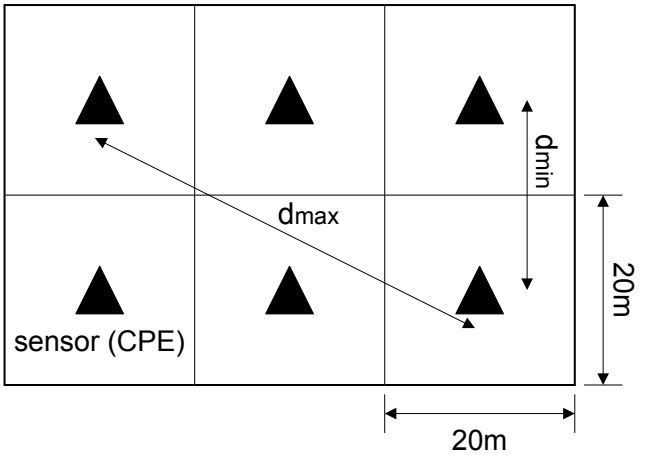


Figure 3.6: Sensor cluster: An illustration of sensor cluster with 6 sensors in an 802.22 WRAN cell.

under various attack scenarios.

3.6.1 Simulation Setup

To demonstrate the effectiveness of ADSP, we consider an IEEE 802.22 WRAN environment with a single DTV transmitter with 6 MHz bandwidth and multiple sensors (i.e., CPEs) located at the edge of the *keep-out radius* of 150.3 km from the DTV transmitter [127]. An 802.22 cell of radius 30 km is considered for our evaluation, and we generate a two-dimensional shadowing field (using the method discussed in Section 3.3.3) with a unit grid of $20 \times 20\,\text{m}^2$ to emulate a realistic shadow fading environment in a cell. Throughout the simulation, we assume 5 sensor clusters located randomly within the cell, with 6 sensors in each cluster; the sensors are located in different grids, and the distances between sensors within a cluster range from $d_{min} = 20\,\text{m}$ to $d_{max} = 20\sqrt{5}\,\text{m}$, as shown in Fig. 3.6. We consider the attack scenario where one-third of the sensors are malicious in each cluster. Table 3.1 lists the system parameters used in our simulation. Each simulation is run on 5×10^4 randomly-generated shadowing fields, and their average values are taken as the performance measures.

Table 3.1: System parameters used in simulations

Parameter	Value	Comments
N_s	30	Number of collaborating sensors
N_c	5	Number of clusters
T_S	1 ms	Sensing duration
M	6×10^3	# of signal samples per sensing
σ_{dB}	4.5 dB	Shadow fading dB-spread
D_{corr}	150 m	Decorrelation distance
Δ	20 m	Dimension of a grid
N_o	-95.2 dBm	Noise power
γ	-20 dB	Signal-to-noise ratio (SNR)
Q^*_{FA}	0.01	Target false-alarm probability
β	0.34	Attack detection threshold

3.6.2 Impact of Sensor Clustering

While ADSP exploits shadowing correlation via sensor clustering, correlated sensor readings are, in general, known to degrade detection performance as they limit diversity gain [55, 56, 106]. Therefore, we first study the effect of sensor clustering on detection performance to understand the efficiency vs. robustness tradeoff in ADSP. Fig. 3.7 compares the achieved incumbent detection probabilities (Q_D) with and without sensor clustering (i.e., sensors are randomly selected by the BS). As expected, cooperative sensing without clustering yields higher detection probability than with sensor clustering with -20 dB SNR. However, the performance gap decreases as more sensors are involved in cooperative sensing, e.g., sensing with 5 clusters achieves $94\,\%$ of that without clustering. Note that this performance with clustering gets even closer to that of random selection as the SNR increases. Therefore, we can conclude that sensor clustering is not critical to incumbent detection, while it provides an efficient means of attack detection.

3.6.3 Attack Detection Performance

As a first line of defense, the attack detector in ADSP must be able to correctly identify any abnormal sensors within each cluster and discard their reports before making a final

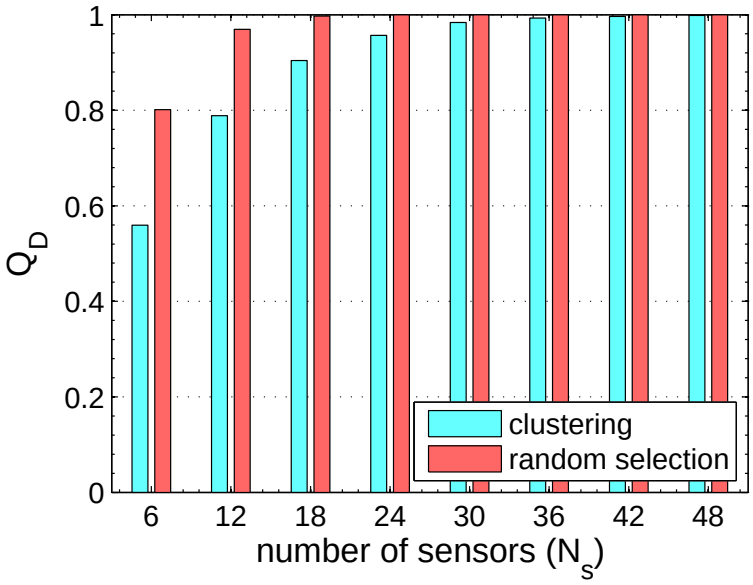
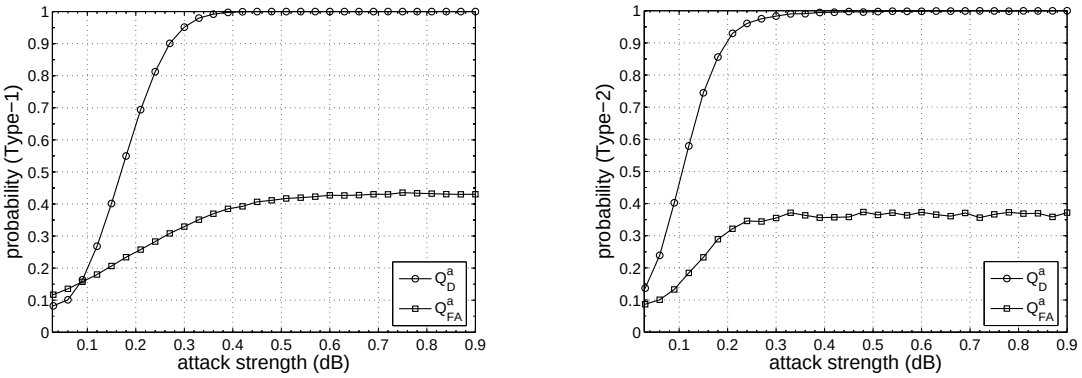


Figure 3.7: Impact of sensor clustering: Sensor clustering with $N_c = 5$ achieves $94\,\%$ of the detection performance without clustering.

decision. Fig. 3.8 shows the performance of our correlation-based filter under both types of attacks. The lower and upper thresholds (i.e., $TH_{\{L,U\}}$) for the correlation filter are set using Eq. (3.17) with a $99\,\%$ confidence interval, i.e., $\epsilon = 0.01$. The figures indicate that the attack detection rate, i.e., probability that a manipulated sensing report will be correctly filtered, increases with attack strength under both attack types. This is because the larger the deviation from the normal profile, the easier it is to identify them. However, the attack false-alarm rate also increases with attack strength because normal sensing reports will be mistakenly flagged more frequently by the manipulated sensing reports, and as a result, normal sensing reports will be classified as attacks more frequently. The figures show that ADSP performs well against both types of attacks.

3.6.4 Attack-Tolerance for One-Time Sensing

We now demonstrate the robustness of ADSP to both type-1 and type-2 attacks for one-time sensing. Fig. 3.9 plots the incumbent false-alarm (Q_{FA}) and detection (Q_D) probabilities under type-1 and type-2 attacks, respectively. Note that Q_{FA} and Q_D are *normalized* with respect to the maximum achievable values in the absence of attacks. The figure shows that the correlation filter is efficient in mitigating the effect of attacks on incumbent detec-



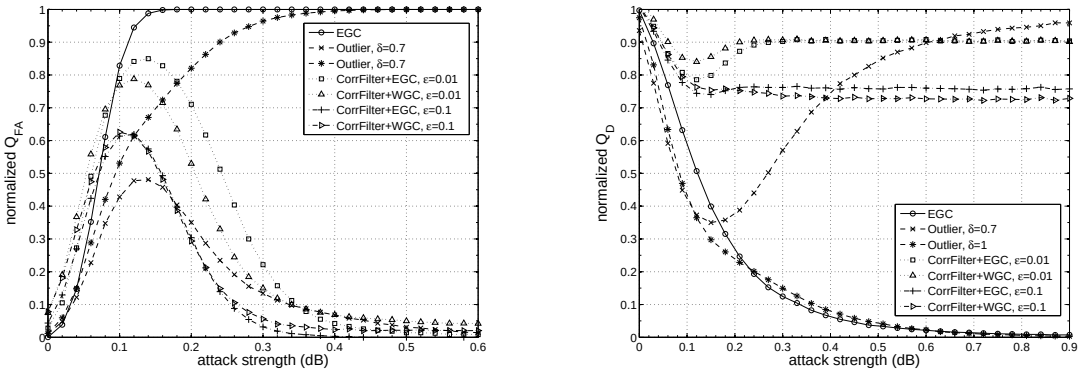
(a) Type-1 attack (b) Type-2 attack

Figure 3.8: Attack detection performance of the correlation filter: The detection and false-alarm probabilities of our correlation filter increase with attack strength under both types of attacks.

tion performance, e.g., $99.2\,\%$ for type-1 and $97.4\,\%$ for type-2 attacks, thanks to its ability to accurately filter out manipulated sensing reports. By contrast, without ADSP (denoted by EGC in Fig. 3.9), Q_{FA} and Q_D rapidly converge to 1 and 0, respectively, as attack strength increases, i.e., attacks have maximal influence on the data-fusion results.

We make the following four main observations. First, the performance of ADSP suffers in cases of low attack strengths (e.g., <0.4 dB for type-1 attack). This is because such low attack strengths do not exhibit deviations significant enough to be detected (thus causing *under-filtering*), yet they affect data-fusion decisions. The proposed weighted gain combining (WGC) mitigates this performance deficiency for both types of attacks by adaptively adjusting sensing reports' weights based on their statistical significance. However, WGC performs as well as, or even worse than, EGC when the attack strength is either (i) extremely low so that most attacks will not be filtered out or (ii) large enough so that most (or all) attacks are filtered out, as can be seen in Fig. 3.9 with $\epsilon = 0.01$. This is because, in the first case, the unfiltered attacks will decrease the weights of the legitimate sensing reports, while sharing large weights among themselves. On the other hand, in the second case, the legitimate sensing reports with extreme values are likely to be assigned small weights despite their critical role in accurate detection of incumbents.

Second, ADSP outperforms the statistics-based filtering method proposed in [79] (denoted by Outlier in Fig. 3.9). The fusion center filters out the sensing reports outside the

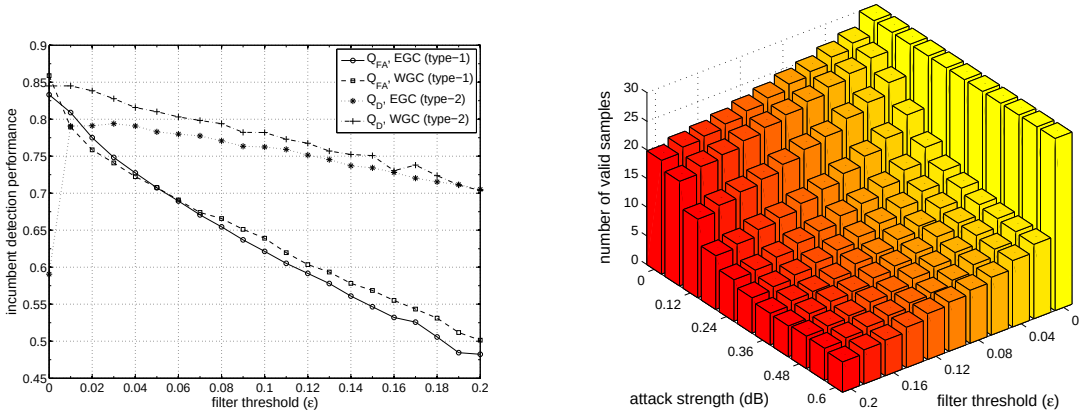


(a) Type-1 attack (b) Type-2 attack

Figure 3.9: Attack-tolerance of ADSP: ADSP (a) minimizes the false-alarm probability by up to $99.2\,\%$ for type-1 attacks, and (b) achieves $97.4\,\%$ of maximum achievable detection probability (i.e., with 20 normal sensing reports in 5 clusters) for type-2 attacks.

range $[\, e_1 - \delta \cdot e_{iqr}, e_3 + \delta \cdot e_{iqr} \,]$ where e_1 and e_3 represent the first and third quartile of the samples, respectively, and $e_{iqr} = e_3 - e_1$ is the interquartile range (see Eq. (4) in [79]). This method does not require sensor clustering, and thus, one might think that it performs well when attack strength is strong enough to be easily detected as an outlier. However, the performance depends strongly on the filtering range, i.e., the choice of δ, the result of which varies with attack scenarios. For example, when $\delta = 0.7$, performance suffers from *over-filtering* with a high attack mis-detection rate. On the other hand, when $\delta = 1$, performance suffers from *under-filtering*, and as a result, Q_{FA} and Q_D converge to 1 and 0, respectively, even in the case of high attack strength. In contrast, ADSP accurately detects manipulated sensing reports by considering shadowing correlation.

Third, even in the case of high attack strength, ADSP does not completely eliminate the effects of attacks for the following reasons. First, the fixed threshold parameter ϵ does not work optimally for all attack strengths, thus causing either over- or under-filtering, both of which degrade detection performance. The over-filtering caused by a large threshold value (e.g., $\epsilon = 0.1$) turned out to lower both Q_{FA} and Q_D, as shown in Fig. 3.9. Second, as a result of filtering, the fusion center will have fewer samples to be used for data fusion. Since data fusion is sensitive to the number of samples used, especially in very low SNR environments (as shown in Fig. 3.7), incumbent detection performance degrades. For example, with 20 sensing reports remaining after filtering out all 10 manipulated sensing reports, the



(a) Incumbent detection performance

(b) Number of valid sensing reports

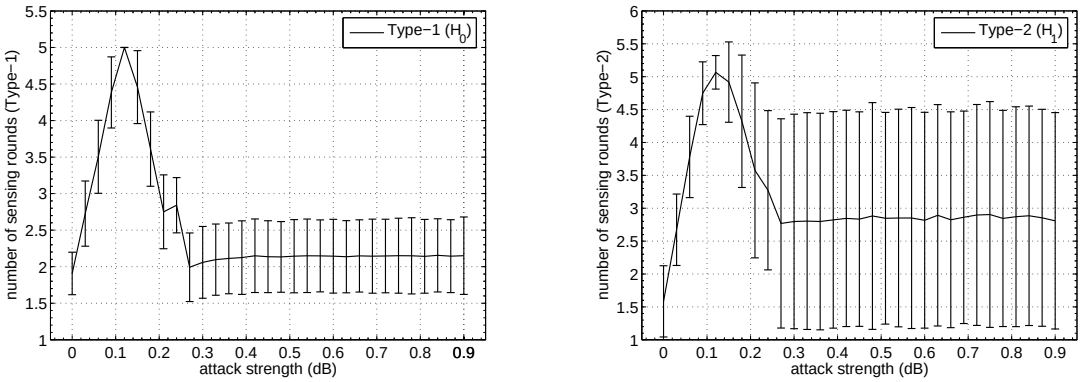
Figure 3.10: Impact of threshold parameter (ϵ): (a) Q_{FA} and Q_D exhibit different behaviors under various ϵ values, and (b) the number of valid sensing reports for data fusion depends on both filter threshold and attack strength.

average achievable Q_D is 0.88, which corresponds to the normalized Q_D of 0.93 in Fig. 3.9.

Fourth, in the absence of attacks, the correlation filter incurs a small increase in both Q_{FA} and Q_D. This is caused by the inaccuracy in the log-normal approximation of sensing reports, which causes over-filtering even in the case of no attacks. We observed that this performance anomaly can be mitigated by reducing the sensing duration T_S (e.g., <1 ms), which makes the approximation more accurate because the distribution of sensing reports more closely resembles a normal distribution.

3.6.5 Tradeoff in Setting the Detection Threshold

We now study the impact of the filtering threshold on attack detection performance. Fig. 3.10(a) plots the impact of the filtering threshold ϵ on incumbent detection performance. In this simulation, we fixed the attack strength at 0.1 dB for both types of attacks. The figure shows that Q_{FA} monotonically decreases as ϵ increases for both fusion rules, implying that filtering out more sensing reports always helps to lower the false-alarm rate of incumbents. For the same reason, however, a large ϵ degrades the detection probability Q_D. This can be explained by the heavy-tail of a log-normal distribution of shadow fading; filtering out high RSSs at the tail lowers the decision statistics significantly, thus reducing the chance of generating false-alarms (or detecting incumbents). Another observation is



(a) Type-1 attack (b) Type-2 attack

Figure 3.11: Average number of sensing rounds under various attack strengths: the number of sensing rounds needed to meet the detectability requirement, i.e., $Q_{FA}, Q_{MD} \leq 0.01$, both under the filter threshold $\epsilon = 0.1$ and -20 dB SNR.

that WGC outperforms EGC for type-2 attacks, thanks to its ability to adjust the weights for sensing reports based on their significance. However, the performance gain decreases as ϵ increases. For type-1 attacks, WGC also outperforms EGC in case of under-filtering, e.g., $\epsilon \in [0.01, 0.06]$, as discussed in Section 3.6.4.

Fig. 3.10(b) shows the average number of valid sensing reports (i.e., those that passed the filter). It clearly indicates that the filter becomes more aggressive in rejecting sensing reports as ϵ increases, thus reducing the number of sensing reports to be used for making a final fusion decision. Therefore, the filter must be carefully designed to make the tradeoff between false-alarm and detection probabilities, while considering their dependency on attack strength.

3.6.6 Meeting the IEEE 802.22 Detection Requirements via Sensing Scheduling

Here we evaluate the performance of the sensing scheduling algorithm in ADSP in terms of the number of sensing rounds (i.e., detection delay). Fig. 3.11 shows the number of sensing rounds needed to meet the detectability requirement of $Q_{FA}, Q_{MD} \leq 0.01$, which is below the requirements of IEEE 802.22, i.e., $Q_{FA}, Q_{MD} \leq 0.1$. Figs. 3.11(a) and 3.11(b) plot the mean and standard deviation of the number of sensing rounds. The figures indi-

cate that the average number of sensing rounds is maximized when the attack strength is relatively small, i.e., 0.12 dB, thus confirming the observation made in Fig. 3.9. In 802.22, sensing rounds can be scheduled as frequent as once every 10 ms, i.e., one MAC frame size in 802.22. Therefore, Fig. 3.11 implies that ADSP can meet the incumbent detection timing requirement of 802.22, i.e., the returning primary signal must be detected within 2 seconds, since the maximum required number of sensing rounds remains below 5.

3.7 Related Work

The problem of ensuring robustness in distributed sensing has been studied in [28, 79, 108]. Chen *et al.* [28] proposed a robust data-fusion scheme that dynamically adjusts the reputation of sensors based on the majority rule. Similarly, in the IEEE 802.22 standard draft, a voting rule [108] has been proposed for secure decision fusion. However, the voting rule may not work well in a very low SNR environment where a majority of sensors fail to detect the primary signal. Kaligineedi *et al.* [79] presented a pre-filtering scheme based on a simple outlier method that filters out extremely low or high sensor reports. However, their method may not be suitable for a very low SNR environment such as 802.22 WRANs, where a final data-fusion decision is very sensitive to small deviations in RSSs. The defense against Primary User Emulation Attack (PUEA) has also been studied in [11, 26]. Chen *et al.* [26] proposed an RSS-based location verification scheme to detect a fake primary transmitter. This scheme, however, requires the deployment of a dense sensor network for estimating the location of a signal source, and thus, incurs high system overhead. Anand *et al.* [11] analyzed the feasibility of PUEA and presented a lower-bound on the probability of a successful PUEA. However, they did not address the impact of PUEA on the performance of cooperative sensing.

The problem of enforcing/enticing secondary users to observe spectrum etiquette has also been studied. Woyach *et al.* [150] studied how to entice secondary users to observe spectrum etiquette by giving them incentives via a game-theoretic approach. In a similar context, Liu *et al.* [91] studied the problem of detecting unauthorized use of a licensed spectrum. They exploited the path-loss effect as a main criterion for detecting anomalous

spectrum usage and presented a machine-learning approach for more general cases. In contrast, we focus on intelligent filtering of suspicious sensor reports.

In a broader context, our work is related to work on secure data aggregation [43, 157, 163] and insider attack detection [90] in wireless sensor networks. However, the problem we consider differs in that it focuses on an important, realistic case where attackers manipulate sensor reports to mislead the fusion center in making a final decision on detection of a primary signal.

In summary, ADSP differs from previous work in several key aspects. First, we exploit shadow-fading correlation for anomaly detection, which has not been considered before. Second, ADSP is unique in that it enables normal spectrum sensing operation even in a hostile environment by *proactively* filtering out suspicious sensor reports, and scheduling sensing multiple times, while most previous work focuses on one-time sensing. Third, ADSP can detect attacks that purposely lower the RSS to obscure the existence of a primary signal (i.e., type-2 attacks), while most previous work focused on detecting spoofed primary signals (i.e., type-1 attacks).

3.8 Conclusion

The design of reliable distributed sensing for opportunistic spectrum use is a major research challenge in DSA networks. To address this challenge, we have developed a novel attack-tolerant distributed sensing protocol (ADSP) that selectively filters out abnormal sensor reports, and thus maintains the accuracy of incumbent detection. The key idea behind this mechanism is that the measured primary signal strength at nearby sensors should be correlated due to shadow fading, which has not been considered before. To realize this idea, we proposed a sensor clustering method and designed filters and data-fusion rules based on the correlation analysis of sensor reports. We also proposed a sensing scheduling scheme based on sequential hypothesis testing that finds an optimal stopping time for sensing, while meeting the detection requirements of 802.22. ADSP can be readily implemented in 802.22 WRANs, incurring very low processing and communication overhead. We evaluated ADSP in realistic shadowing environments of 802.22 WRANs, demonstrating

its ability to tolerate both type-1 and type-2 attacks.

Algorithm 3 ATTACK-TOLERANT DISTRIBUTED SENSING WITH WEIGHTED GAIN COMBINING

Procedure **ADSP_WGC**($\{R_i\}, Q_{FA}, \beta$)

while each sensing round n **do**
  $T_{\Sigma,n} \leftarrow 0$ /* Decision statistic */
  $N_{normal} \leftarrow 0$ /* Number of normal sensing reports */
  // Step 1. Check (ab)normality of sensing reports
  for each sensor cluster $\mathcal{S}_k$ $k = 1, \ldots, N_c$ **do**
    for each sensor $i \in \mathcal{S}_k$ **do**
      $(\texttt{Isnormal}(i), \mathbf{w}_i) \leftarrow$ **CorrFilter**$(i, \{R_j\}_{j \in N(i)}, \beta)$
    end for
  end for
  // Step 2. Update decision statistic
  for each sensor cluster $\mathcal{S}_k$ $k = 1, \ldots, N_c$ **do**
    for each sensor $i \in \mathcal{S}_k$ **do**
      if $\texttt{Isnormal}(i) == 1$ **then**
        Update w_i using Eq. (3.19)
        $T_{\Sigma,n} \leftarrow T_{\Sigma,n} + w_i R_i$
        $N_{normal} \leftarrow N_{normal} + 1$
      end if
    end for
  end for
  $T_{\Sigma,n} \leftarrow T_{\Sigma,n} \times N_{normal} / \sum w_i$ /* Normalization */
  Calculate the decision threshold η_n using Eq. (3.4)
  if $T_{\Sigma,n} > \eta_n$ **then**
    $\Lambda_n \leftarrow \Lambda_{n-1} + \ln \frac{\phi_1'}{\phi_0'}$
  else
    $\Lambda_n \leftarrow \Lambda_{n-1} + \ln \frac{1-\phi_1'}{1-\phi_0'}$
  end if
  // Step 3. Make a final decision
  if $\Lambda_n \geq B$ **then**
    return 1 /* Primary exists */
  else if $\Lambda_n < A$ **then**
    return 0 /* Primary does not exists */
  else
    Schedule another sensing round and wait for the observation
  end if
end while

Algorithm 4 FILTERING ALGORITHM BASED ON CORRELATION ANALYSIS

Procedure **CorrFilter(**$i, \{R_j\}_{j \in N(i)}, \beta$**)**

1: $\texttt{blacklist\_counter}(i) \leftarrow 0$ /* Initialize the counter */
2: $\mathbf{w}_i \leftarrow [0, \ldots, 0]^T$ /* Initialize the weight vector */
3: $\texttt{Isnormal} \leftarrow 1$ /* Initialize the indicator */
4: **for** each neighbor $j \in N(i)$ **do**
5: Update w_{ij} using Eq. (3.19)
6: **if** $Corr(R_i, R_j) \neq \rho(d_{ij})$ using Eq. (3.17) **then**
7: $++\texttt{blacklist\_counter}(i)$
8: **end if**
9: **end for**
10: **if** $\texttt{blacklist\_counter}(i) > \beta \cdot N(i)$ **then**
11: $\texttt{Isnormal} \leftarrow 0$ /* Mark it as abnormal */
12: **end if**
13: **return** $(\texttt{Isnormal}, \mathbf{w}_i)$

CHAPTER 4

Efficient Detection of Small-Scale Primary Users

4.1 Introduction

Unlike the detection of *large-scale* primary signals (e.g., TV signals), detection of small-scale primary devices, such as wireless microphones (WMs), is very difficult and still remains to be an open problem for the following reasons. First, while a TV signal has a large transmission range (up to 150 km), the WM signal has a small spatial footprint due to its weak transmission power (typically 10-50 mW) [118]. This indicates that the 802.22 needs a separate dense sensor network for WM detection [107], or more preferably, an efficient cooperative sensing mechanism tailored to WM detection, which is the main focus of this chapter. Second, the ON-OFF patterns of WMs have high spatial and temporal variations [14]. WMs can be turned on at any location and at any time without prior notification to secondary users. They are usually mobile and used at each location for a short period of time. Therefore, it is practically infeasible to maintain a database for WMs [63] or to profile all the possible locations and schedules of WM usage in real time. More importantly, this unpredictability makes it hard for the base station (BS) to select proper sensors for cooperative sensing. Third, despite its small footprint, a WM must be detected according to the strict sensitivity requirement imposed by the FCC. For example, the 802.22 standard draft specifies that sensors must be able to detect WM signals as weak as -114 dBm over a 200 KHz band within 2 seconds, with both false-alarm and mis-detection probabilities less than 0.1. However, a recent measurement study [115] indicates that sensors suffer from a

high false-alarm rate when detecting WM signals due to their weak signal strengths [33].

The detection of WMs is important for efficient spectrum utilization, especially in the space domain. For example, when a WM signal is detected by the sensors without knowing/estimating the transmitter's location, all the secondary users located within the cell of typical radius 33 km (up to 100 km) may need to vacate the channel. Considering the small transmission range of a WM signal, i.e., 100-150 m, this can cause significant under-utilization of spectrum in the space domain. Therefore, secondary users in 802.22 must be able to accurately detect the presence of a WM signal, and also estimate the WM transmitter's location.

Despite its practical importance, however, little work has been done on the detection of small-scale primary signals. To the best of our knowledge, the disabling beacon protocol, recently proposed by the 802.22 Task Group 1 (TG 1) [36, 86], is the only known solution. The disabling beacon protocol aims to enhance WM detection by transmitting a specially-designed signal before starting WM devices. It is suitable for carrying additional information, such as the signature/authentication and geo-location of WMs, which helps improve spectrum efficiency via better spatial [34, 58, 107] and frequency reuse [24]. However, the disabling beacon protocol still has the following limitations. First, we do not expect that all WM users will be equipped with a separate beacon device in the near future in view of the fact that most users have not even registered their WMs. Second, the transmit power of the beacon message is limited to the same level as the WM's (i.e., 250 mW in a UHF band), and thus, beacons cannot compensate for the low sensor density in 802.22 [36]. Finally, the disabling beacon protocol incurs a significant sensing-time overhead (i.e., 5-100 ms) [36] compared to simple energy detection, which may take only 1 ms.

Motivated by these practical needs and problems, we propose an efficient sensing framework for detection of small-scale primaries using cooperative sensing. To cooperatively detect small-scale primary signals, the BS must carefully select a set of sensors by estimating the primary transmitter's characteristics, such as its location and transmit-power. We first assume this information is available to secondary users, and derive the optimal fusion-range within which the sensors cooperate to minimize detection delay, i.e., the number of sensing rounds needed for detecting a primary signal. Based on our analytical find-

ings, we then design—without assuming the availability of information on the primary transmitter's characteristics—a practical framework, called DeLOC, which performs joint cooperative sensing and location/transmit-power estimation, in order to meet detectability requirements, while minimizing detection delay.

4.1.1 Contributions

This chapter makes the following main contributions.

- Introduction of a novel *spatio-temporal* data-fusion scheme with the following salient features: it (i) exploits physical-layer signal propagation characteristics in the *space* domain by finding an optimal fusion range for cooperative sensing, and (ii) makes statistics-based decisions in the *time* domain by identifying an optimal time to stop scheduling sensing. This spatio-temporal fusion provides useful and practical insights and can be used as a general framework for designing sensing schemes.

- Identification and characterization of the impact of data-fusion range and sensor density on the performance of small-scale primary detection in CRNs. We derive a closed-form expression for the *optimal* fusion range that minimizes the average detection delay. Moreover, we show that the optimal fusion range does not depend on sensor density and that the minimum required sensor density for given detectability constraints decreases inversely proportional to the average detection delay.

- Development of a framework for *joint* small-scale primary detection and location/transmit-power estimation, called DeLOC. DeLOC iteratively performs cooperative sensing and location/transmit-power estimation until the fusion center (i.e., the BS) collects a sufficient amount of information to make a final decision. This approach allows sensing and estimation to refine each other over multiple scheduled periods.

- Design of a new data-fusion rule tailored to small-scale primary detection. Specifically, we propose a *sequential probability ratio test with ascending weight* (SPRT-AW) for DeLOC that intentionally delays decision-making at the BS by assigning small weight to decision statistics in early detection stages when location and transmit-

power estimates are inaccurate. Our simulation results show that DeLOC combined with SPRT-AW achieves high detection accuracy, while minimizing detection delay in a realistic 802.22 WRAN environment.

4.1.2 Organization

The remainder of this chapter is organized as follows. Section 4.2 describes the network, signal-propagation and spectrum sensing models, and briefly introduces our approach to WM signal detection. Section 4.3 studies the impact of the data-fusion range on the performance of WM detection and location/transmit-power estimation. Section 4.4 formulates the sequential hypothesis testing problem for WM detection and derives an optimal fusion range that minimizes average detection delay. Section 4.5 details our proposed iterative sensing framework, DeLOC, which incorporates location and power estimation, and presents the SPRT-AW based data-fusion rule. Section 4.6 evaluates the performance of DeLOC, and Section 4.7 reviews related work. Section 4.8 concludes the chapter.

4.2 Preliminaries

In this section, we introduce the network model, the wireless signal-propagation model, the WM sensing model, and the data-fusion model.

4.2.1 Network Model

We consider a CRN consisting of primary and secondary users in the same geographical area. In general, there are two types of PUs: large-scale (e.g., TV transmitters) and small-scale (e.g., WMs). Here we focus on detecting small-scale PUs. While the techniques that we propose can be applied to other small-scale primary transmitters, we will focus on WM detection in IEEE 802.22 WRANs. WMs use a weak transmit power of around 10-50 mW, or below [34, 118], and its transmission range is only 150-200 m, which is much smaller than the typical 802.22 cell radius of 33 km. We assume that WMs can use any UHF band and are turned on at random locations and at any time for relatively short periods of

time. In 802.22, the secondary spectrum users are called *consumer premise equipments* (CPEs), which represent households in rural areas. Such CPEs are stationary and their locations are known to the BS. CPEs transmit/receive data to/from the BS, and function as spectrum sensors during the quiet periods reserved for primary detection. All the CPEs within the cell must be silent during quiet periods, and employ the spectrum sensors to measure the received signal strengths (RSSs) and report them to the BS for data-fusion. We assume secondary users have been deployed in an area A, i.e., an IEEE 802.22 WRAN cell, following a point Poisson process with density ρ, i.e., $n_A \sim Poi(n; \rho|A|)$. We also assume that sensor density ρ, as the typical density of CPEs (i.e., households) in rural areas is very low (around $1.25/\text{km}^2$) [139].

4.2.2 Signal-Propagation and WM Sensing Models

We assume that sensor n's received primary signal strength can be characterized by the following propagation model:

$$P_n = P_o\Big(\frac{d_o}{d_n}\Big)^{\alpha} e^{X_n} e^{Y_n} \quad \text{(Watt)}, \tag{4.1}$$

where d_o is the reference distance (e.g., 1 m), P_o the received primary signal strength at the reference distance, α the path-loss exponent, and d_n the distance from the primary transmitter to sensor n. Shadow fading and multi-path fading are accounted for in e^{X_n} and e^{Y_n}, respectively, where $X_n \sim \mathcal{N}(0, \sigma^2)\ \forall n$. The log-normal shadow fading is often characterized by its dB-spread, σ_{dB}, which has the relationship $\sigma = 0.1\ln(10)\sigma_{dB}$.

We make the following assumptions regarding the WM signal detection: Sensors

A1) use energy detection for sensing, and

A2) sense an entire 6 MHz-wide TV channel.

Regarding A1, feature detection cannot be applied for WM detection because, unlike TV signals, there is no standard modulation specified by the FCC Report and Order (R&O) for WM signals [49]. The test statistic at sensor n can be approximated as Gaussian using

the Central Limit Theorem (CLT) as in [127]:

$$T_n \sim \begin{cases} \mathcal{N}\big(N_o, \frac{N_o^2}{M_s}\big) & \mathcal{H}_0 \text{ (no primary signal)} \\ \mathcal{N}\big(P_n + N_o, \frac{(P_n+N_o)^2}{M_s}\big) & \mathcal{H}_1 \text{ (primary signal exists)}, \end{cases} \tag{4.2}$$

where P_n is the power of a received primary signal at sensor n, N_o the noise power, i.e., -95.2 dBm for a TV channel with 6 MHz bandwidth [137], and M_s the number of signal samples, e.g., 6×10^3/ms for 6 MHz TV band at the Nyquist rate.

Regarding A2, WMs use a relatively narrow frequency band, i.e., 200 KHz, compared to a 6 MHz TV band. Therefore, sensing the entire TV channel simplifies the sensing design at the cost of decreased measured signal-to-noise ratio (SNR) due to the increased noise level over a 6 MHz-wide channel.

In each sensing round (i.e., quiet periods), the BS directs a set of sensors to perform sensing for a sensing duration of T_S (e.g., 1 ms), and the sensors report their readings to the BS for data fusion at the end of each sensing round.

4.2.3 Data-Fusion Model

For the data fusion rule at the BS, we assume Equal Gain Combining (EGC) for a single-round sensing. EGC is known to have near-optimal performance without requiring the estimation of channel gains [141], and has the following decision statistic $T_\Sigma \triangleq \sum_{n=1}^{n_s} T_n$, where T_n is the test statistic (i.e., measured RSS) of the energy detector at sensor n, and n_s is the number of cooperative sensors. EGC will be used to characterize the impact of the fusion range on detection performance (in Section 4.3). In `DeLOC`, the BS performs a sequential hypothesis testing for primary detection, and the test statistic of EGC, i.e., sum of the RSSs measured at cooperating sensors, is used in updating the decision statistic for hypothesis testing (in Section 4.4).

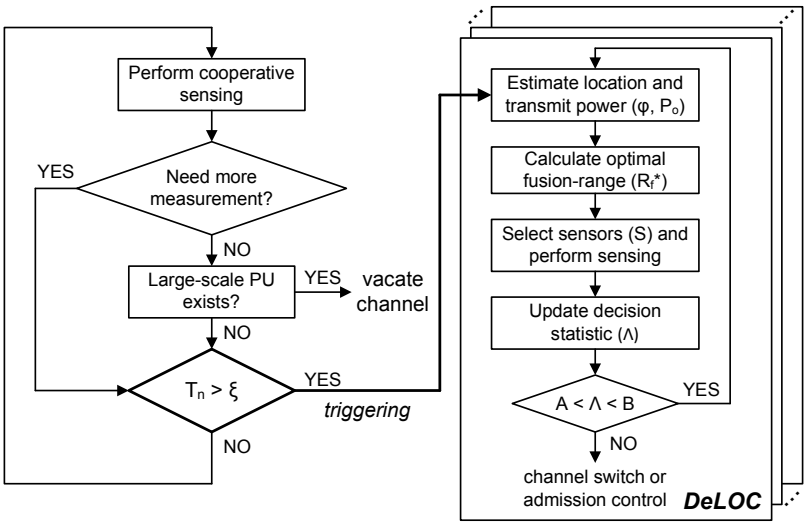


Figure 4.1: The DeLOC framework: When a sensor reports a test statistic above a predefined threshold (ξ), DeLOC for a small-scale primary signal, i.e., WM, is triggered by the BS that initiates and repeats the iteration between cooperative sensing and location/transmit-power estimation until the BS collects enough information to make a final decision.

4.2.4 The Proposed Approach

Fig. 4.1 illustrates our proposed spectrum-sensing framework, DeLOC, which is tailored to the detection of small-scale primary signals such as WMs. When a large-scale primary signal exists, all the sensors within the network (e.g., an 802.22 cell) must vacate the channel regardless of the presence of small-scale primary signals. Thus, when a large-scale primary signal exists, DeLOC will not be triggered.[1] To minimize energy consumption and communication overhead, DeLOC for WM sensing is triggered only when a sensor reports a test statistic above a predefined threshold (ξ) during the normal sensing mode for detection of large-scale primaries (i.e., TV signals). In our simulation study, we set $\xi = N_o + 3.5\,\sigma_o$ where N_o and $\sigma_o = \frac{N_o}{\sqrt{M_s}}$ are the mean and standard deviation of the test statistics under $\mathcal{H}_0$. Note that the BS can run multiple instances of DeLOC in parallel, corresponding to different triggering events at different geographical locations.

Upon triggering the detection process, the BS iteratively performs the location/transmit-power estimation and cooperative sensing until it collects a sufficient amount of informa-

[1] Note that large-scale primary signals can be reliably detected using either existing sensing schemes (e.g., [100]) or a geo-location database [63]. The detection of large-scale primary signals is not within the scope of this chapter.

tion, i.e., measured test statistics, to make a final decision on the presence of a WM. In each sensing stage, the BS first estimates the location and transmit power of a WM, and based on this estimation, it computes an optimal fusion range (R_f^*) for cooperative sensing. Then, based on the test statistics reported from the sensors, the BS updates the decision statistic, and compares it with predefined lower (A) and upper (B) thresholds, to make a final decision. The thresholds are designed to guarantee the desired false-alarm and misdetection probabilities (see Eq. (4.7) in Section 4.4). If the test statistic is below the lower threshold, then the BS assumes the absence of a primary transmitter, e.g., the event was falsely triggered by measurement error. If the test statistic exceeds the upper threshold, the BS assumes the presence of a primary transmitter at the estimated location, and then takes an appropriate action, e.g., vacating the channel or disabling nearby secondary users. Otherwise, the BS schedules another sensing event with the sensors within the optimal sensing range, thus accumulating detection confidence in the temporal domain.

4.3 Cooperative Sensing for Small-scale Primary Detection

In this section, we first study the impact of sensor cooperation on the detection of small-scale PUs. In particular, we investigate the impact of the data-fusion range and localization error on the performance of signal detection.

4.3.1 To Cooperate or Not?

Although cooperative sensing is shown to help improve sensing performance of large-scale PUs [100, 101], its relevance for small-scale primary detection is less obvious. On one hand, a large number of sensors may be needed for cooperation, because WM signals usually have small footprints, and their spatial-temporal ON-OFF patterns are highly unpredictable. On the other hand, those sensors located far from the WM will report only noise power. Thus, employing a large number of noisy reports may adversely affect detection performance, since the energy detector cannot extract the primary signal from the

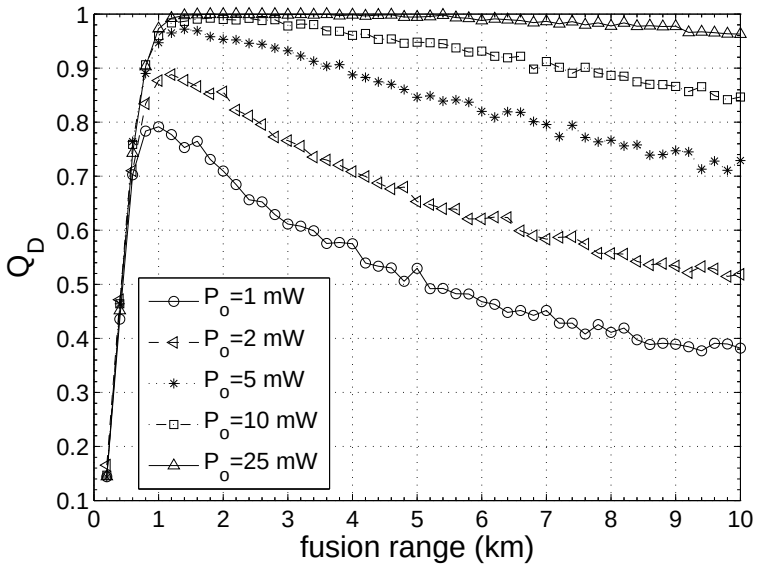


Figure 4.2: Impact of data-fusion range: The figures show the existence of an optimal fusion range R_f in terms of maximizing detection probability Q_D. Sensor density, sensing duration, shadow fading dB-spread, path-loss exponent, and the required false-alarm probability are set to $\rho = 1.25 \times 10^{-6}/\text{m}^2$, $T_S = 1$ ms, $\sigma_{dB} = 5.5$ dB, $\alpha = 4$, and $Q_{FA} = 0.01$, respectively.

noise. The set of sensors chosen for cooperative sensing may also affect the accuracy of location and transmit-power estimations, which play an important role in detecting WMs. In what follows, we thus investigate the impact of fusion range and location uncertainty on detection performance.

4.3.2 Impact of the Data-Fusion Range

Fig. 4.2 shows the impact of data-fusion range on the detection probability Q_D subject to a given false-alarm probability $Q_{FA} = 0.01$ using MATLAB-based simulation. Intuitively, when the range is small, enlarging the range increases sensor diversity, thus improving sensing performance. However, as the range increases further, the test statistics measured from the sensors more closely resemble the noise level, adversely affecting detection performance. This implies the existence of an optimal fusion range that maximizes sensing performance. Fig. 4.2 also indicates that the optimal range depends on the transmit power of the primary transmitter. (Also, see Fig. 4.5(a) in Section 4.6 for more detail.)

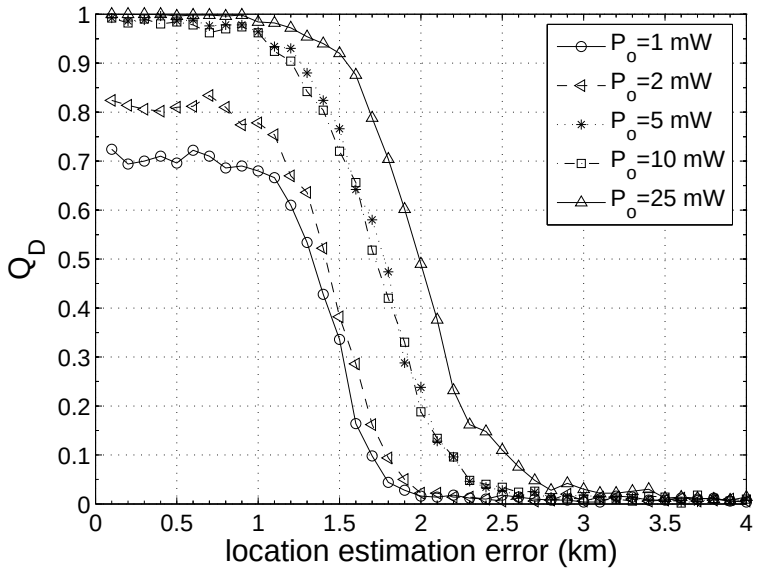


Figure 4.3: Impact of localization error: Detection performance degrades as location estimation becomes less accurate. The simulation parameters are set to $\rho = 1.25 \times 10^{-6}/\text{m}^2$, $T_S = 1$ ms, $\sigma_{dB} = 5.5$ dB, $\alpha = 4$, $Q_{FA} = 0.01$, and $R_f = 2$ km, respectively.

4.3.3 Impact of Location-Estimation Error

Ideally, the BS performs data fusion with a set of sensors located within the data-fusion range centered around the primary transmitter. In practice, however, the unpredictability of a primary's location can significantly degrade the quality of incumbent detection because it makes it difficult to select a proper set of sensors for data fusion.

Fig. 4.3 plots WM detection probability (Q_D) with one-time sensing for various location estimation errors. The figure shows that detection performance degrades drastically as the localization error increases beyond a certain level, e.g., 1 km. Even a small difference in one-time detection performance can greatly affect the average number of sensing rounds to achieve the desired false-alarm and mis-detection probabilities. Moreover, accurate location estimation is necessary for efficient coexistence between a WM and secondary users once DeLOC detects the presence of the WM signal. Therefore, reasonably accurate localization is necessary in our design of small-scale primary detection.

4.4 Detection of Small-scale Primary via Spatio-Temporal Data-fusion

In this section, we first formulate the small-scale primary detection problem as a sequential hypothesis testing problem. We then derive the optimal data-fusion range that minimizes average detection delay. We finally show that the sensor density required to meet a certain detectability constraint decreases inversely proportional to the average number of sensing rounds scheduled for detection.

4.4.1 Hypothesis Testing

Let $\boldsymbol{\theta}_t = [T_1, \ldots, T_{|S_t|}]^T$ denote the vector of test statistics (i.e., RSSs) measured at the sensing stage t by a set S_t of cooperating sensors. A sensor is selected by the BS if it is within the fusion range R_f from the WM transmitter. Note that the fusion range, and hence the set of cooperating sensors, can differ in each sensing stage according to the WM's estimated location and transmit-power level. Let $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_N^T]^T$ denote the $M \times 1$ vector of test statistics measured at sensors over N sensing stages, where $M = \sum_{t=1}^{N} |S_t|$. As shown in Eq. (4.2), the test statistics can be estimated to be Gaussian regardless of the existence of a primary signal [127].

Our detection problem is thus a binary Gaussian classification problem where the observed test statistic $\boldsymbol{\theta}$ belongs to one of two classes, $\mathcal{H}_0$ or $\mathcal{H}_1$, where:

$$\mathcal{H}_0 : \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) \quad \text{(no primary signal)}$$
$$\mathcal{H}_1 : \boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \quad \text{(primary signal exists)},$$

where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are the mean vector and covariance matrix of the test statistics under $\mathcal{H}_k, k \in \{0, 1\}$. The average test statistics under each hypothesis are $\boldsymbol{\mu}_0 = N_o \times \mathbf{1}$ and $\boldsymbol{\mu}_1 = (P_R + N_o) \times \mathbf{1}$, where N_o and P_R are the average noise power and received primary signal power at sensors, respectively.[2]

[2] Since the BS does not have the exact distribution of the received primary signal strengths, the BS can set P_R to -107 dBm, which is the detectability requirement in 802.22 [36].

The covariance matrix $\mathbf{\Sigma}_0$ can be expressed as $\mathbf{\Sigma}_0 = \sigma_o^2\,\mathbf{I}$, where $\mathbf{I}$ is an $M \times M$ identity matrix and $\sigma_o^2 = \frac{N_o^2}{M_s}$. Note that the correlation among sensor reports is negligible under the assumption that the locations of the sensors and WM transmitter are fixed during the detection process [100]. Moreover, in a very low SNR environment, it is reasonable to assume $P_n + N_o \approx N_o \ \forall n$, and hence the covariance matrix $\mathbf{\Sigma}_1$ can be approximated as $\mathbf{\Sigma}_o \approx \mathbf{\Sigma}_1 = \sigma_o^2\,\mathbf{I}$.

4.4.2 Sensing Scheduling via Sequential Probability Ratio Test

In DeLOC, the BS schedules the sensing periods (stages) until it obtains a sufficient amount of information for making a final decision. Via sensing scheduling, the BS receives a sequence of test statistics $\{\boldsymbol{\theta}_t\}_{t=1}^{N}$ from the sensors. We adopt Wald's *Sequential Probability Ratio Test* (SPRT) [147] to process the statistics and determine when to stop sensing. SPRT is optimal in the sense of minimizing the average number of observations, given bounded false-alarm probability Q_{FA} and mis-detection probability Q_{MD}. It enables the BS to reduce erroneous triggering of WM detection by optimizing its decision thresholds.

The decision statistic Λ_N is the log-likelihood ratio derived from a sequence of test statistics $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N$ as follows:

$$\Lambda_N \triangleq \lambda(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N) = \ln \frac{f_1(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)}{f_0(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)}, \tag{4.3}$$

where $f_k(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_N)$ is the joint p.d.f. of the sequence of test statistics under the hypothesis $\mathcal{H}_k \ k \in \{0, 1\}$.

With SPRT, a decision is made based on the observed sequence of test statistics, $\{\boldsymbol{\theta}_t\}_{t=1}^{N}$, using the following rules:

$$\begin{aligned} \Lambda_N \geq B &\ \Rightarrow\ \text{accept } \mathcal{H}_1 \text{ (primary signal exists)} \\ \Lambda_N < A &\ \Rightarrow\ \text{accept } \mathcal{H}_0 \text{ (no primary signal)} \\ A \leq \Lambda_N < B &\ \Rightarrow\ \text{take another observation,} \end{aligned}$$

where A and B $(0 < A < B < \infty)$ are the detection thresholds that depend on the desired

values of Q_{FA} and Q_{MD}.

Let λ_t be the log-likelihood ratio at sensing stage t, i.e., $\lambda_t = \ln \frac{f_1(\boldsymbol{\theta}_t)}{f_0(\boldsymbol{\theta}_t)}$. Recall that $\{\boldsymbol{\theta}_t\}_{t=1}^N$ are Gaussian, and assume they are independent and identically distributed (i.i.d.). Then, Eq. (4.3) becomes:

$$\Lambda_N = \sum_{t=1}^{N} \lambda_t = \sum_{t=1}^{N} \ln \frac{f_1(\boldsymbol{\theta}_t)}{f_0(\boldsymbol{\theta}_t)} = \sum_{t=1}^{N} \sum_{n=1}^{|S_t|} \ln \frac{f_1(\widetilde{T}_n)}{f_0(\widetilde{T}_n)}, \tag{4.4}$$

where the test statistic can be approximated as Gaussian using the Central Limit Theorem (CLT) as $\widetilde{T}_n \sim \mathcal{N}(\mu_k, \sigma_o^2)$ under $\mathcal{H}_k$, as shown in Eq. (4.2).

We now consider the *normalized* test statistics (i.e., RSSs) to simplify the derivation of the average number of sensing rounds. Let $\widetilde{T}_n \triangleq T_n \cdot \sigma_o^{-1}$ denote the normalized test statistic, i.e., $t_{n|\mathcal{H}_k} \sim \mathcal{N}(\phi_k, 1)$ where $\phi_k = \frac{\mu_k}{\sigma_o},\ \forall k$. Then, we have:

$$\lambda_t = \sum_{n=1}^{|S_t|} \ln \frac{h_1(\widetilde{T}_n)}{h_0(\widetilde{T}_n)} = (\phi_1 - \phi_0) \sum_{n=1}^{|S_t|} \widetilde{T}_n + \frac{1}{2} \sum_{n=1}^{|S_t|} (\phi_0^2 - \phi_1^2), \tag{4.5}$$

where $h_k(\cdot)$ is the p.d.f. of $\widetilde{T}_{n|\mathcal{H}_k}$.

Based on Eqs. (4.4) and (4.5), the decision statistic Λ_N can be expressed as:

$$\begin{aligned} \Lambda_N &= (\phi_1 - \phi_0) \sum_{t=1}^{N} \sum_{n=1}^{|S_t|} \widetilde{T}_n + \frac{1}{2} \sum_{t=1}^{N} \sum_{n=1}^{|S_t|} (\phi_0^2 - \phi_1^2) \\ &= (\phi_1 - \phi_0) \sum_{n=1}^{M} \widetilde{T}_n + \frac{M}{2} (\phi_0^2 - \phi_1^2), \end{aligned} \tag{4.6}$$

where $M = \sum_{t=1}^{N} |S_t|$ is the total number of test statistics collected by the BS through N sensing stages.

SPRT can meet the desired detectability requirements by carefully setting the detection thresholds A and B. Let a^* and b^* denote the desired values of Q_{FA} and Q_{MD}, respectively. Then, the decision boundaries are given by [147]:

$$A = \ln \frac{b^*}{1 - a^*} \quad \text{and} \quad B = \ln \frac{1 - b^*}{a^*}, \tag{4.7}$$

and the actual achievable error probabilities can only be slightly larger than the desired values a^* and b^*.

4.4.3 Minimization of the Average Detection Delay

Recall that our goal is to minimize the number of sensing rounds that the BS has to schedule to meet the desired detection performance requirements, e.g., $Q_{FA}, Q_{MD} \leq 0.01$. Thus, we first derive a closed-form expression for the average number of sensing rounds required until a decision is made (i.e., either boundary A or B is reached).

The average number of sensing rounds required for making a decision (denoted by $\mathbb{E}[N]$) can be computed as [147]:

$$\mathbb{E}[N] = \mathbb{E}[\lambda \,|\, \mathcal{H}_k]^{-1} \times \mathbb{E}[\Lambda_N]. \tag{4.8}$$

First, using Eq. (4.5), the average value of the log-likelihood ratio λ under hypothesis $\mathcal{H}_k$ can be derived as:

$$\mathbb{E}[\lambda \,|\, \mathcal{H}_k] = (\phi_1 - \phi_0)\, \mathbb{E}\Big[\sum_{n=1}^{|S_t|} \widetilde{T}_{n|\mathcal{H}_k}\Big] + \frac{1}{2}\mathbb{E}\Big[\sum_{n=1}^{|S_t|} (\phi_0^2 - \phi_1^2)\Big]. \tag{4.9}$$

Next, the expectation of Λ_N in Eq. (4.8) can be found as follows. Suppose $\mathcal{H}_1$ holds, then Λ_N will reach the decision boundary A with the desired mis-detection probability b^*; otherwise, it will reach B. Thus, using Eq. (4.7), we have:

$$\mathbb{E}[\Lambda_N \,|\, \mathcal{H}_1] = b^* \, \ln \frac{b^*}{1-a^*} + (1-b^*) \, \ln \frac{1-b^*}{a^*}. \tag{4.10}$$

Based on Eqs. (4.8), (4.9) and (4.10), we can derive the average number of sensing rounds needed for decision-making as:

$$\mathbb{E}[N \,|\, \mathcal{H}_1] = \frac{b^* \, \ln \frac{b^*}{1-a^*} + (1-b^*) \, \ln \frac{1-b^*}{a^*}}{(\phi_1 - \phi_0)\mathbb{E}\big[\sum_{n=1}^{|S_t|} \widetilde{T}_{n|\mathcal{H}_1}\big] + \frac{1}{2}(\phi_0^2 - \phi_1^2)\mathbb{E}\big[|S_t|\big]}. \tag{4.11}$$

Similarly, the average number of sensing rounds under $\mathcal{H}_0$, i.e., $\mathbb{E}[N \,|\, \mathcal{H}_0]$, can be derived.

Eqs. (4.9), (4.10), and (4.11) indicate that the average number of sensing rounds $\mathbb{E}[N]$ depends on: (i) the average number of sensors within the fusion range, which can be easily calculated as $\mathbb{E}[|S_t|] = \rho\pi R_f^2$, under the assumption of the point Poisson distribution of sensors, i.e., $|S_t| \sim Poi(n; \rho\pi R_f^2)$, and (ii) the sum of their reported test statistics, i.e., $\mathbb{E}[\sum_{n=1}^{|S_t|} \widetilde{T}_{n|\mathcal{H}_k}]$.

As will be shown below, the sum of test statistics is affected mainly by three parameters: (i) sensor density (ρ), (ii) transmit-power level of the primary device (P_o), and (iii) data-fusion range (R_f), assuming other parameters remain constant. In general, sensor density is known at the BS at the time of system deployment, and the transmit power can be estimated based on measurements (which will be detailed in Section 4.5.1). Therefore, we opt to derive an optimal fusion range R_f^* that minimizes the average number of sensing rounds, thus minimizing detection delay.

4.4.4 Approximation of the Sum of Test Statistics

Unfortunately, it is infeasible to derive a closed-form expression for the exact distribution of the sum of test statistics. This is because it depends on various random factors including the number of sensors within the fusion range, their locations relative to the primary transmitter, channel gains between the primary transmitter and the sensors, and the measurement error of the energy detector. Therefore, as a first step to derive an optimal fusion range, we approximate the sum of test statistics in Eq. (4.11) as a shifted log-normal random variable.

Let $T_{\Sigma(\rho,R_f)}$ denote the sum of the test statistics measured at the sensors within the fusion radius R_f from the WM transmitter, in a network with sensor density ρ. Then, under $\mathcal{H}_1$, it can be approximated as:

$$\begin{aligned}
\mathbb{E}\big[T_{\Sigma(\rho,R_f)}\big] &= \mathbb{E}\Big[\sum_{n\in S_t} T_{n|\mathcal{H}_1}\Big] \\
&= \mathbb{E}\Big[\sum_{n\in S_t} \mathcal{N}(P_n + N_o, \sigma_o^2)\Big] \\
&\approx \mathbb{E}\Big[\sum_{n\in S_t} P_n\Big] + \mathbb{E}\Big[\sum_{n\in S_t} N_o\Big], \qquad (4.12)
\end{aligned}$$

where P_n is the received primary signal strength at sensor n and $S_t \equiv S(\rho, R_f)$ for brevity. The approximation in Eq. (4.12) is made based on the observation that, assuming the sensing duration of 1 ms, the measurement error of the energy detector is relatively smaller than the average received primary signal strength, i.e., $\sigma_o^2 \ll (P_n + N_o)$.

Based on Eq. (4.12), we now focus on approximation of the sum of received primary signal strengths, which can be rewritten as $\mathbb{E}[\sum_{n\in S_t} P_n] = P_o\, \mathbb{E}[\sum_{n\in S_t} g(d_n) e^{X_n} e^{Y_n}]$ where P_o is the primary's transmit power, $g(d_n)$ is the sensor n's channel gain due to path-loss, i.e., $g(d_n) = (d_o/d_n)^\alpha$, and e^{X_n} and e^{Y_n} are the channel gains from shadowing and multi-path fading, respectively. We approximate the sum of channel gains due to path-loss, denoted by $\mathcal{G}_\Sigma(\rho, R_f) = \sum_{n\in S_t} g(d_n)$, as a log-normal random variable. Previous numerical studies have shown that the aggregate interference of Poisson-distributed transmitters to a single receiver can be accurately approximated as a log-normal distribution [99]. Conversely, assuming the reciprocity of the RF path, we can also approximate the sum of received primary signal strengths at sensors as a log-normal random variable. It has been shown that the impact of fading on received signal strengths is not a critical factor in such an approximation [69]. The effects of log-normal shadowing and multi-path fading in an average sense will be incorporated later (see Eq. (4.17)).

Denote $\mathcal{G}_\Sigma(\rho, R_f) \sim Log\text{-}\mathcal{N}(\mu_G, \sigma_G^2)$. Then, the p.d.f. of $\mathcal{G}_\Sigma(\rho, R_f)$ is given as:

$$p_{\mathcal{G}(\rho,R_f)}(x) = \frac{1}{x\sigma_G\sqrt{2\pi}} \exp\left(- \frac{(\ln x - \mu_G)^2}{2\sigma_G^2} \right), \tag{4.13}$$

where the μ_G and σ_G^2 have the following relationships [99]:

$$m_1(\rho, R_f) = e^{\mu_G + \frac{1}{2}\sigma_G^2} \;\text{ and }\; m_2(\rho, R_f) = e^{2\mu_G + \sigma_G^2}(e^{\sigma_G^2} - 1). \tag{4.14}$$

Here $m_k(\rho, R_f)$ is the kth cumulant of $\mathcal{G}_{(\rho,R_f)}$, given as:

$$\begin{aligned} m_k(\rho, R_f) &= \rho\pi(R_f^2 - \epsilon^2) \int_\epsilon^{R_f} \frac{2r}{(R_f^2 - \epsilon^2)} g(r)^k \mathrm{d}r \\ &= \frac{2\rho\pi\, d_o^{k\alpha}}{(k\alpha - 2)} \left(\frac{1}{\epsilon^{k\alpha-2}} - \frac{1}{R_f^{k\alpha-2}} \right), \end{aligned} \tag{4.15}$$

where d_o is the reference distance and ϵ is the minimum separation between the primary transmitter and the sensors, which is set to $\epsilon = 75$ m in our simulation.[3]

From Eqs. (4.14) and (4.15), the log-normal random variable $\mathcal{G}_\Sigma(\rho, R_f) \sim Log\text{-}\mathcal{N}(\mu_G, \sigma_G^2)$ can be approximated as:

$$\mu_G = \frac{1}{2}\ln\left(\frac{m_1^4}{m_1^2 + m_2}\right) \quad \text{and} \quad \sigma_G^2 = \ln\left(1 + \frac{m_2}{m_1^2}\right). \tag{4.16}$$

Therefore, from Eqs. (4.12) and (4.16), and by incorporating the effects of shadowing and multi-path fading assuming the fading is i.i.d. for each sensor, the sum of received primary power at the cooperating sensors S_t can be expressed as:

$$\mathbb{E}\Big[\sum_{n\in S_t} P_n\Big] = P_o \cdot \mathbb{E}[e^X] \cdot \mathbb{E}[e^Y] \cdot \mathbb{E}[\mathcal{G}_\Sigma(\rho, R_f)], \tag{4.17}$$

where $\mathbb{E}[e^X] = e^{\frac{1}{2}\sigma^2}$, $\sigma = 0.1\ln(10)\sigma_{dB}$, and $\mathbb{E}[\mathcal{G}_\Sigma(\rho, R_f)] = e^{\mu_G + \frac{1}{2}\sigma_G^2}$. For multi-path fading, we assume Rayleigh fading with zero mean, and thus, $\mathbb{E}[e^Y] = 1$.

Then, from Eqs. (4.12) and (4.17), the average of the sum of normalized test statistics can be expressed as:

$$\begin{aligned}\mathbb{E}\Big[\sum_{n=1}^{|S_t|} \widetilde{T}_{n|\mathcal{H}_1}\Big] &= \mathbb{E}\Big[T_{\Sigma(\rho,R_f)}\,\sigma_o^{-1}\Big] \\ &= \Big(P_o\, e^{\frac{1}{2}\sigma^2}\mathbb{E}[\mathcal{G}_\Sigma(\rho, R_f)] + N_o\rho\pi R_f^2\Big)\,\sigma_o^{-1}.\end{aligned} \tag{4.18}$$

Finally, based on Eqs. (4.9) and (4.18), the first term in Eq. (4.8) for calculating the average number of sensing rounds $\mathbb{E}[N \mid \mathcal{H}_1]$ can be derived as:

$$\begin{aligned}\mathbb{E}[\lambda \mid \mathcal{H}_1] = \frac{1}{2}(\phi_0^2 - \phi_1^2)\,\rho\pi R_f^2 + (\phi_1 - \phi_0) \\ \times \left(P_o\, e^{\frac{1}{2}\sigma^2}\,\mathbb{E}[\mathcal{G}_\Sigma(\rho, R_f)] + N_o\rho\pi R_f^2\right)\sigma_o^{-1},\end{aligned} \tag{4.19}$$

where $\phi_0 = \frac{N_o}{\sigma_o}$ and $\phi_1 = \frac{N_o + P_R}{\sigma_o}$ are the average normalized test statistics under both hy-

[3]This is reasonable because the probability that there exists at least one sensor within $\epsilon = 75$ m from the WM transmitter is $1 - Poi(0; \rho\pi\epsilon^2) \approx 0.02$ given sensor density of $\rho = 1.25 \times 10^{-6}/\text{m}^2$.

potheses.

The average number of sensing rounds $\mathbb{E}[N \mid \mathcal{H}_1]$ can be derived by substituting Eqs. (4.10) and (4.19) into Eq. (4.8).

4.4.5 Optimal Data-Fusion Range

Based on the analyses above, we now derive an optimal data-fusion range that minimizes the average detection delay, i.e., the number of sensing rounds needed to meet the detection performance requirements.

Proposition 4.1 *Let $\mathcal{J}(R_f) \triangleq \mathbb{E}[\lambda \mid \mathcal{H}_1]$ in Eq.* (4.19)*. Then, the optimal fusion range that minimizes the average number of sensing rounds $\mathbb{E}[N]$ is given as:*

$$R_f^* = \arg\max_{R_f} \mathcal{J}(R_f) = R_f\Big|_{\frac{\partial J(R_f)}{\partial R_f}=0} = \left(\frac{a_1(\alpha-2)}{2a_2}\right)^{\frac{1}{\alpha}}, \tag{4.20}$$

where

$$a_1 = \frac{2(\phi_1-\phi_0)P_o\, e^{\frac{1}{2}\sigma^2}\rho\pi d_o^{\alpha}}{\sigma_o(2-\alpha)}, \tag{4.21}$$

and

$$a_2 = \frac{1}{2}(\phi_0^2-\phi_1^2)\rho\pi + \frac{(\phi_1-\phi_0)N_o\rho\pi}{\sigma_o}. \tag{4.22}$$

Proof. See Appendix A. □

Proposition 4.1 indicates that the optimal fusion range that minimizes detection delay depends on various system parameters, such as transmission power (P_o), noise power (N_o), shadow fading (σ), and path-loss exponent (α).

Based on Proposition 4.1, we have the following counter-intuitive observation:

Corollary 4.1 *The optimal fusion range (R_f^*) is independent of the sensor density ρ.*

One might think that the optimal fusion range should decrease as sensor density increases, since more sensors (near the WM transmitter) with high RSSs become available for data fusion. However, this is not the case because the performance of EGC depends on how far a cooperating sensor's report (i.e., the measured RSS) is from the noise power level, which is independent of sensor density (see Fig. 4.6 for details).

4.4.6 Impact of Sensing Scheduling

We now show that sensing scheduling can reduce the minimum sensor density required for given detection constraints. While the achievable performance gain via cooperative sensing has been studied extensively [100, 101, 106], the impact of sensing scheduling on the sensor density requirement has not been studied.

Proposition 4.2 *The minimum sensor density required to meet certain detectability requirements is* inversely *proportional to the average number of times to sense.*

Proof. Based on Eqs. (4.8), (4.10), (4.15) and (4.19), the required sensor density for a given average number of sensing rounds $\mathbb{E}[N]$ can be expressed as:

$$\rho = \frac{\sigma_o^2}{\mathbb{E}[N]} \times \frac{b^* \ln \frac{b^*}{1-a^*} + (1-b^*) \ln \frac{1-b^*}{a^*}}{P_R\left[P_o e^{\frac{1}{2}\sigma^2}\left(\frac{1}{\epsilon^2} - \frac{1}{R_f^2}\right) - \frac{1}{2}\pi R_f^2 P_R\right]}, \tag{4.23}$$

where a^* and b^* are the desired false-alarm and mis-detection probability values. Eq. (4.23) indicates that the sensor density ρ is inversely proportional to the average number of sensing rounds. Therefore, the proposition follows. □

Propositions 4.1 and 4.2 are derived based on the assumption that the WM's location and transmit-power level are known *a priori* to secondary users. However, such information may not be available in practice, and thus the benefits of our analytical findings cannot be realized without an efficient way of estimating the WM's location and transmit power. This motivates our approach of integrating sensing with location and transmit-power estimation, which we discuss next.

4.5 DeLOC: The Iterative Approach

We now introduce DeLOC, an iterative algorithm that expedites the detection of small-scale primary signals via joint data-fusion and location/transmit-power estimation. We first describe the estimation techniques, and then the proposed data-fusion rule and the iteration method employed by DeLOC.

4.5.1 Estimation Techniques

4.5.1.1 Estimation of WM Location

As we observed in Fig. 4.3, a reasonable approximation, if not accurate, of the WM's location is sufficient for improving detection performance. In DeLOC, the BS estimates and updates the WM's location based on the RSSs reported by the sensors. In particular, the BS employs a *weighted centroid method* proposed in [151], which estimates the WM's location via a weighted average of the sensors' locations, where the weight equals the corresponding sensor's report. The BS further refines the estimation via an exponential moving average over multiple sensing stages.

More specifically, let $\widehat{\boldsymbol{\vartheta}}_t = (\widehat{x}_t, \widehat{y}_t) \in \mathbb{R}^2$ denote the estimated location of the primary at sensing stage t. Then, the WM's location is estimated as [92, 151]:

$$\widehat{\boldsymbol{\vartheta}}_t = (1-\beta)\,\widehat{\boldsymbol{\vartheta}}_{t-1} + \beta \left(\sum_{n \in S_t} \frac{P_n}{\sum_{m \in S_t} P_m}\, \psi_n \right),$$

where P_n is the received primary signal power at sensor n, $\boldsymbol{\psi}_n = (x_n, y_n)$ the location of sensor n, and $\beta \in (0,1)$ the smoothing factor.

Note that DeLOC uses a simple existing localization method to estimate the PU's location in each round, but it is not restricted to any specific localization algorithm, so other localization methods, such as the semi range-based method proposed by Ma *et al.* [96], can also be used.

4.5.1.2 Estimation of Transmit Power

In DeLOC, the BS estimates the WM's transmit-power based on the WM's estimated location and the reported RSSs using the method proposed in [159] as:

$$\widehat{P}_{o,t}(dB) = 10 \log_{10}\left(\frac{1}{k}\right) + \frac{10}{|S_t|} \sum_{n \in S_t} \Big(\log_{10}(P_n) + \alpha \log_{10}(d_n) \Big),$$

where $k = p_o\, d_o^{\alpha}\, P_o^{-1}$. p_o is the measured signal power at reference distance d_o, P_n the received primary signal strength at sensor n, and d_n the distance between the WM trans-

Algorithm 5 DeLOC: ALGORITHM FOR JOINT DETECTION AND ESTIMATION OF SMALL-SCALE PRIMARY USERS

At the end of a sensing period, the BS does the following

for Each triggering event **do**
  $t \leftarrow 1$ // Initialization
  while $t \leq$ MaxNumIter **do**
    $t \leftarrow t + 1$
    $\theta_t \leftarrow$ Receive sensing results from cooperating sensors S_t
    $\Lambda_t \leftarrow \Lambda_{t-1} + \lambda_t^{f(t)}$ // Update the decision statistic
    if $\Lambda_t \geq B$ **then**
      A primary exists and hence returns the estimated location and transmit-power level
    else if $\Lambda_t < A$ **then**
      A primary does not exist (i.e., the event is triggered by a ghost primary) and hence terminates the iteration
    else
      $(\widehat{\vartheta}_{t+1}, \widehat{P}_{o,t+1}) \leftarrow$ Estimate the location and transmit power of the primary transmitter
      $R^*_{f,t+1} \leftarrow$ Calculate the optimal fusion range
      $S_{t+1} \leftarrow$ Select a set of sensors located within $R^*_{f,t+1}$ from the estimated primary transmitter location
      Schedule another sensing round and wait for the observation
    end if
  end while
  return No primary signal exists
end for

mitter and sensor n, i.e., $d_n = \sqrt{(\hat{x}_t - x_n)^2 + (\hat{y}_t - y_n)^2}$. Note that the test statistics of the energy detector include noise power, so the received primary signal strength P_n needs to be estimated from the test statistics by subtracting the average noise power N_o from the measurements.

While DeLOC employs simple location and power estimation techniques, the estimation accuracy can be further improved by using more sophisticated techniques at the cost of more computation.

4.5.2 The Proposed Data-Fusion Rule

While DeLOC improves small-scale primary detection performance via iterative cooperative sensing and estimation, we observed that it often terminates in the early stages mis-detecting the WM. This is because, initially, the BS's location and transmit-power estimates are inaccurate, resulting in many noisy sensor reports during data fusion. This

prevents DeLOC from fully exploiting its unique feature—an iterative refinement of estimation and fusion.

To overcome this problem, we propose a new data-fusion rule for DeLOC, a *sequential probability ratio test with ascending weight* (SPRT-AW), to prevent the BS from making biased decisions in early stages. The idea is to assign smaller weights to the decision statistics in early stages, and gradually increase the weights as the location and transmit-power estimates become more accurate. Specifically, we use the following rule to update the decision statistic:

$$\Lambda_t = \Lambda_{t-1} + \lambda_t^{f(t)} \quad \text{where} \quad f(t) = \frac{1}{1+e^{1-t}} \quad t \in \mathbb{N}, \tag{4.24}$$

where we use the sigmoid function $f(t)$ such that the exponent of test statistics increases from 0.5 to 1 as t increases. Consequently, the test statistics in later stages count more in decision-making. The resulting decision statistics will be used in updating the Λ_t in Eq. (4.4), and compared with the upper (B) and lower (A) thresholds to make a decision.

4.5.3 Description of DeLOC Protocol

As described in Fig. 4.1, DeLOC is triggered only when a sensor's report is above a certain predefined threshold ξ, which is suspected as a WM signal. The triggering threshold must be chosen carefully by the BS to balance the false- and mis-triggering of DeLOC. Upon triggering, the BS assumes the triggering sensor's location as the WM's location, and employs additional sensors within the fusion range for WM detection in the next scheduled sensing round. If there are multiple triggering sensors in close proximity, the BS considers the sensor with highest RSS. In each sensing round, the BS iterates the following two steps: (i) location and transmit-power estimation and (ii) data-fusion, until the decision statistic for data fusion Λ reaches one of the thresholds. The BS also terminates the iteration after scheduling sensing rounds for MaxNumIter. Algorithm 5 details DeLOC.

4.6 Performance Evaluation

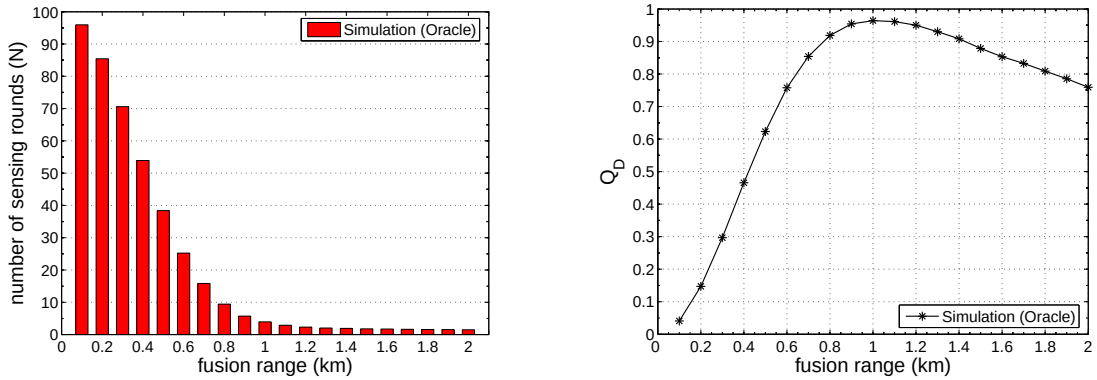
In this section, we evaluate DeLOC using MATLAB-based simulation. We first describe the simulation setup and then present the impact of the fusion range on detection performance and its dependency on transmit-power. We also show the relationship between detection delay and sensor density. Finally, we demonstrate the performance of DeLOC in comparison with other testing schemes.

4.6.1 Simulation Setup

In the simulation, we consider a realistic 802.22 environment where sensors are randomly distributed over a 30 km$\times 30$ km area. The average sensor density is set to $1.25/\text{km}^2$, as typically used in 802.22 WRANs [107], unless specified otherwise. We assume a WM randomly located in the area with effective transmit-power below 25 mW, as indicated by the measurement study in [36]. The maximum number of sensing rounds scheduled within the 2-second *channel detection period* (CDT) is limited to $\mathsf{MaxNumIter} = 100$.[4] The duration of a single sensing period is assumed to be 1 ms. The path-loss exponent is $\alpha = 4$, and the shadow fading dB-spread is $\sigma_{dB} = 5.5$ dB, which is typically assumed for rural areas. We also assume that the signal-propagation parameters are known *a priori* to the secondary system. The triggering threshold in DeLOC is configured as $\xi = N_o + 3.5\,\sigma_o$, which gives the false-triggering rate of 2.3×10^{-4}. The simulation results are obtained from 5×10^3 randomly-generated topologies.

To evaluate the efficacy of DeLOC, we compare the performance of the following four sensing schemes: (i) Oracle (the ideal case), (ii) DeLOC with SPRT-AW, (iii) DeLOC, (iv) DeLOC without localization, and (v) DeLOC without transmit-power estimation. In *Oracle*, the location and transmit-power information is available to the BS, so the BS always uses the optimal fusion range for sensing without the need for estimation. Thus, Oracle will be used as a performance reference. In DeLOC *without localization*, the location of the triggering sensor is regarded as the primary's location. In DeLOC *without power estimation*,

[4] This is reasonable since the BS can schedule sensing as frequently as once every 10 ms, i.e., one MAC frame size in 802.22.



(a) average number of sensing rounds (N) (b) detection probability (Q_D)

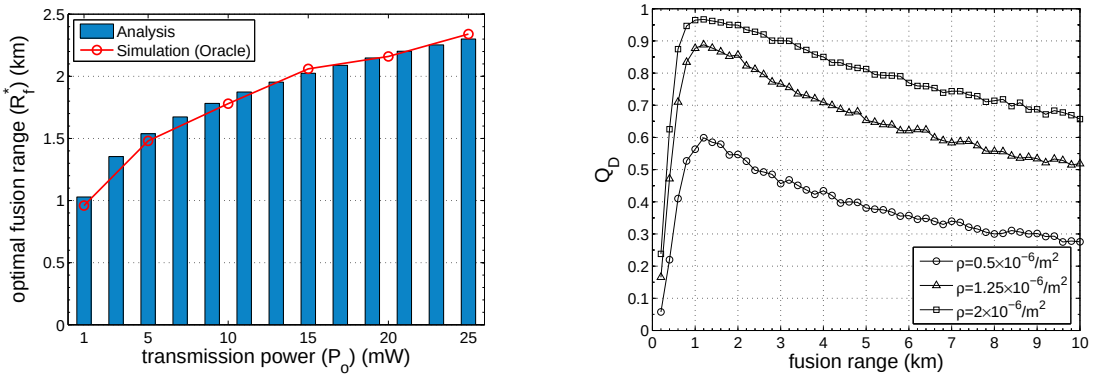
Figure 4.4: Impact of the data-fusion range on detection performance: (a) the average number of sensing rounds decreases as the fusion range increases, whereas (b) the detection probability is maximized when $R_f = 1$ km, which is close to the analytical result. Here the WM's transmit power is set to 1 mW.

the power level is assumed to be randomly chosen in $[0, 25]$ mW.

4.6.2 Impact of Fusion Range

Fig. 4.4 shows the impact of the data fusion range on detection performance in terms of detection delay and accuracy. The figures indicate that too small a fusion range suffers from the lack of cooperating sensors, which makes it difficult for the BS to collect enough information, i.e., measured RSSs, to make a decision within MaxNumIter, resulting in a low detection probability. On the other hand, too large a fusion range, i.e., beyond 1 km, suffers from having many noisy reports, misleading the BS to promptly conclude that there is no primary signal, increasing the chance of mis-detection. Fig. 4.4(b) shows that the detection probability Q_D is maximized when the fusion range is 1 km, which closely matches the analytical result, i.e., 1.03 km.

An additional observation from our simulation results is that false-alarms occur only 16 times over 5×10^3 iterations, i.e., $Q_{FA} = 16/5000 = 0.0032$, thus achieving the false-alarm requirement of $Q_{FA} \leq 0.01$.



(a) Impact of transmit-power

(b) Impact of sensor density

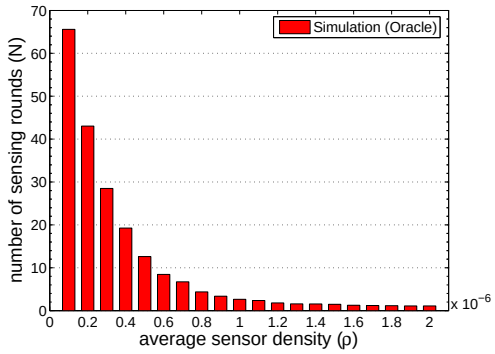
Figure 4.5: Optimal Fusion Range: (a) shows the impact of transmit-power on the optimal fusion range, indicating that the analytical results (the bar graph) closely match the simulation results (the line graph), thus validating the analytical results. (b) shows that the optimal fusion range R_f^* that maximizes detection performance (Q_D) does not depend on sensor density.

4.6.3 Optimal Fusion Range

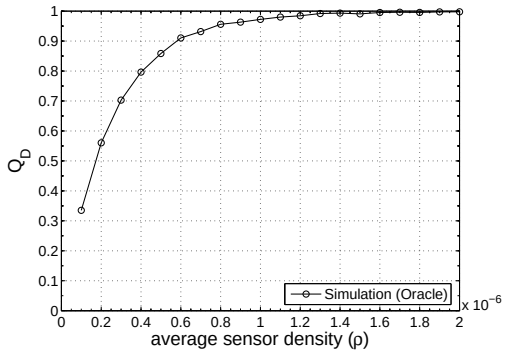
Fig. 4.5(a) plots the optimal fusion range for various transmit-power levels, and indicates that our analytical results (Proposition 4.1 in Section 4.4.5) closely match the simulation results. The figure also shows that to detect a high transmit-power WM, it is better (in the sense of reducing sensing delay) to extend the fusion range, thus increasing the number of cooperating sensors. On the other hand, to detect a WM with very weak transmit-power, it is better to have a small number of sensors, thus reducing the number of noisy reports. Fig. 4.5(b) indicates that the optimal fusion range (in the sense of maximizing the detection probability Q_D) remains the same over different sensor densities, thus confirming Corollary 1 in Section 4.4.5.

4.6.4 Impact of Sensor Density

Although sensor density does not affect the optimal fusion range, a higher sensor density (hence more sensors within fusion range) can still improve sensing performance by exploiting diversity of measurement. Fig. 4.6(a) shows the average number of sensing rounds (i.e., detection delay) required to meet the detection performance $Q_{FA}, Q_{MD} \leq 0.01$, which obviously decreases with sensor density. The figure also indicates that the average num-



(a) Impact of sensor density on detection delay



(b) Impact of sensor density on detection accuracy

Figure 4.6: Impact of sensor density on detection performance: (a) the detection delay decreases rapidly as sensor density increases, and (b) the detection accuracy increases as the sensor density increases. In the simulation, the transmit power is set as $P_o = 2\,\text{mW}$.

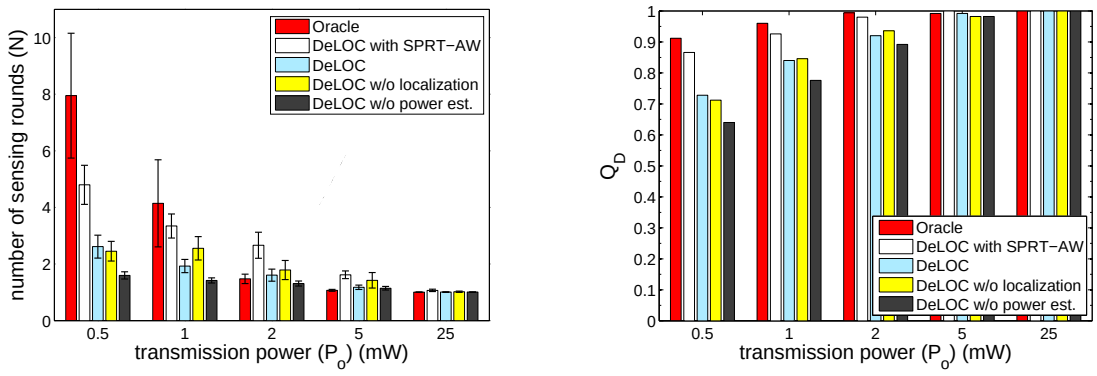
ber of sensing rounds is almost inversely proportional to sensor density, thus confirming Proposition 4.2 in Section 4.4. Fig. 4.6(b) further shows that the detection probability Q_D increases with increasing sensor density.

4.6.5 Performance of DeLOC

To demonstrate the efficacy of DeLOC, we compare its performance with the other four testing schemes under the detection constraints $Q_{FA}, Q_{MD} \leq 0.01$. As shown in Fig. 4.7, when the WM's transmit-power increases, detection performance (with respect to delay and detection probability) increases for all testing schemes. We make three additional observations.

First, Fig. 4.7(a) shows that the average number of sensing rounds for decision-making is below 10, which may take only $100\,\text{ms}$ as the BS can schedule sensing periods as frequently as every $10\,\text{ms}$, i.e., one MAC frame size in 802.22. In addition, the detection probability of DeLOC with SPRT-AW meets the detection requirement of 802.22, i.e., $Q_{MD} \leq 0.1$, even for a very weak transmit-power of $1\,\text{mW}$, as indicated in Fig. 4.7(b).

Second, Fig. 4.7(b) shows that DeLOC with SPRT-AW performs close to Oracle in terms of detection rate, and outperforms all other schemes that use regular SPRT. As mentioned earlier, the SPRT in DeLOC often makes a wrong decision (mis-detection of a WM)



(a) Average detection delay

(b) Achieved detection probability

Figure 4.7: Performance of DeLOC: DeLOC (a) requires only a small number of sensing rounds for WM detection, and (b) achieves a high detection rate even for a very weak signal power, e.g., $P_o = 1$ mW.

in early detection stages because of the large number of noisy reports due to the inaccurate location and power estimates. DeLOC with SPRT-AW mitigates this problem by discounting the decision statistics in early stages.

Third, Fig. 4.7(b) shows that DeLOC without localization outperforms the DeLOC without transmit-power estimation. This is because power estimation plays an important role in finding the optimal fusion range, and thus, errors in power estimation results in significant performance degradation. On the other hand, the location-estimation error is small compared to the typical fusion range, and thus it does not cause significant performance degradation, as already shown in Fig. 4.3.

These simulation results clearly demonstrate that the joint design of data-fusion and location and transmit-power estimation maximizes the benefits of spatial-temporal sensing for detecting small-scale primaries, such as WMs in 802.22.

4.7 Related Work

Despite its practical importance, there has only been limited research in MAC-layer solutions to WM detection. Most existing work focuses on PHY-layer signal detection techniques [26, 161], which have short sensing ranges and require a separate dense sensor network for WMs. Mishra *et al.* [107] studied the minimum sensor density required for

detecting WM signals based on energy detection. They showed that when the path-loss exponent is 4 or higher, the average sensor density in rural areas (i.e., $1.25/\text{km}^2$) is not sufficient for detecting WMs. Recently, the 802.22 Working Group established Task Group 1 to develop a standard for the disabling beacon protocol [120]. Although the disabling beacon can protect WM signals better, it has several practical limitations as discussed in Section 4.1. Moreover, the disabling beacon is restricted to WMs because the sensing relies on specialized signal features. In contrast, our DeLOC algorithm is a *generic* MAC-layer sensing scheme for small-scale primary detection, which can be incorporated into the beacon protocol while overcoming its limitations. Chouinard [34] proposed a coexistence model between WMs and 802.22 WRANs by exploiting the WM signal's small footprint and its narrow bandwidth, i.e., 200 KHz. However, they do not account for location estimation, which greatly affects the spatial reuse of spectrum.

Sequential detection of PUs has been studied by others [28, 85, 100]. Chen *et al.* [28] proposed a weighted sequential probability ratio test (WSPRT) that assigns different weights to sensor reports based on the sensors' reputations in order to minimize the impact of manipulated (or erroneous) sensor reports in making a final decision on the presence/absence of a primary signal. By contrast, SPRT-AW introduced in DeLOC is designed to intentionally defer the final decision at the BS, so as to reduce the effects of any wrong decision made in early stages when localization and power estimation are relatively inaccurate.

Chen *et al.* [29] proposed a scheme for verifying a PU's location, called *LocDef*. Its main idea is that if the estimated location of the signal source differs significantly from the known location of the primary transmitter, i.e., a TV transmitter, then the BS assumes that the signal is transmitted from a fake PU. By contrast, the location and transmit-power estimation introduced in DeLOC aim to improve detection performance of small-scale primary signals, e.g., WMs, by helping the BS select an optimal set of cooperating sensors. In addition, when there is a WM signal, DeLOC returns the estimated location and transmit-power of the detected WM, so that the BS may use this information for admission control and transmit-power control of the secondary users to achieve better spectrum reuse in the space domain. Another key difference is that DeLOC makes use of a sparse sensor network,

whereas *LocDef* requires a dense sensor network for location verification.

The work presented in this chapter is also related to the existing body of research on the coverage of sensor networks. Using a theoretical analysis, Xing *et al.* [153] showed that data fusion via sensor cooperation can improve the coverage of sensor networks over the conventional detection schemes based on a disc model. While our approach also emphasizes the importance of data fusion, we adopt sensing scheduling to improve the detection performance of small-scale primaries. We characterize the impact of various factors on WM detection, and establish a practical framework that accounts for the unpredictability of each WM's transmit-power and location.

4.8 Conclusion

The detection of small-scale primary signals is a critical, but challenging problem in realizing DSA in CRNs. To address this problem, we proposed a novel spatio-temporal fusion scheme that exploits (i) spatial diversity by cooperative sensing with an optimal fusion range, and (ii) temporal diversity by scheduling a series of sensing stages with an optimal stopping time. We modeled the detection problem as a hypothesis test, approximated the sum of sensor readings as a log-normal random variable, and then solved a convex optimization problem, to obtain the optimal fusion range that minimizes the average detection delay. We also proposed a new sensing algorithm called `DeLOC` that iterates between cooperative sensing and location/transmit-power estimation to further improve sensing performance under realistic settings. Our evaluation results show that `DeLOC` reduces detection delay significantly while achieving high detection performance.

CHAPTER 5

Robust Tracking of Mobile Small-Scale Primary Users

5.1 Introduction

Unlike the detection of large-scale primaries, e.g., DTV users, where localization is not the primary concern in opportunistic spectrum reuse, accurately tracking the physical location of *mobile* small-scale primaries, such as wireless microphones (WMs), is crucial in achieving the core objectives and functionalities of CRNs, such as spatial spectrum reuse [34], interference management [64, 146], routing decisions [37], and falsified primary signal detection [29, 91]. For example, knowing the location of the primary transmitter enables secondary users (SUs) to reuse licensed spectrum more efficiently without causing excessive interference to primary communications [34, 64, 126, 146]. Without knowing the location of a WM, however, all the SUs (also called CPEs) in an 802.22 cell (of radius up to 100 km) must immediately vacate the current operating channel upon detection of the WM, resulting in significant waste of spatial spectrum opportunities [34]. Furthermore, location information is also very useful for cooperative sensing by enabling the base station (or fusion center) to select an optimal set of sensors, especially when detecting a very weak primary signal, like a WM signal [105].

However, CRN faces unique challenges, such as the absence of primary-secondary coordination and low sensor density, that make it difficult to accurately track mobile primaries. According to the FCC, opportunistic spectrum access should require no modification to the primary system [47]. Thus, SUs (sensors) must rely solely on measured received

signal strengths (RSSs) (obtained via spectrum sensing) for primary tracking. This makes the primary tracking vulnerable to attacks, since the tracking process can be disrupted by malicious or faulty sensors that report incorrect RSSs. A sensing report falsification attack can be easily launched by attackers due to the open nature of low-layer protocol stacks in SDR devices, such as USRP [3] and Sora [84]. Moreover, low sensor density in CRNs hampers the accurate tracking of mobile PUs, e.g., the average sensor density in 802.22 WRANs is only about $1.25/\text{km}^2$ [107]. Inaccurate location estimation may ultimately cause SUs to generate excessive interference to the primary system, violating the basic premise of CRNs and discouraging PUs from sharing their licensed spectrum bands with SUs. Therefore, there is a clear need for an efficient and secure tracking scheme for small-scale mobile PUs in CRNs.

In this chapter, we address the problem of reliably tracking small-scale mobile PUs in CRNs. Specifically, we design an RSS-based tracking scheme, called `SOLID`, which allows accurate, attack/fault-tolerant tracking of mobile PUs by jointly estimating the location of a primary and shadow-fading gains in the RSSs. The shadowing estimation in `SOLID` greatly improves localization performance. Besides, by monitoring temporally-correlated shadow fading, `SOLID` accurately detects manipulated or erroneous sensor reports, thus achieving high robustness. The key motivation behind exploiting temporal shadowing correlation in attack detection is based on the observation that malicious sensors cannot control the physical-layer signal-propagation characteristics. While we focus on the robust tracking of WMs' location in 802.22 WRANs, our proposed techniques are *generic* and can be used for detecting other types of small-scale primaries or, in a broader context, target tracking in wireless sensor networks.

5.1.1 Contributions

This chapter makes the following main contributions.

- Development of a new tracking scheme, `SOLID`, that jointly estimates the mobile PU's location and shadow-fading gains using an adaptive filter. By exploiting the temporal correlation in shadow fading, `SOLID` (i) improves localization accuracy

and (ii) accurately identifies abnormal sensing reports. To the best of our knowledge, this is the first attempt to incorporate shadow fading into cooperative localization.

- In-depth evaluation of SOLID in a realistic shadow-fading environment under various attack scenarios. Our extensive simulation study shows that, under no attack, SOLID lowers the average localization error by up to $88\,\%$ compared to the conventional Sequential Monte Carlo (SMC) based tracking scheme, since the two components of SOLID—SMC-based localization and shadow-fading estimation—refines each other throughout the tracking process.

- High attack- and fault-tolerance of SOLID. Our evaluation results show that SOLID can detect compromised sensing reports with high accuracy, e.g., attack-false alarm and mis-detection rates below 1 % and 7 %, respectively. It also shows that in a realistic shadowing and multi-path environment, SOLID lowers the average error by up to $89\,\%$ even under "slow-poisoning" attacks.

- Investigation of the tradeoff in the design of the attack detector in SOLID. When the base station (BS) filters out sensors or sensing reports too aggressively (or conservatively), the localization can suffer from lack of samples (compromised sensing reports). Via in-depth simulation, we identify the impact of attack detection thresholds, and the results provide practical guidelines for the design of a robust and efficient tracking scheme.

5.1.2 Organization

The rest of this chapter is organized as follows. Section 5.2 describes the system models and assumptions, and introduces the attack models. Section 5.3 presents our proposed approach for attack detection, and the underlying localization protocol. Section 5.4 details our approach for the estimation of shadow fading, and the design of SOLID's attack detector. Section 5.5 evaluates the performance of SOLID, and Section 5.6 reviews related work. Section 5.7 concludes the chapter.

5.2 System and Attack Models

In this section, we describe the network, spectrum sensing, and signal-propagation models that we use throughout this chapter. We then provide an overview of our model of tracking a small-scale mobile primary transmitter and introduce the attack model.

5.2.1 CR Network Model

We consider a CRN that consists of primary and secondary users/devices in the same geographical area. The secondary network is an infrastructure-based network, such as an IEEE 802.22 WRAN, in which each cell consists of a base station (BS) and multiple sensors, called *customer premise equipments* (CPEs). The main goal of IEEE 802.22 WRANs is to provide Internet access to rural areas by reusing unused TV spectrum bands, without causing excessive interference to PUs. In an 802.22 WRAN, the BS manages the DSA of the SUs in the network by (i) scheduling sensors to perform spectrum sensing, and (ii) performing data fusion and primary location estimation to determine the presence or absence of a primary signal based on the sensing reports. For spectrum sensing, the BS employs the sensors located within a fusion range centered at the estimated primary location for cooperative spectrum sensing [105].

We assume that sensors are stationary and that the BS has the location information of the sensors within its own cell. For example, the IEEE 802.22 WRANs standard draft requires the BS to know the sensor locations. We assume that the sensors have been deployed in an area A, e.g., an IEEE 802.22 WRAN cell, following a point Poisson process with average density ρ. Unlike in a typical wireless sensor network environment, where sensors are densely distributed, we assume a low sensor density ρ because the typical density of CPEs in rural areas is only $1.25/\text{km}^2$ [139]. We assume that the BS and sensors communicate over a common control channel.

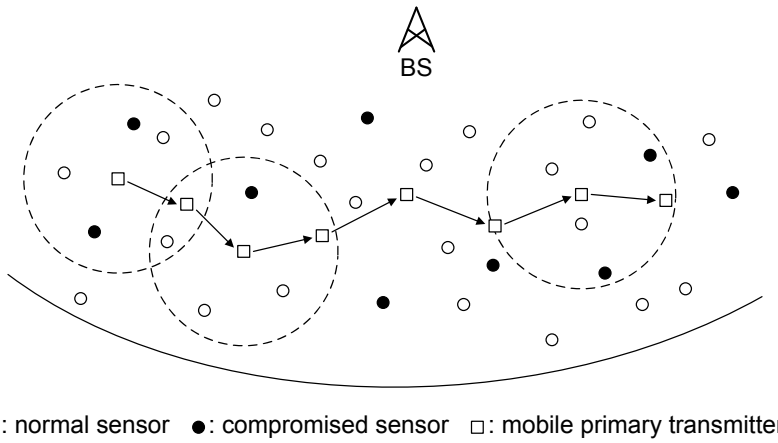


Figure 5.1: An illustrative example of small-scale primary tracking: The BS tracks the location of a mobile PU (e.g., a WM) based on the sensing reports (i.e., received primary signal strengths) from the sensors located within the fusion range (the dotted circle).

5.2.2 Spectrum-Sensing and Signal-Propagation Models

Due to the lack of primary-secondary cooperation, primary tracking must be done based only on the received primary signal strengths measured at cooperative sensors.[1] We consider energy detection [127] for spectrum sensing in the PHY-layer. Energy detection is the most widely-used PHY-layer sensing technique due to its simple design and low sensing overhead. The test statistics of the energy detector are an estimate of the sum of received primary signal and noise power [127]. We assume that the BS employs only the sensors located close to the primary transmitter, i.e., located within the fusion range R_s from the estimated location of a primary transmitter, for location tracking. The BS directs the cooperative sensors to perform spectrum sensing at a periodic time interval $t \in \mathcal{T}$, and reports their sensing results to the BS for localization. Fig. 5.1 depicts an example scenario of tracking a mobile primary transmitter in a CRN.

Assuming that the noise power is much smaller than the received primary signal strength, sensor n's measurement in sensing time slot t can be expressed (in dB) as [46]:

$$P_{t,n} = P_o + \alpha 10 \log(d_o) - \alpha 10 \log(d_{t,n}) + X_{t,n} + Y_{t,n} \tag{5.1}$$

[1]Cooperative sensors refer to a set of sensors in a 802.22 WRAN, which are employed by the BS for spectrum sensing.

where α is the path-loss exponent, d_o the reference distance, P_o the received primary signal strength at the reference distance, $d_{t,n}$ the distance between the primary transmitter and sensor n in time slot t. Log-normal shadow fading, denoted by $X_{t,n}$, can be characterized by dB-spread, σ_{dB}, where $X_{t,n} \sim \mathcal{N}(0, \sigma_{dB}^2)$.[2] We assume that non-fading components, such as antenna and device losses, are approximated as an i.i.d. Gaussian random variable with zero mean and variance σ_m^2, denoted as $Y_{t,n} \sim \mathcal{N}(0, \sigma_m^2)\ \forall n$.

Let S_t denote a set of cooperating sensors in time slot t. Then, the received primary signal strength at cooperating sensors in Eq. (5.1) can be expressed in a vector form as:

$$\mathbf{P_t} = \mathbf{H}(\mathbf{d_t}) + \widehat{\mathbf{X}}_\mathbf{t} + \mathbf{Y_t}, \tag{5.2}$$

where $\mathbf{H}(\mathbf{d_t}) = [\, h(d_{t,1}), \ldots, h(d_{t,|S_t|}) \,]^T$ represents the channel gain due to path-loss, where $h(d_{t,i}) = P_o + \alpha 10 \log(d_o) - \alpha 10 \log(d_{t,i})$. The shadow fading gain and noise vectors are denoted by $\widehat{\mathbf{X}}_\mathbf{t}$ and $\mathbf{Y_t}$, respectively.

5.2.3 Attack Model

The main objective of attackers (compromised sensors) is to disrupt the primary transmitter localization/tracking process by manipulating their sensing reports. Specifically, we consider the following two attack scenarios: a sensor is

- *compromised*, reporting manipulated (i.e., higher or lower than real) RSSs to the BS,
- *malfunctioning or faulty*, generating sensor readings that significantly deviate from the true RSS.

The above two cases render the sensing reports to the fusion center (i.e., the BS) inaccurate, degrading the localization/tracking performance. Such large localization error will require SUs to be more conservative in reusing spectrum opportunities, resulting in a waste of *spatial* spectrum opportunities (see Section 5.5.7). Therefore, we opt to design an *attack- or fault-tolerant* primary tracking mechanism that successfully tolerates such ma-

[2] Measurement studies [10] indicate that a typical σ_{dB} values are 4-8 dB depending on geographical environments, e.g., urban or suburban.

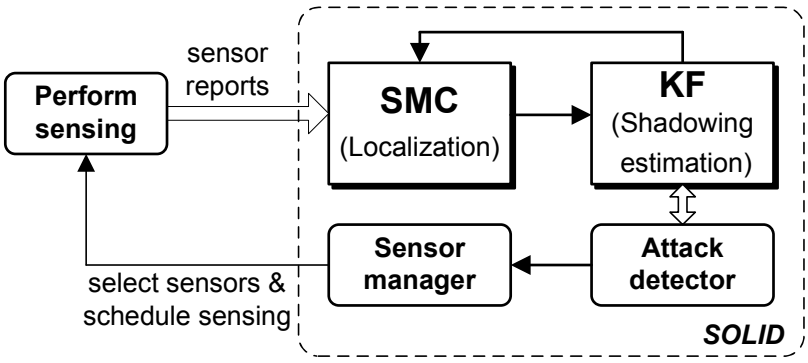


Figure 5.2: The SOLID framework: SOLID provides high accuracy and robustness in mobile primary tracking by (i) estimating/monitoring shadow-fading gains between the primary transmitter and sensors using the Kalman Filter, and (ii) detecting and filtering out abnormal sensing reports based on the shadowing-correlation profile.

nipulated (or erroneous) sensing reports. Although there exist other security threats, such as jamming or denial-of-service attacks, in the primary tracking process, the sensing report manipulation attack that we consider in this chapter is stealthier due to the attacker's ability to control the sensing reports in a finer-grained manner.

5.3 The Proposed Approach

We first describe the overall architecture of SOLID and present its design rationale. We then introduce the *sequential Monte Carlo* (SMC) localization process that underlies SOLID.

5.3.1 SOLID Architecture

SOLID (Fig. 5.2) resides at the BS and consists of the following three building blocks:

- **location estimator** that tracks the location of a small-scale mobile primary transmitter based on sensing reports,
- **shadowing estimator** that tracks the shadowing gain at cooperative sensors using the Kalman Filter (KF), and

- **attack detector** that detects and discards abnormal sensing reports, and updates the normal profile.

The above three components synergistically interact with each other and collectively enable accurate and robust PU tracking. Based on the estimated primary location, the sensor manager selects sensors to cooperate with each other based on their (ab)normality and proximity to the primary transmitter.[3]

In particular, the shadowing estimator introduced in SOLID offers two main benefits:

- It improves localization accuracy by mitigating the effect of shadow fading in RSSs ((a) in Fig. 5.2), and
- It enables accurate detection of abnormal sensing reports ((b) in Fig. 5.2).

SOLID also minimizes communication and processing overhead since it exploits physical-layer signal-propagation characteristics, extracted from the cooperative sensing results.

5.3.2 Design Rationale for Attack Detection

To maximize attack-tolerance and preserve localization accuracy, SOLID exploits the *temporal* correlation in shadow fading in received primary signal strengths. The key insight behind the attack detector is that, in shadow-fading environments, the sequence of RSSs measured at each sensor is highly likely to be correlated as indicated in measurement studies (e.g., [10, 60]). Thus, the attack detector takes an *anomaly-detection* approach to identifying and discarding abnormal sensing reports in the localization process. So, if attackers raise or lower the sensing results (i.e., RSSs) reported to the BS in order to influence the localization result, SOLID can easily detect them by examining the consistency of the sensing reports with the estimated primary location and the previous history of sensing reports. Hence, the attacker must lower its attack strength to evade detection by SOLID, exerting only a negligible impact on localization.

One important, but not so obvious feature of the attack detector in SOLID is that it is *cooperative*, in the sense that the accuracy of shadowing-gain estimation depends heavily

[3] Although there are many sophisticated sensor-selection methods for target tracking (e.g., [30]), optimal sensor-selection is not our focus.

on the location estimate, which is updated based on reports from all the cooperating sensors. In other words, the robustness of attack detection is directly correlated with localization accuracy.

5.3.3 SOLID: Sequential Monte Carlo Combined with Shadow-Fading Estimation

SOLID employs *sequential Monte Carlo* (SMC) [70] as the baseline scheme for tracking small-scale mobile PUs. SMC has been widely used as a localization method in mobile wireless systems [13, 121]. The key idea of SMC is to represent the required posterior density function by a set of random samples (or particles) with their associated weights, and then compute the estimated location by taking their weighted average. SOLID augments the conventional SMC with shadow-fading estimation to further improve the tracking accuracy and achieve robustness against malicious/faulty sensors.

Let $\{\boldsymbol{\phi} \mid \boldsymbol{\phi_t} = (x_t, y_t)\ t \in \mathbb{N}\}$ denote the sequence of a mobile primary's locations in two-dimensional coordinates where t is the index for (sensing) time slot. The BS estimates the primary transmitter's location based on the vector of received primary signal strengths, denoted by $\mathbf{P_t}$ in Eq. (5.2).

Let the particle set denote the set of tuples $\{(\boldsymbol{\theta_t^{(i)}},\ w_t^{(i)})\}_{i=1}^{N_s}$ where each sample $\boldsymbol{\theta_t^{(i)}}$ is associated with its weight $w_t^{(i)}$, where $\sum_{i=1}^{N_s} w_t^{(i)} = 1$, and N_s is the number of particle samples. Then, the PU tracking process in SOLID consists of the following 6 steps.

Step 1: At the end of sensing period t, SOLID draws N_s new samples[4] using transition probabilities $p(\boldsymbol{\theta_t^{(i)}}|\boldsymbol{\theta_{t-1}^{(i)}})$, given by:

$$p(\boldsymbol{\theta_t^{(i)}}|\boldsymbol{\theta_{t-1}^{(i)}}) = \begin{cases} \frac{1}{\pi(v_{max}+\beta)^2} & \text{if } d(\boldsymbol{\theta_t^{(i)}}, \boldsymbol{\theta_{t-1}^{(i)}}) < v_{max} \\ 0 & \text{otherwise,} \end{cases} \tag{5.3}$$

where v_{max} is the maximum speed of the mobile primary transmitter, and β is used to generate better samples [121]. We set $\beta = 0.2\, v_{max}$ empirically in our simulations.

[4]Initially, SOLID randomly selects N_s sample points $\boldsymbol{\theta_0} = \{\boldsymbol{\theta_0^{(i)}}\}_{i=1}^{N_s}$ in the detection region to represent candidate locations of the mobile PUs.

Step 2: After generating N_s new samples using Eq. (5.3), SOLID updates the weights associated with the samples as:

$$w_t^{(i)} = w_{t-1}^{(i)}\, \mathcal{L}(\mathbf{P_t} \,|\, \boldsymbol{\theta_t^{(i)}}), \tag{5.4}$$

where the likelihood $\mathcal{L}(\mathbf{P_t} \,|\, \boldsymbol{\theta_t^{(i)}})$ can be calculated based on multivariate Gaussian in Eq. (5.2), i.e., $\mathcal{L}(\mathbf{P_t} \,|\, \boldsymbol{\theta_t^{(i)}}) \sim \mathcal{N}(\mathbf{H}(\mathbf{d_t}) + \widehat{\mathbf{X}}_\mathbf{t}, \sigma_m^2\, \mathbf{I}_{N\times N})$, and $\mathbf{I}_{N\times N}$ is an identity matrix where $N = |S_t|$ is the number of cooperating sensors in time slot t. The weights are normalized such that $\sum_{i=1}^{N_s} w_t^{(i)} = 1$.

Step 3: Based on Eqs. (5.3) and (5.4), SOLID approximates the posterior density $p(\boldsymbol{\phi_t}|\mathbf{P_{1:t}})$ as:

$$p(\boldsymbol{\phi_t}|\mathbf{P_{1:t}}) \approx \sum_{i=1}^{N_s} w_t^{(i)}\, \delta(\boldsymbol{\phi_t} - \boldsymbol{\theta_t^{(i)}}), \tag{5.5}$$

where $\delta(\cdot)$ is the *Dirac delta measure*.

Step 4: Then, SOLID estimates the location of the primary transmitter by taking the weighted average of the samples:

$$\widehat{\boldsymbol{\phi}}_\mathbf{t} \triangleq (\hat{x}_t, \hat{y}_t) = \Big(\sum_{i=1}^{N_s} w_t^{(i)} x_t^{(i)}, \sum_{i=1}^{N_s} w_t^{(i)} y_t^{(i)} \Big). \tag{5.6}$$

Step 5: SOLID then calculates the effective number of particles, i.e., $\widehat{N}_{eff} = (\sum_{i=1}^{N_s} (w_t^{(i)})^2)^{-1}$, and compare it against the given threshold N_{thr}. If $\widehat{N}_{eff} < N_{thr}$, SOLID re-samples the particles using the posterior probability in Eq. (5.5) to replace the current particle set with this new one, and sets the weights $w_t^{(i)} = 1/N_s$ for $i \in \mathcal{S}_t$. Steps 1-4 repeat themselves until the effective number of particles, $\widehat{N}_{eff}$, is equal to, or greater than a given threshold N_{thr}.

Step 6: Given the estimated primary transmitter in Eq. (5.6), SOLID estimates the shadow-fading gains $\widehat{\mathbf{X}}_\mathbf{t}$ between the primary transmitter and the sensors using the Kalman filter. We will detail this in Section 5.4.

Algorithm 6 describes the primary tracking process of SOLID.

Algorithm 6 SMC WITH SHADOW-FADING ESTIMATION

At the end of each sensing round $t \in \mathcal{T}$, SOLID does

// 1. Localization

Initialization

$\boldsymbol{\theta_0^{(i)}} \sim p(\boldsymbol{\theta_0}), w_0^{(i)} = 1/N_s$ for $i = 1, \ldots, N_s$

$\widehat{N}_{eff} \leftarrow 0$ // Effective number of particles

while $(\widehat{N}_{eff} < N_{thr})$ **do**

  for $i = 1$ to N_s **do**

    Draw $\boldsymbol{\theta_t^{(i)}} \sim p(\boldsymbol{\phi_t} \mid \boldsymbol{\theta_{t-1}^{(i)}})$ using Eq. (5.3)

    Update $w_t^{(i)}$ using Eq. (5.4)

  end for

  Calculate the total weight $W_t = \sum_{i=1}^{N_s} w_t^{(i)}$

  for $i = 1$ to N_s **do**

    $w_t^{(i)} \leftarrow w_t^{(i)}/W_t$ // Normalization

    $(\hat{x}_t, \hat{y}_t) \leftarrow \left(\sum_{i=1}^{N_s} w_t^{(i)} x_t^{(i)}, \sum_{i=1}^{N_s} w_t^{(i)} y_t^{(i)} \right)$

    $\widehat{N}_{eff} \leftarrow (\sum_{i=1}^{N_s} (w_t^{(i)})^2)^{-1}$

  end for

end while

return $(\hat{x}_t, \hat{y}_t)$

// 2. Shadowing Estimation

Estimate the shadowing gains $\widehat{\mathbf{X}}_\mathbf{t}$ using Eq. (5.11)

5.4 Detection of Abnormal Sensor Reports via Monitoring Shadowing Correlation

In this section, we describe the shadowing-estimation component in SOLID, and discuss the attack-detection algorithm.

5.4.1 Monitoring Shadow Fading for Attack Detection

For the analysis and simulation of SOLID, we need a method to generate temporally-correlated shadow fading that accurately represents real-world shadowing environments. For this, we use the Gudmundson's empirical shadow fading model [60] to generate temporally-correlated shadowing gains between the primary transmitter and sensors.

5.4.1.1 Construction of Shadowing Profile

SOLID constructs and maintains the *profile* of normal shadow-fading behavior for each cooperative sensor n, based on the history of reports from the sensors during the primary transmitter tracking process. We define the basic *profile element* (PE) of sensor n as the shadowing component in the received primary signal strengths in Eq. (5.1), i.e.,

$$X_{t,n} = P_{t,n} - P_o - \alpha 10 \log(d_o) + \alpha 10 \log(\widehat{d}_{t,n}) - Y_{t,n}, \tag{5.7}$$

where $P_{t,n}$ is the sensor n's measurement report at sensing period t, $\widehat{d}_{t,n}$ the estimated distance between the primary transmitter and sensor n, which is obtained via SMC, and $Y_{t,n} \sim \mathcal{N}(0, \sigma_m^2)$ the noise power.

Suppose that, at time t, SOLID has processed $k(\geq 1)$ PEs for sensor n. Note that k may vary with sensors based on the time they joined the cooperative sensor set. This sequence of PEs exhibit a strong temporal correlation, because SOLID keeps track of each sensor's shadowing gain at each sensing period (e.g., once every 2 seconds). To exploit the temporal correlation in PEs, we define a profile vector consisting of the entire history of PE records:

$$\mathbf{X}_{t,n}(k;1) = [X_n(t), \ldots, X_n(t-k+1)]^T, \quad 1 \leq n \leq N. \tag{5.8}$$

Thus the estimates of the shadowing gain $X_{t,n}$ provide a compact description of the normal shadowing profile. We henceforth omit the subscript t for brevity.

5.4.1.2 Shadowing Estimation Using Kalman Filter

We now describe how SOLID accurately estimates the PE (i.e., shadowing gain) from the observed primary signal strengths. Specifically, the attack-detector in SOLID wants to find the shadow-fading estimator that minimizes the mean squared errors (MSE):

$$MSE_n(k;1) = \mathbb{E}\left\{ \sum_{\tau=t-k+1}^{t} \left| \mathbf{X}_n(\tau) - \widehat{\mathbf{X}}_n(\tau) \right|^2 \right\}, \tag{5.9}$$

where k is the index of the sensing stage since sensor n joined the set of cooperative sensors. We thus need an efficient estimator that minimizes the MSE in Eq. (5.9).

To meet this requirement, SOLID employs the Kalman filter (KF) [66], a recursive estimator that produces optimal estimates by minimizing the mean squared errors. In the KF, the system can be modeled as:

$$\mathbf{S}_n(k+1) = \mathbf{\Phi}_n(k)\,\mathbf{S}_n(k) + \mathbf{W}_n(k), \tag{5.10}$$

where $\mathbf{S}_n(k)$ represents the state (i.e., shadowing gain) of the system, $\mathbf{\Phi}_n(k)$ is the state-transition matrix that relates the state $\mathbf{S}_n(k)$ to the next state $\mathbf{S}_n(k+1)$, $\mathbf{W}_n(k) \sim \mathcal{N}(0, \mathbf{Q})$ is the system noise vector where the covariance matrix $\mathbf{Q}$ represents the degree of variability in the state variables.

The measurement of the system is defined as:

$$\mathbf{X}_n(k) = \mathbf{H}_n(k)\,\mathbf{S}_n(k) + \mathbf{V}_n(k), \tag{5.11}$$

where the matrix $\mathbf{H}_n(k)$ represents an observation model that relates the true state variable $\mathbf{S}_n(k)$ to the measurements $\mathbf{X}_n(k)$. In our model, $\mathbf{H}_n(k)$ is the channel gain between the primary transmitter and sensor n in Eq. (5.2), and it is updated in each sensing time slot based on the estimated location of the primary transmitter. The measurement noise is denoted as $\mathbf{V}_n \sim \mathcal{N}(0, \mathbf{R})$, where the covariance matrix $\mathbf{R}$ represents the measurement uncertainty. We consider the measurement noise in spectrum sensing due to noise power (i.e., $Y_{t,n}$ in Eq. (5.1)) by setting $\mathbf{R} = \sigma_m^2$, and setting $\mathbf{Q} = 0.1^2$ empirically.

5.4.2 Attack Detection and Filtering

A compromised or malfunctioning sensor node may report a falsified sensing value to the BS. Such manipulated sensing reports may render the localization less reliable, hampering an efficient reuse of spectrum opportunities in the spatial domain. To mitigate this problem, SOLID verifies the trustworthiness of sensing reports and filters out or penalizes the bad ones before performing the localization.

Algorithm 7 ATTACK-DETECTION ALGORITHM IN SOLID

For every newly joint cooperating sensor n, the BS performs

Initialization
$k \leftarrow 0$
$\mathsf{blacklist\_count}(n) \leftarrow 0$
while $n \in S_t$ **do**
  $k \leftarrow k + 1$ // Start the k^{th} iteration
  The BS estimates $X_n(k)$ using Kalman filter
  Compute $PED_n(k)$ using Eq. (5.13)
  if $PED_n(k) > \eta$ **then**
    if $++\,\mathsf{blacklist\_count}(n) \geq N_B$ **then**
      blacklist n
    end if
    if Sensor n is blacklisted **then**
      Exclude sensor n from localization
    end if
  end if
end while

SOLID activates an instance of attack-detection scheme whenever the BS employs a sensor for cooperative sensing. The attack detector in SOLID quantifies the deviation of a sensor's shadowing gain from the value predicted from its history by monitoring the prediction error, which can be computed as:

$$e_n(k) = \mathbf{X}_n(k) - \mathbf{H}_n(k)\,\widehat{\mathbf{S}}_n(k \mid k-1), \tag{5.12}$$

where $\mathbf{X}_n(k)$ is the observed shadow fading in Eq. (5.7).

We introduce a metric for attack detection in SOLID, called *prediction error distance* (PED) that indicates the Euclidean distance in two consecutive prediction errors, i.e.,

$$PED_n(k) = \big|\, e_n(k) - e_n(k-1) \,\big|. \tag{5.13}$$

This is a very useful, yet simple, metric because the prediction error is correlated under no attack, and consequently, the difference in two consecutive errors is kept small (see Fig. 5.5(a)). We also observed from our simulation results that $PED_n(k)$ is smaller than the prediction error itself.

The attack detector in SOLID raises a flag on sensor n's report as compromised (or

abnormal) if:

$$PED_n(k) \geq \eta, \tag{5.14}$$

where $\eta \in \mathbb{R}$ is a pre-defined threshold for detecting anomalies. SOLID classifies a sensor as malicious and excludes it from the localization process if the cumulative number of flags raised is greater than N_B, which is a design parameter. Algorithm 7 describes the pseudocode of the attack-detection algorithm in SOLID.

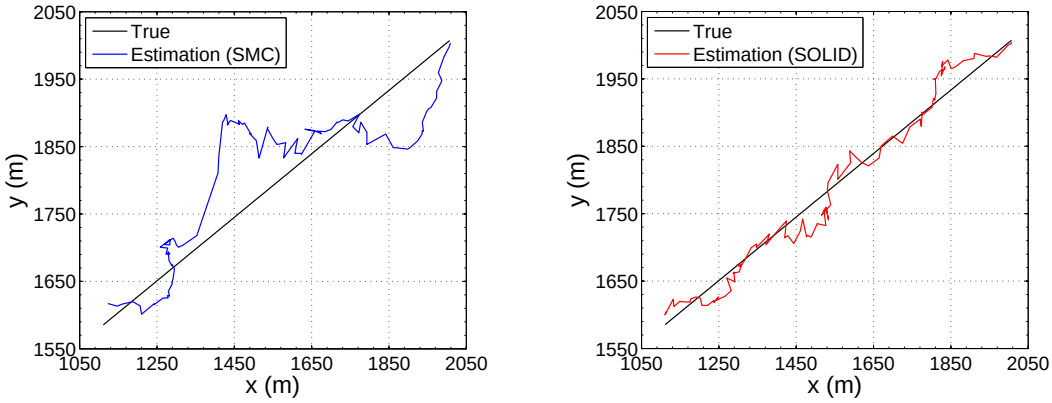
5.5 Performance Evaluation

SOLID is evaluated using MATLAB-based simulation. We first describe the simulation setup and show the efficacy of shadow-fading estimation in SOLID in the absence of attacks. We then demonstrate SOLID's robustness against various attack scenarios including *slow-poisoning* attacks, and show the tradeoff in determining the attack-detection threshold. Finally, we show SOLID's efficacy in spatial spectrum reuse.

5.5.1 Simulation Setup

We consider a CRN where sensors are randomly distributed according to a point Poisson process in a $6\,\text{km} \times 6\,\text{km}$ area with the average sensor density of $3/\text{km}^2$, unless otherwise specified. We assume a WM with a transmit-power of $250\,\text{mW}$, which is the maximum transmit-power allowed by the FCC in the UHF band [36]. For WM's mobility, we assume a Random Waypoint model without pause time [158], which is frequently used in simulations in wireless networks. We assume that the WM moves at a fixed speed of $5\,\text{m/s}$ with a destination randomly selected in the simulated network area. For each testing scenario, we ran simulations over at least 60 randomly-generated secondary network topologies to study average behavior.

For WM sensing, we fix the sensing interval at 2 seconds, and during each sensing period, sensors measure the RSS using the energy detector for $1\,\text{ms}$, as is typically assumed in 802.22 WRANs [100]. The radius of the fusion range for cooperative sensing is fixed at $R_s = 1\,\text{km}$, which is shown to be near-optimal for WM sensing in an 802.22 WRAN



(a) SMC (b) SOLID

Figure 5.3: Comparison of the tracking performance under no attack: SOLID outperforms conventional SMC-based tracking thanks to its ability to accurately estimate shadow-fading gains.

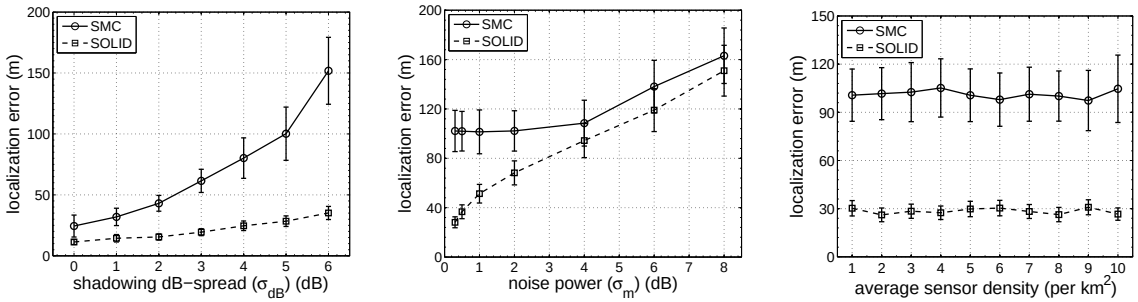
[105]. The shadow fading dB-spread σ_{dB} is assumed to be 5 dB, as it is typically assumed in IEEE 802.22 WRANs. The shadowing-decorrelation distance is set to 150 m,[5] and the path-loss exponent α is 4. We assume these parameters are estimated at the time of system deployment, and thus known *a priori* to the secondary system.

For WM tracking, we set the number of samples for SMC to $N_s = 40$ and set the re-sampling threshold N_{thr} empirically in the range $N_{thr} \in [3, 5]$, depending on the network environment. In what follows, the figures of localization error plot the average as well as $\pm 0.25\,\sigma$ interval.

5.5.2 Performance of SOLID under No Attack

SOLID's accurate localization of a primary transmitter will not only allow better spatial spectrum reuse, but will also enable high robustness against malicious/malfunctioning sensors. Here we demonstrate, in the absence of attacks, the efficacy of shadow-fading estimation in SOLID. Fig. 5.3 plots examples of mobile primary transmitter tracking during a period of 100 s. The SMC-based tracking suffers from large tracking error due to the shadow-fading-induced unpredictability in RSSs (Fig. 5.3(a)), whereas SOLID closely tracks the primary transmitter's location for the entire tracking process (Fig. 5.3(b)). In

[5] A measurement study [10] indicates that a typical decorrelation distance is in the range of 120-150 m in suburban areas.



(a) Effects of shadow fading (σ_{dB})

(b) Effects of noise power (σ_m)

(c) Effects of sensor density (ρ)

Figure 5.4: Tracking performance under no attack: SOLID (a) successfully withstands shadow fading-induced unpredictability, (b) achieves high performance gain when the measurement noise (σ_m) is small, and (c) outperforms SMC-based tracking for various sensor densities.

what follows, we demonstrate the efficacy of shadow-fading estimation under various network parameters.

5.5.2.1 Effects of Shadow Fading

Fig. 5.4(a) shows that SMC-based tracking suffers from the unpredictability in RSSs due to shadow fading, resulting in a fast increase of error as σ_{dB} increases. By contrast, SOLID maintains a small average localization error (< 35 m) for all simulated scenarios thanks to its estimation of the primary location and shadow-fading gains, which refine each other throughout the tracking process.

5.5.2.2 Effects of Noise Power

The measurement noise (including the effects of multi-path fading) in RSSs can adversely affect the accuracy of shadow-fading estimation. Fig. 5.4(b) shows that the average localization error increases with noise power (σ_m) since a large σ_m makes the shadow-fading estimation becomes less accurate. Therefore, it is crucial to combat or reduce the effect of noise power σ_m at each cooperative sensor in order to fully benefit from shadow-fading estimation in SOLID.

Although the standard deviation of Rayleigh fading, σ_m, can be as large as 5.5 dB in practice, one can use many existing techniques to significantly reduce the effect of multi-

path fading, e.g., exploiting antenna diversity [9]. For sensors with a single transceiver, this can be accomplished by extending the sensing time (longer than the channel coherence time) [135] at the expense of increased sensing overhead (e.g., time and energy). In what follows, we assume the standard deviation of the noise power is fixed at $\sigma_m = 0.3$ dB.

5.5.2.3 Effects of Sensor Density

Fig. 5.4(c) plots the localization error for various average sensor densities. The figure shows that the average localization error decreases as the sensor density increases for both schemes. However, the error drops faster with SOLID, significantly outperforming the SMC-based tracking scheme thanks to its ability to accurately track the shadow fading gains. When the average sensor density is $\rho = 3.5/\text{km}^2$, SOLID reduces the error by up to $88\,\%$ compared to the SMC-based tracking.

5.5.3 Performance of Attack Detector

To illustrate the performance of the SOLID's attack detector, we consider a simple exemplary scenario where a malicious sensor injects manipulated sensing reports at time slot 50. The malicious sensor introduces a deviation (or *attack strength*) from its actual measured RSSs by 0 (no attack), 1, 3, and 5 dB, where the deviation direction (i.e., $\pm$) is randomly chosen. Fig. 5.5 shows that the deviation injected by an attacker at the 50-th iteration increases the *prediction error distance* (PED) proportionally to the attack strength, yielding high detection accuracy. The figures show that even a small deviation (e.g., 1 dB) causes an abrupt increase in PED, and can thus be easily detected by SOLID, thanks to its ability to closely estimate/track temporally-correlated shadow fading in the measured RSSs.

5.5.4 Attack-Tolerance of SOLID

We now demonstrate SOLID's attack-tolerance while varying two key attack parameters; *attack strength* and *attack population*. We fix the attack frequency at 0.3, i.e., compromised sensors launch attacks independently with probability 0.3 in each sensing stage. We

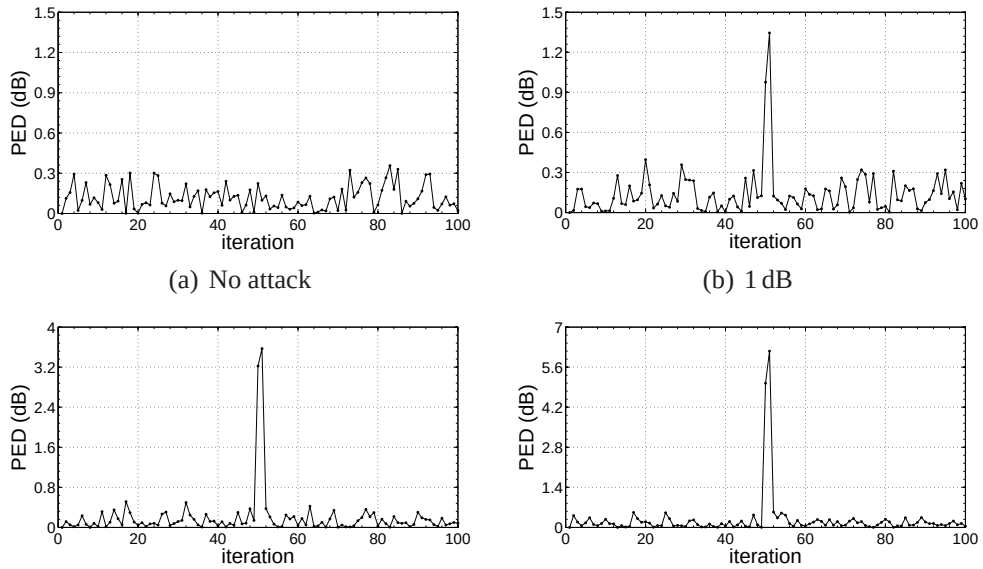


Figure 5.5: The attack-detection capability of SOLID: SOLID can accurately detect even a small deviation in sensor reports (i.e., RSSs) since such a deviation boosts the prediction error distance (PED), which makes it easy for SOLID to detect any abnormal sensing reports.

set the detection and blocking thresholds to $\eta = 5$ dB and $N_B = 2$, respectively.

To demonstrate the efficacy of SOLID, we compare the following three testing schemes: (i) *SMC-based tracking*, (ii) SOLID *without attack detector*, and (iii) SOLID *with attack detector*.

5.5.4.1 Impact of Attack Strength

Here we show the impact of attack strength on the localization accuracy, while varying the attack strengths in the range between 0 and 10 dB. We assume that the attack population is $30\,\%$, i.e., each sensor is compromised with probability 0.3.

Fig. 5.6 shows that the localization performance of SMC-based tracking suffers from large attack strengths due to the lack of ability to detect and filter out manipulated sensing reports. For a similar reason, the localization error of SOLID *without attack detector* also increases with increasing attack strengths. However, this scheme significantly lowers the average error compared to the SMC-based tracking, thanks to its ability to accurately track the shadowing gains.

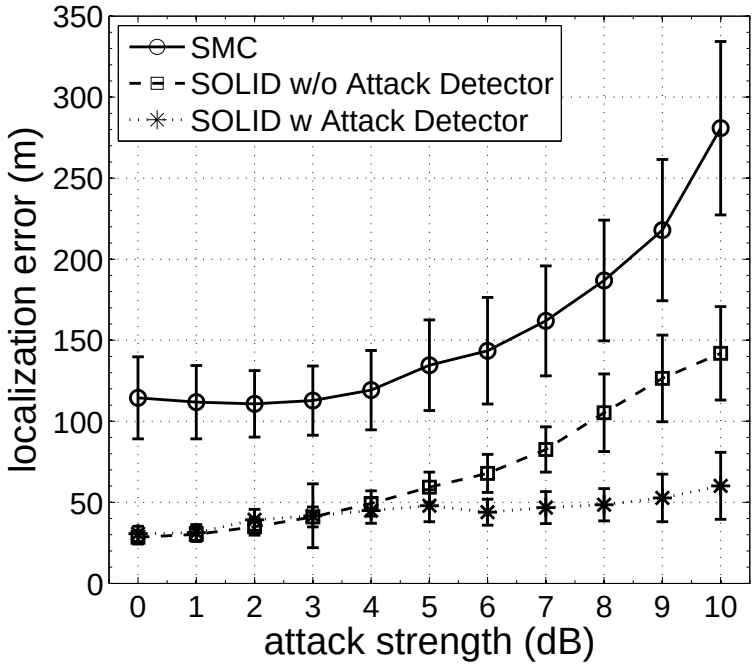


Figure 5.6: Attack-tolerance of SOLID: SOLID successfully tolerates attacks thanks to its ability to exploit temporal shadowing correlation to accurately detect abnormal sensing reports.

In contrast, SOLID *with attack detector* maintains a low localization error even in the case of large attack strengths. This performance superiority can be explained as follows. On one hand, the attack detector in SOLID successfully withstands weak attacks, i.e., $<$ $\eta = 5$ dB, because such attacks do not influence the localization outcome much even though they can evade the attack detector. On the other hand, the attack detector can easily detect strong attacks, i.e., $> \eta = 5$ dB, thanks to its ability to detect large deviations in shadowing estimation caused by manipulated sensing reports.

However, Fig. 5.6 shows that the localization error of SOLID *with attack detector* still increases slowly with increasing attack strength for the following two reasons. First, the detection delay (i.e., N_B) allows an attacker to influence the localization outcome. Second, the localization error induced by the attackers increases the attack false-alarm rate, i.e., misclassifying legitimate sensors as malicious/faulty, thus increasing the fraction of attackers in the set of cooperative sensors.

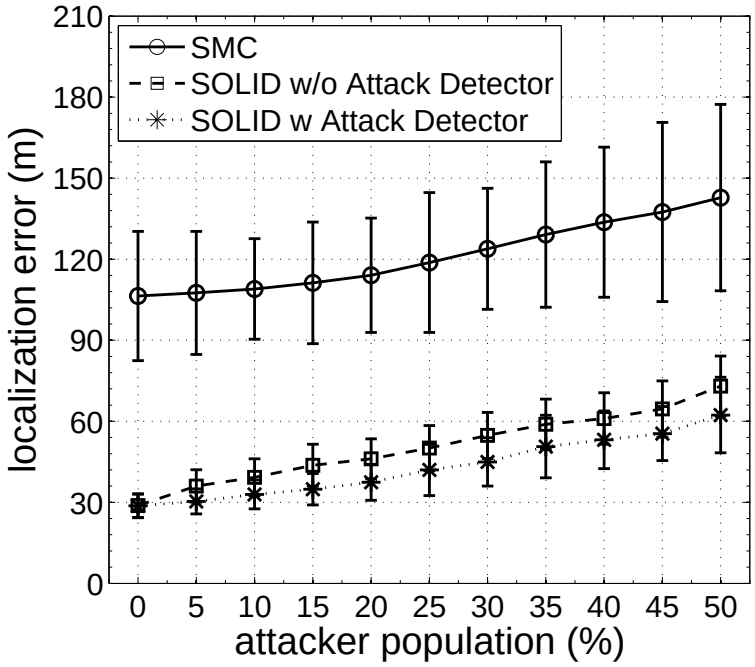


Figure 5.7: Impact of attacker population: The localization accuracy of SOLID depends on the design of attack detection threshold η, making tradeoff between under- and over-filtering.

5.5.4.2 Impact of Attacker Population

Next, we examine the impact of the attacker population by varying the fraction of compromised sensors from $0\,\%$ to $50\,\%$. We fix the attack strength at 5 dB. As expected, Fig. 5.7 shows that a larger attacker population degrades localization performance because it is harder to identify compromised sensors. Moreover, a large fraction of compromised sensors will remove a large number of sensors from the cooperating group, which, in turn, negatively affects the localization performance. Nevertheless, the localization error is significantly lowered by SOLID *with attack detector* compared to the conventional SMC-based tracking scheme even with a large fraction of compromised sensors, demonstrating its robustness against attacks.

5.5.5 Tolerance against "Slow-Poisoning" Attack

To further demonstrate SOLID's high attack-tolerance, we evaluate SOLID's tracking performance under a challenging, *slow-poisoning* attack, such that malicious sensors incre-

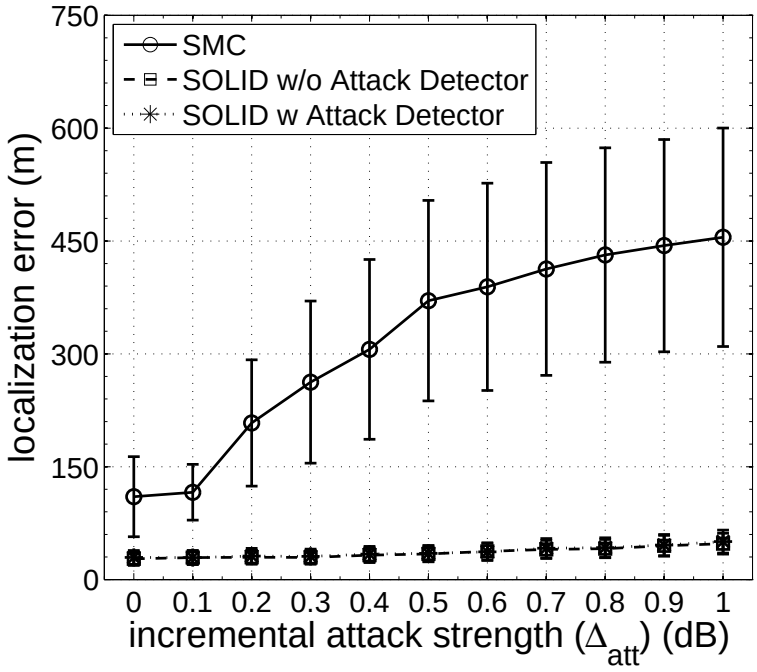


Figure 5.8: Attack-tolerance of SOLID against slow-poisoning attack: SOLID successfully tolerates slow-poisoning attacks, successfully safeguarding the tracking process.

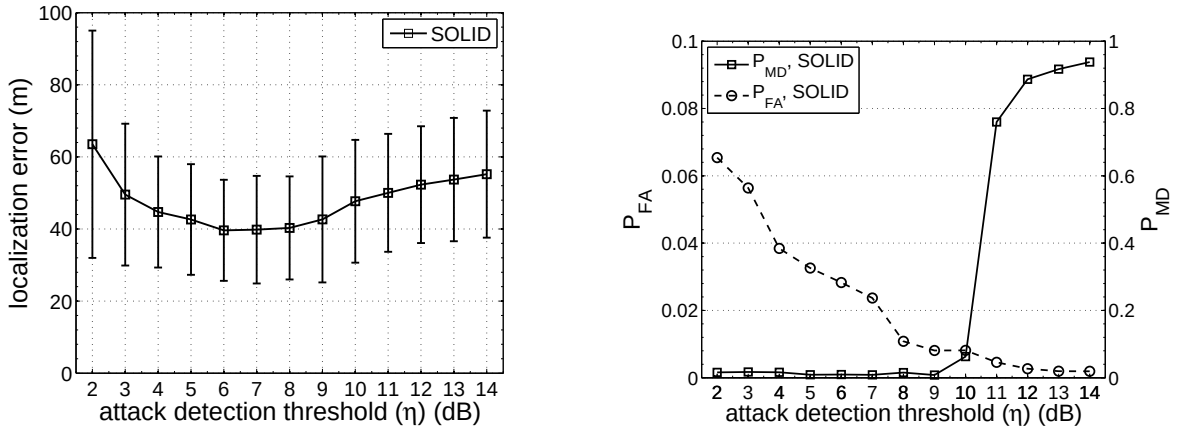
mentally raise the attack strength by Δ_{att} (dB) in order to evade detection, while disrupting the localization process. Specifically, we assume that a malicious sensor reports the falsified value $P^a_{t,n}(k)$ in the k^{th} sensing stage after joining the set of cooperative sensors, i.e., $P^a_{t,n}(k) = P_{t,n} + k \cdot \Delta_{att}$.

Fig. 5.8 shows that SOLID performs well under a slow-poisoning attack, even without the attack detector, while the performance of the SMC-based tracking suffers greatly from the attack. Thus, the figure demonstrates that SOLID efficiently mitigates the effects of a slow-poisoning attack.

5.5.6 Tradeoff in Determining the Attack Detection Threshold

We now study the impact of detection threshold η. In our simulation, we fixed the attack strength at 5 dB, and measure the localization accuracy and attack detection performance (in terms of false-alarm and mis-detection probabilities), while varying the detection threshold in the range $\eta \in [2, 14]$ dB.

Fig. 5.9(a) indicates that the localization performance of SOLID suffers in the case



(a) Localization error (b) Attack detection performance

Figure 5.9: Impact of attack detection threshold: The attack detection threshold η affects (a) the localization accuracy, as well as (b) false-alarm and mis-detection probabilities. In simulations, the attack strength is fixed at 5 dB.

of low detection thresholds, i.e., $\eta < 6$ dB, due mainly to *over-filtering*, i.e., some of the well-behaving sensors are flagged as malicious and then their reports are discarded. On the other hand, too high a detection threshold, i.e., $\eta > 6$ dB, also degrades localization performance because of *under-filtering*, where some of the attackers evade detection, thus adversely influencing the localization process.

Fig. 5.9(b) clearly shows the tradeoff in determining the attack-detection threshold η in terms of false-alarm (denoted by P_{FA}) and mis-detection (denoted by P_{MD}) probabilities. `SOLID` is shown to achieve near-zero P_{MD} and to maintain a low false-alarm rate, i.e., $P_{FA} < 6\,\%$, unless the detection threshold is significantly larger than the attack strength, i.e., $\eta > 10$ dB.

Therefore, the attack detection threshold must be carefully chosen to make the tradeoff between false-alarm and mis-detection probabilities, while considering their dependency on attack strengths and `SOLID`'s tolerance to weak attacks, as observed in Fig. 5.6.

5.5.7 Improvement in Spatial Spectrum Reuse

The SUs located within a keep-out-radius of R_e from a small-scale PU (e.g., a WM) must vacate the channel to avoid excessive interference to primary communications [34]. The keep-out-radius needs to be enlarged when the localization is inaccurate, thus reducing

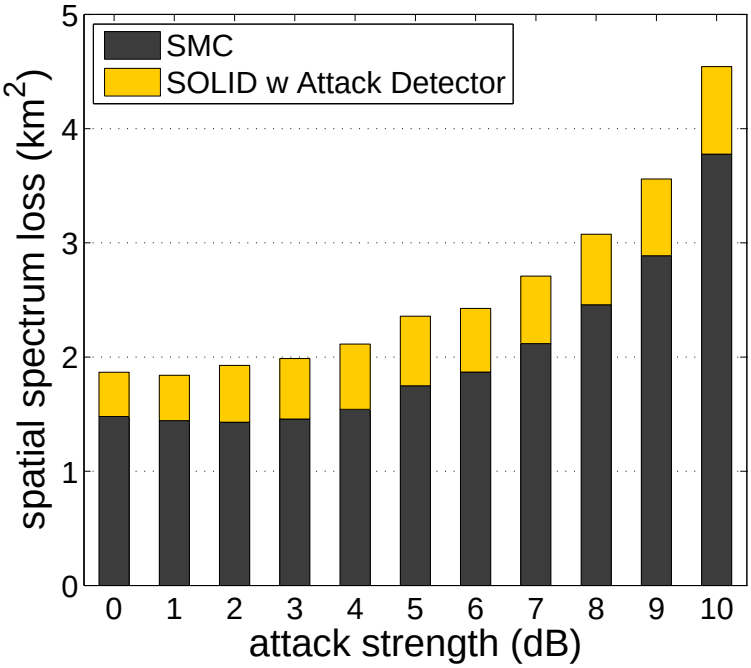


Figure 5.10: Spatial spectrum opportunity loss (SSOL) due to localization error: SOLID significantly reduces the spatial spectrum loss thanks to its ability to accurately track the location of the mobile primary transmitter.

spatial spectrum utilization. SOLID achieves high spatial spectrum efficiency by providing accurate location of mobile primary transmitters. We quantify the improvement in spectrum efficiency made by SOLID by introducing the metric of *spatial spectrum opportunity loss* (SSOL), which is defined as the extended area for PU protection due to the inaccuracy of PUs' localization. Assuming a localization error of ϵ, the spatial spectrum opportunity loss due to the inaccuracy of the tracking process can be roughly approximated as $SSOL \approx \pi(R_e+\epsilon)^2 - \pi R_e^2 = \pi\epsilon^2 + 2\pi R_e\epsilon$.

Fig. 5.10 compares the spatial spectrum opportunity loss of the SMC-based tracking and SOLID, assuming the keep-out-radius of $R_e = 2\,\text{km}$, which is reasonably sufficient to give a typical WM transmission range of 100-$150\,\text{m}$. The figure clearly indicates that SOLID maintains small $SSOL$, improving spatial spectrum efficiency substantially. Note that the improved spectrum efficiency can be translated to other performance metrics, such as bandwidth of SUs.

5.6 Related Work

Spatial Spectrum Reuse in CRNs: Reusing *spatial* spectrum opportunities is important for efficient utilization of spectrum resources, but has received little attention. The IEEE 802.22 Working Group (WG) proposed a coexistence model with a wireless microphone (WM) to maximize spatial spectrum reuse [34], assuming that the location of WMs is available to secondary systems. This, however, may not be valid for mobile WMs where the geo-location database is not available. Recently, we proposed a small-scale PU detection/localization scheme, called `DeLOC` [104]. While `DeLOC` provides an efficient mechanism for initial detection of PUs, this chapter focuses on attack-tolerant tracking of mobile PUs.

Secure Spectrum Sensing in CRNs: Chen *et al.* [29] proposed an RSS-based location verification scheme, called *LocDef*, to detect a fake primary signal. Liu *et al.* [93] developed a primary signal verification scheme by jointly exploiting the location-dependent link signature, i.e., multi-path fading profile, and conventional cryptographic authentication methods. The problem of ensuring the robustness in distributed sensing has also been studied [28, 79, 103]. Min and Shin [103] proposed an attack-tolerant secure cooperative sensing scheme that exploits shadow-fading correlation in RSS among close-by sensors. Unlike these, we focus on a new type of attack, i.e., disruption of location tracking of a mobile primary transmitter by falsifying sensor reports.

Secure Mobile Target Tracking: The problem of node localization and target tracking has been studied extensively in the area of wireless sensor networks [15, 42, 61, 136, 162]. The primary tracking in CRNs, however, faces unique challenges. In CRNs, it is not desirable to modify the primary system, and thus, the received primary signal strengths information obtained via spectrum sensing is only available to the secondary system. The solution approach taken by `SOLID` to overcome this challenge differs from others in that it only relies on the PHY-layer signal-propagation characteristics (i.e., temporally-correlated shadow fading) to accurately detect malicious sensors, which has not been considered before.

5.7 Conclusion

In this chapter, we have introduced SOLID, which enables accurate and robust location tracking of small-scale mobile PUs in CRNs. By jointly performing localization and shadow-fading estimation, SOLID significantly improves the accuracy of mobile PU tracking and masks the effect of manipulated sensing reports by accurately detecting and filtering out manipulated sensing reports. Our in-depth evaluation results, in realistic wireless environments, show that SOLID reduces localization error significantly both in the absence/presence of attacks, including the "slow-poisoning" attack. The enhanced primary tracking capability offered by SOLID enables the secondary system to make a great improvement in overall spectrum efficiency.

In the future, we would like to study scenarios in which multiple attackers collude to disrupt the tracking process. It would also be interesting to devise an optimal attack strategy that can maximally influence the primary tracking performance. We also plan to design secure primary tracking in ad-hoc CRNs.

CHAPTER 6

Opportunistic Spectrum Access for Mobile Cognitive Radios

6.1 Introduction

In this chapter, we study the problem of enabling DSA for *mobile* CR devices by identifying and addressing three fundamental challenges. First, existing spectrum-availability models are derived based solely on PUs' temporal traffic statistics and might thus be unsuitable for CRNs with mobile CRs/SUs. Unlike in stationary CRNs (e.g., [1]), in which spectrum opportunity (or availability) is mostly affected by PUs' temporal channel usage patterns, in mobile CRNs, availability can also change as SUs move towards or away from PUs that are actively transmitting data. To overcome this limitation, we model channel availability—that reflects the fluctuation of spectrum opportunities induced by SU mobility—as a two-state continuous-time Markov chain (CTMC) and verify its accuracy via in-depth simulation.

Second, protecting PUs from SU mobility-induced interference is a challenging problem that calls for an efficient spectrum-sensing strategy tailored to mobile CRNs. Mobile SUs may need to sense spectrum more frequently to avoid interfering with PU communications. However, frequent spectrum sensing may not only incur significant time overhead [88], but also quickly drain the battery of mobile CR devices due to the power-intensive nature of spectrum sensing [7, 68]. To address this challenge, we propose the use of *guard*

distance to minimize the required spectrum sensing for mobile SUs, while providing sufficient protection to primary communications. Guard distance is an additional separation between PUs and SUs to prevent mobile SUs from causing excessive interference. Further, based on our proposed channel-availability model, we jointly optimize the guard distance and spectrum-sensing interval to maximize the reuse of spectrum opportunities in the space and time domains.

Third, mobile SUs experience heterogeneous spectrum opportunities across the space and time domains based on the geographical distribution of PUs and SUs' mobility patterns. To better utilize such heterogeneous spectrum opportunities, we derive an optimal, distributed channel-access strategy in a closed form within the convex optimization framework. Our channel-access strategy incorporates three key factors that diversify spectrum access opportunities across different channels: (i) SU-mobility-aware spectrum sensing adaptation, (ii) heterogeneity in PUs' spatial distributions and channel-usage patterns, and (iii) spectrum sharing among SUs. Our proposed channel-access strategy is shown to significantly improve secondary network throughput, fairness and energy-efficiency in spectrum sensing.

The three challenges mentioned above are interrelated. Hence, to fully realize the benefits of DSA for mobile SUs, they must be considered jointly. To the best of our knowledge, our work is the first to extensively investigate SU mobility in regard to the channel-availability model, spectrum sensing and access strategies.

6.1.1 Contributions

In summary, this chapter makes the following main contributions.

- *Introduction of a novel spectrum-availability model*: We show via analysis that the channel availability experienced by a mobile SU can be accurately modeled as a two-state Markov model under reasonable assumptions. We further verify the accuracy of this model via in-depth simulations.

- *Design of mechanisms for protection of primary communications*: We identify *guard distance* as a key enabler of efficient spectrum reuse while protecting PU communica-

tions. We jointly consider guard distance and spectrum sensing interval to maximize reusable spatio-temporal spectrum opportunities.

- *Optimal distributed channel-selection strategy*: Based on our channel availability model, we derive an optimal, distributed channel-selection strategy that maximizes secondary network throughput. We also show how the optimal strategy is influenced by parameters, such as SU density and speed.

6.1.2 Organization

The remainder of this chapter is organized as follows. Section 6.2 introduces the system models that will be used throughout this chapter. Section 6.3 presents our new channel-availability model for mobile SUs. Sections 6.4 and 6.5 detail the design of spectrum sensing and access schemes that maximize secondary network throughput. Section 6.6 evaluates the performance of the proposed schemes, and Section 6.8 concludes the chapter.

6.2 System Model

In this section, we present a mobile CRN model, along with distributed spectrum sensing and channel-access models.

6.2.1 Mobile CRN Model

We consider a CRN with infrastructure-based *fixed* primary networks and *mobile* ad-hoc secondary networks in the same geographical area, as shown in Fig. 6.1. We assume that each cell of the primary system consists of a single central node (e.g., access point) and receivers. From now on, we refer to each primary cell as a PU. We assume that there is a non-empty set $\mathcal{K}$ of licensed channels, and that PUs operating on the same channel belong to the same type of system and have the same *temporal* channel-usage statistics, e.g., channel `busy`/`idle` durations.[1] Primary transmitters are assumed to be distributed,

[1]We use the terms `busy`/`idle` to indicate PUs' temporal traffic patterns, and use ON/OFF to indicate the availability of a channel seen from a mobile SU's perspective.

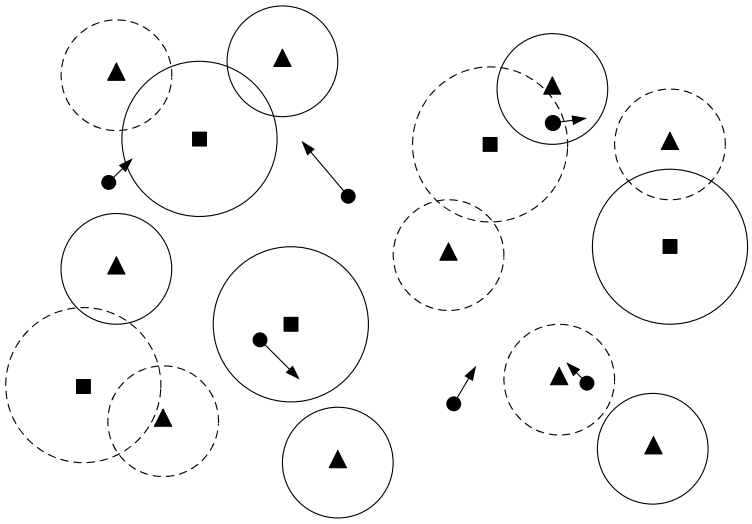


Figure 6.1: Illustration of a mobile CRN: Mobile CR devices (solid dots with arrow) can opportunistically use the licensed channels only when the distance from any active PUs (triangles and rectangles) is greater than a certain threshold (i.e., protection region) so as to avoid excessive interference to PUs. The circles with solid (dotted) lines indicate the protection region of active (inactive) PUs with (without) data transmission.

following a point Poisson process, with a different average density for each channel, i.e., $n_{p,i} \sim Poisson(k; \rho_{p,i})$, where $n_{p,i}$ is the number of primary transmitters and $\rho_{p,i}$ is the average PU density on channel $i \in \mathcal{K}$. We assume that primary transmitters on the licensed channel $i \in \mathcal{K}$ are separated by at least twice their transmission range in order to avoid interference [156]. Such a PU distribution can be obtained by eliminating overlapping PUs in the original Poisson process, resulting in a *Marten Hardcore Process* [140]. We assume that SUs know the average density of PUs on each channel, and PUs' temporal channel-usage characteristics. We further assume that SUs do not know the availability of a channel at a specific time and location unless they perform spectrum sensing.

6.2.2 Distributed Spectrum Sensing & Access Models

We assume that SUs are *mobile* devices with CR-functionality that allows them to access any licensed channels in the set $\mathcal{K}$. However, they do not have the capability of accessing a geo-location spectrum database to obtain local spectrum-availability information.[2]

[2]The FCC specifies two types (Mode I and II) of portable devices that can access TV white space [49]. Mode I devices are required to access the geo-location database, whereas Mode II devices are not required to

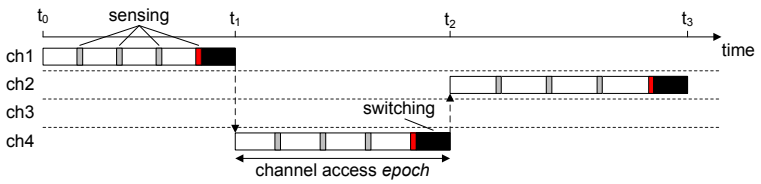


Figure 6.2: Opportunistic channel access model: An SU periodically senses its current operating (in-band) channel (the gray block) until it detects a primary signal, followed by channel switching (the black block). The sensing interval is dynamically adapted based on the SU's speed and PUs' spatio-temporal channel usage statistics.

Therefore, we assume that SUs rely on local spectrum sensing (e.g., feature detection) to detect channel availability—i.e., the presence/absence of primary signals—at a given time and location. SUs are assumed to use feature detection (e.g., [51]) for PHY-layer sensing. Feature detection is known to provide high accuracy without collaboration amongst SUs even at a low SNR [142]. Thus, it is better suited for ad-hoc secondary networks, in which SU collaboration may not be feasible due to the needs for information exchange and global time synchronization [112].

Once an SU identifies available channels via spectrum sensing, it contends with neighboring SUs to access the channel via a random access scheme such as CSMA. SU channel access behavior is depicted in Fig. 6.2. We assume that SUs always have packets to transmit and always use the maximum transmission power allowed by a regulatory body.

6.3 Modeling Channel Availability for Mobile Secondary Users

In this section, we characterize the spectrum opportunity that corresponds to PUs' spatio-temporal channel usage patterns, propose a new SU mobility-aware channel availability model, and demonstrate its accuracy via simulation.

have such access capability.

6.3.1 Characterizing Spatio-Temporal Spectrum Opportunity

We first introduce the keep-out-radius and guard distance for protecting PUs from increased interference caused by SU mobility. We then quantify the spatio-temporal spectrum opportunities available to mobile SUs.

Definition 6.1 *(Keep-out radius) The keep-out radius is defined as the minimum distance between a primary transmitter and SUs under the* interference temperature limit *(*ITL*) set by the regulatory body (e.g., the FCC), i.e.,*

$$R_{e,i} = \inf\left\{ d \in \mathbb{R} \,\middle|\, I_{tot}(\rho_{s,i}, d) \leq \texttt{ITL} \right\}, \tag{6.1}$$

where $I_{tot}(\rho_{s,i}, d)$ is the average interference generated by SUs (separated by least distance d from the primary transmitter) at a primary receiver located at the edge of the primary coverage area and $\rho_{s,i}$ is the density of SUs on channel i.

The aggregate SU interference at a primary receiver located at the edge of the primary transmission range (i.e., at distance R_o from the primary transmitter) can be bounded as [145]:

$$I_i^U(\rho_{s,i}, R_{e,i}) = \frac{2\pi P_o d_o^{\alpha} \rho_{s,i}}{\alpha - 2} \big(R_{e,i} - R_o\big)^{2-\alpha}, \tag{6.2}$$

where P_o is the transmission power of SUs, d_o the short reference distance (e.g., 5 m), α the path-loss exponent, $\rho_{s,i}$ the average SU density on channel i, R_o the PUs' transmission range, and $R_{e,i}$ the primary keep-out radius.

From Eq. (6.2), the keep-out radius necessary for channel i to meet the interference constraint, $I_i^U \leq \texttt{ITL}$, is given as:

$$R_{e,i}^*(\rho_{s,i}) = \left[\left(\frac{(\alpha - 2)}{2\pi P_o d_o^{\alpha} \rho_{s,i}} \cdot \texttt{ITL} \right)^{\frac{1}{2-\alpha}} \right]^+ + R_o, \tag{6.3}$$

where $[\bullet]^+ \triangleq \max\{\bullet, 0\}$.

One important observation from Eq. (6.3) is that the keep-out radius of channel i increases with the density of channel-i SUs, $\rho_{s,i}$, as shown in Fig. 6.3(a). This is because

as SU density increases (i.e., more SUs access channel i), the keep-out radius must be expanded to meet the interference constraint.

The keep-out radius in Eq. (6.3), however, assumes *stationary* SUs, and thus, it may not be sufficient to protect PUs from interference caused by mobile SUs. To protect PUs further from such SU mobility-induced interference, we introduce an additional protection layer (*guard distance*), denoted by ϵ_i.

Definition 6.2 *(Primary protection region) Let $\mathcal{P}_i$ denote a set of primary transmitters on channel i. A* primary protection region *(PPR) of primary transmitter $j \in \mathcal{P}_i$, denoted as $\Omega_{i,j}$, is defined as a unit disk centered at the primary transmitter j located at $(x_{i,j}, y_{i,j})$, i.e.,*

$$\Omega_{i,j} = \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; \|(x_{i,j}, y_{i,j}) - (x,y)\| \leq R_{e,i} + \epsilon_i \Big\}, \tag{6.4}$$

where $R_{e,i}$ is the keep-out radius, *and ϵ_i is the guard distance.*

Thus, if an SU is located within a PPR of active PUs on channel i, it refrains from using the same channel to avoid causing interference.

Then, the average fraction of the union of PPRs on channel i in the entire network is [89]:

$$\chi_i(\rho_{s,i}) = 1 - e^{-\rho_{p,i}\pi(R_{e,i}(\rho_{s,i}) + \epsilon_i)^2}, \tag{6.5}$$

where $\rho_{s,i}$ is the average SU density on channel i.

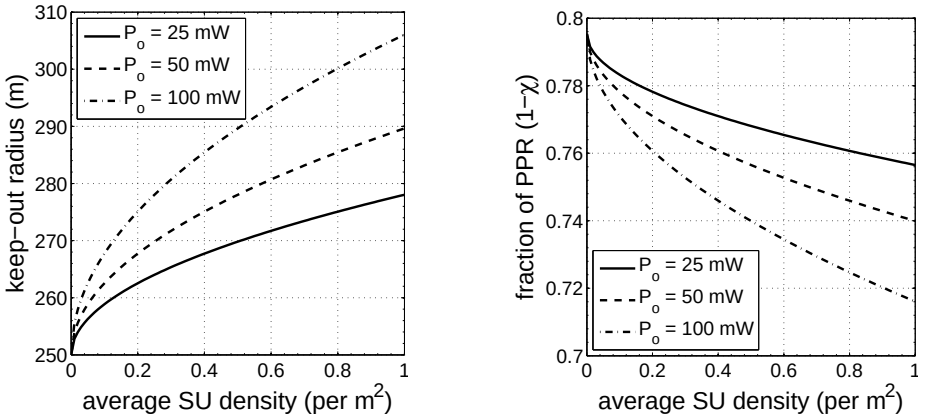
The average fraction of areas where the channel is available at any given time can be approximated as:

$$\gamma_i \approx (1 - \chi_i) + \chi_i\, \varpi_{idle,i} = 1 - \chi_i\, \varpi_{busy,i}, \tag{6.6}$$

where $\varpi_{idle,i} = 1 - \varpi_{busy,i}$ is the steady-state probability that a PU on channel i is in `idle` state, i.e., not transmitting data.

6.3.2 Assumptions for Modeling Channel Availability

To model channel availability from a mobile SU perspective, we make the following three main assumptions:



(a) Keep-out radius (b) Fraction of available area

Figure 6.3: Impact of SU density on spatial spectrum opportunity: The keep-out radius for primary protection (a) increases with increasing SU density, and thus (b) spatial spectrum opportunity decreases. The simulation parameters are set to $R_o = 250\,\text{m}$, $\texttt{ITL} = 0.1\,\text{mW}$, $\rho_p = 1/\text{km}^2$, and $\alpha = 4$.

A1.) PUs' traffic statistics, i.e., `busy`/`idle` periods follow exponential distributions.

A2. The time interval that an SU moves inside a PPR follows exponential distributions.

A3.) The time during which an SU is located within a PPR follows exponential distributions.

Regarding **A1**, the exponential distribution is the most widely used for modeling PU traffic patterns in CRNs. A recent measurement study [149] indicates that the PU channel-usage pattern can indeed be accurately approximated as an exponential distribution unless the average `busy`/`idle` periods are very long.[3]

Regarding **A2**, let T_{hit} denote the first (hitting) time that a mobile SU n moves into an active PU's PPR (i.e., in `busy` state). Then, the analysis of T_{hit} is analogous to the hitting time of a stationary object in wireless sensor networks, which can be considered as a PU in a mobile CRN. By borrowing the analysis in [89], T_{hit} can be approximated as [89]:

$$T_{hit,n} \sim Exp(2(R_{e,i} + \epsilon_i)\bar{v}_n \rho_{p,i} \varpi_{busy,i}), \tag{6.7}$$

where $\bar{v}_n$ is the average speed of SU n.

[3]For such channels with long `busy`/`idle` periods, a long-tail distribution, such as log-normal distribution, is more suitable.

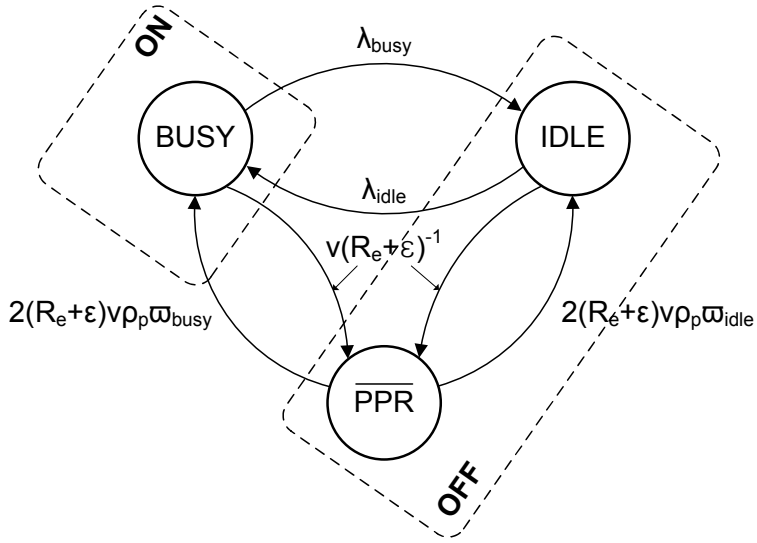


Figure 6.4: Mobility-aware channel availability model as a continuous-time Markov chain (CTMC): A channel is available for a (mobile) SU either when (i) the SU located outside the PPRs (denoted as $\overline{\text{PPR}}$) or (ii) the primary transmitter of the PPR that the SU belongs to is in `idle` state.

Regarding **A3**, the time duration in which an SU stays within a PPR can be derived from the link-lifetime distribution analysis in mobile ad-hoc networks [148]. According to [148], the link lifetime, i.e., the time duration during which the transmitter-receiver pair are located closer than a transmission range, can be accurately approximated as an exponential distribution with intensity, $\frac{\bar{v}}{R}$, where $\bar{v}$ is the average relative speed of the transceiver and R is the transmission range.

6.3.3 Mobility-Aware Channel Availability Model

We now opt to design a mobility-aware channel availability model for mobile CRNs. For this, we first define three states–i.e., `busy`, `idle`, and $\overline{\text{PPR}}$–based on the SU's location relative to the PPRs and PUs' traffic patterns, as shown in Fig. 6.4. We assume that channel i is available (i.e., OFF state) when a mobile SU is located outside the PPR of any *active* primary transmitters on channel i (i.e., `idle` or $\overline{\text{PPR}}$); otherwise, the channel is not available (i.e., ON state). We can thus reduce the Markov chain into a two-state model by merging the states `idle` and $\overline{\text{PPR}}$ into an OFF state, as shown in Fig. 6.4.

The ON/OFF state transitions occur in the following cases.

- **ON→OFF**: An SU moves out of the protection region of an active PU or a PU stops transmitting data.
- **OFF→ON**: An SU moves into the protection region of an active PU or a PU starts transmitting data.

We now derive the distributions of ON and OFF durations based on the Markov model in Fig. 6.4.

6.3.3.1 Distribution of "ON" Period

The sojourn time of the ON state of channel i follows an exponential distribution [148]:

$$T_{on,i} \sim Exp\Big(\lambda_{busy,i} + \frac{\bar{v}_n}{R_{e,i} + \epsilon_i}\Big), \tag{6.8}$$

where $\lambda_{busy,i}$ is the rate at which a PU resumes data transmission, $\bar{v}_n$ the average speed of an SU,[4] and $R_{e,i}$ and ϵ_i are the keep-out radius and the guard distance on channel i, respectively.

6.3.3.2 Distribution of "OFF" Period

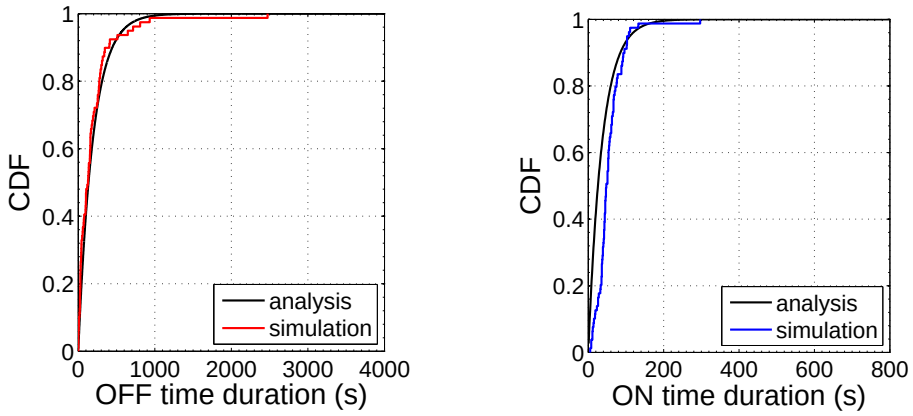
The OFF period duration can be thought of as the hitting time of the `busy` state, having either `idle` or $\overline{\text{PPR}}$ as an initial state. The OFF→ON state transition rate, λ_{off}, can be derived using the detailed balance equation, i.e., $\varpi_{on,i}\lambda_{on,i} = \varpi_{off,i}\lambda_{off,i}$, based on the stationary distributions of ON/OFF states, which can be approximated from Eq. (6.6), i.e., $\varpi_{on,i} = 1 - \gamma_i$ and $\varpi_{off,i} = \gamma_i$, and the ON→OFF transition rate $\lambda_{on,i}$ in Eq. (6.8), i.e.,

$$\lambda_{off,i} = \frac{\chi_i \varpi_{busy,i}}{1 - \chi_i \varpi_{busy,i}} \Big(\lambda_{busy,i} + \frac{\bar{v}_n}{R_{e,i} + \epsilon_i}\Big), \tag{6.9}$$

and thus, the sojourn time of the OFF state is given as:

$$T_{off,i} \sim Exp(\lambda_{off,i}). \tag{6.10}$$

[4] Although the speed of an SU can vary depending on its movement pattern, we consider average speed in the analysis for mathematical tractability.



(a) OFF period (b) ON period

Figure 6.5: Comparison of channel ON/OFF duration distributions: Our analyses on channel ON/OFF durations closely match the simulation results, thus corroborating the validity of the proposed model. In the simulation, we use the Random Waypoint model with no pause time where an SU uniformly chooses its speed in $[1, 10]$ m/s and destination with a fixed interval of 60 seconds. The average PU and SU densities are set to 2 and 10 (per km^2), respectively. We set $\varpi_{idle,i} = 0.4$ and $\lambda_{idle,i} = 0.01\,\forall i \in \mathcal{K}$.

The above analysis for channel modeling will be used for designing efficient spectrum sensing scheduling and distributed access strategy in Sections 6.4 and 6.5.

6.3.3.3 Model Verification

To show the accuracy of the proposed channel-availability model, we measure the channel ON/OFF periods observed from a mobile SU via simulation for 2×10^4 seconds. Fig. 6.5 shows that the empirical results closely match the analytical results, indicating the accuracy of the proposed model. To further quantify the accuracy, we measure the similarity between the empirical c.d.f. and the analytical c.d.f. using *Kullback-Leibler Divergence* (KLD) [83]. The KLD for two exponential distributions with intensities μ_o and μ_1 can be calculated as:

$$\mathcal{D}_{KL}(\mu_o \| \mu_1) = \log(\mu_o) - \log(\mu_1) + \frac{\mu_1}{\mu_o} - 1. \tag{6.11}$$

Table 6.1 summarizes the average and standard deviation of KLD for the ON/OFF durations while varying the maximum speed of SUs in the range of $[2, 10]$ m/s. It shows that the KLD remains low for all simulated scenarios. In fact, the case where $v_{max} = 10$ m/s

Table 6.1: Kullback-Leibler divergence for channel model

v_{max} (m/s)	$\mathcal{D}_{KL,OFF}$		$\mathcal{D}_{KL,ON}$	
	mean	std	mean	std
2	0.0441	0.0513	0.0069	0.0028
4	0.0413	0.0456	0.0202	0.0269
6	0.0301	0.0410	0.0848	0.0511
8	0.0875	0.0485	0.0982	0.0415
10	0.2335	0.0942	0.3134	0.1605

corresponds to the case in Fig. 6.5.

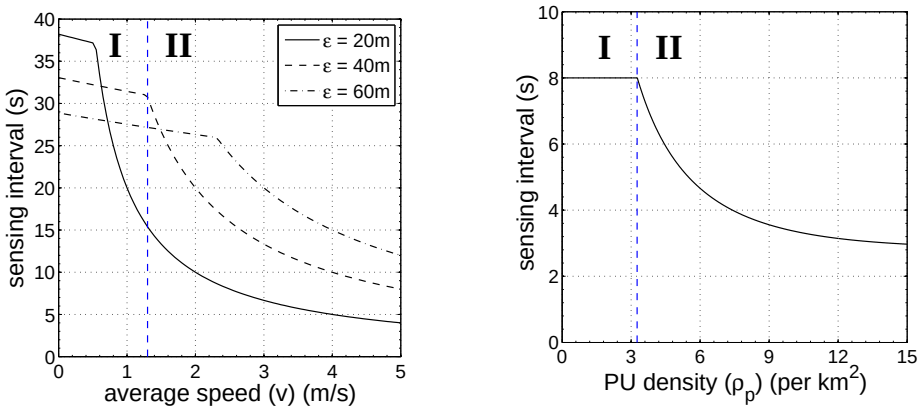
6.4 Primary Protection via Joint Optimization of Spectrum Sensing Interval and Guard Distance

In this section, we jointly design the sensing interval and guard distance to protect PU communications from mobile SUs. We first derive the minimum spectrum sensing interval for mobile SUs, and then the optimal guard distance that maximizes spatio-temporal spectrum opportunities.

6.4.1 Mobility-Aware Spectrum Sensing

In order to avoid causing excessive interference to primary communications, SUs must perform spectrum sensing frequently enough to detect a primary signal before they move into the PPR of active PUs. We assume that SUs can perfectly detect the presence of a primary signal via spectrum sensing when they are located within the PPR of any active PU. In practice, SUs may need to adjust sensing parameters to identify their locations relative to the PPRs, but this is not within the scope of this chapter.

There are two conditions under which an SU performs spectrum sensing: (i) when the c.d.f. of the channel OFF state at a given time exceeds a predefined threshold, ξ ($0<\xi<1$), to detect the returning PUs, or (ii) when an SU travels a certain distance since the previous sensing time, to prevent an SU from moving into the keep-out radius, whichever comes



(a) Impact of SU mobility

(b) Impact of PU density

Figure 6.6: Minimum sensing interval: Sensing interval depends on (a) the SUs' average speed, $\bar{v}$, and (b) the average PU density, ρ_p. In our simulation, we set the parameters $\xi = 0.3$, $\epsilon = 40\,\text{m}$, $\rho_s = 10/\text{km}^2$, $\rho_p = 1/\text{km}^2$, $R_o = 250\,\text{m}$, $\bar{v} = 5\,\text{m/s}$, $\lambda_{idle,i} = 0.1$, and $\varpi_{idle,i} = 0.4\,\forall i \in \mathcal{K}$.

first.

Then, the minimum sensing interval required on channel i is given as:

$$t_i^* = \max\left\{T_{s,i}, \min\left\{-\frac{\ln(1-\xi)}{\lambda_{off}}, \frac{\epsilon_i}{\bar{v}}\right\}\right\}, \tag{6.12}$$

where λ_{off} is the intensity of the channel OFF period distribution in Eq. (6.9), ϵ_i the guard distance, and $\bar{v}$ the average speed of an SU. Note that a lower probability ξ will lead SUs to sense the channel more frequently.

Eq. (6.12) indicates that the minimum sensing interval depends not only on *temporal* features such as primary traffic statistics, but also on *spatial* features such as the SUs' average speed $\bar{v}_n$ and the PU density $\rho_{p,i}$.

Fig. 6.6(a) shows that when an SU moves slowly (Region **I** for the case $\epsilon = 40\,\text{m}$), the sensing interval will be determined by PU traffic patterns, i.e., λ_{busy} and λ_{idle}, whereas, when it moves quickly (Region **II**), the interval will be determined by the speed of SUs. We have made a similar observation regarding PU density in Fig. 6.6(b).

6.4.2 Design of Optimal Guard Distance

The selection of guard distance, ϵ, entails an interesting tradeoff in exploring spectrum opportunities in the time and space domains. That is, a larger guard distance (thus enlarging the areas of PPRs) will reduce spatial spectrum opportunities. However, this allows SUs to perform sensing less frequently and spend more time on data transmission, thus increasing the spectrum opportunities in the temporal domain.

Definition 6.3 *(Average channel utilization)* Average channel utilization *is defined as the average fraction of time a mobile SU can access the channel $i \in \mathcal{K}$, i.e.,*

$$u_{i,n} = \mathbb{E}\left\{ 1 - \frac{\sum_{j=1}^{N_{s,i,n}} T_{s,i} - T_{sw,i}}{T_i} \right\}, \tag{6.13}$$

where $N_{s,i,n}$ is the number of times SU n performs spectrum sensing within the channel access epoch T_i. $T_{s,i}$ and $T_{sw,i}$ are the times spent for a one-time sensing and switching for channel i, respectively. Without loss of generality, we assume $T_s = T_{s,i}\ \forall i$ and $T_{sw} = T_{sw,i}\ \forall i$.

Definition 6.4 *(Spatio-temporal spectrum opportunity) The availability of channel $i \in \mathcal{K}$ in the spatio-temporal domain, denoted as Λ_i, is defined as the long-term average fraction of the time a mobile SU can access the channel, i.e., $\Lambda_i = \gamma_i u_i$ where γ_i and u_i are defined in Eqs.* (6.6) *and* (6.13)*, respectively.*

Fig. 6.7(a) plots the spatio-temporal channel availability Λ_i for various guard distances ϵ_i. As shown in the figure, when ϵ_i is too small (i.e., $\epsilon_i < 3$ m), Λ_i is 0 because of the need to sense the channel continuously, i.e., $t_i^* = T_{s,i}$. When ϵ_i is relatively small, Λ_i suffers from a large (temporal) sensing overhead, whereas when ϵ_i is too large, Λ_i suffers from decreased spatial spectrum opportunities.

Proposition 6.1 *(Optimal guard distance) The optimal guard distance ϵ^* that maximizes spatio-temporal spectrum opportunity, Λ_i, is given as:*

$$\epsilon_i^* = \frac{R_{e,i}\bar{v}\,T_{s,i} + \sqrt{(R_{e,i}\bar{v}T_{s,i})^2 + \frac{2\bar{v}T_{s,i}(R_{e,i} - \bar{v}\,T_{s,i})}{\pi\rho_{p,i}\varpi_{busy,i}}}}{2(R_{e,i} - \bar{v}T_{s,i})}, \tag{6.14}$$

where $R_{e,i}$ is the keep-out radius, $\bar{v}$ the average speed of SUs, $T_{s,i}$ the sensing time, $\rho_{p,i}$ the primary density, and $\varpi_{busy,i}$ the steady-state probability of a `busy` *state for channel i.*

Proof The average fraction of area which is not covered by the PPRs can be approximated as $\gamma_i(\epsilon_i) \approx e^{-f(\epsilon_i)}$ from Eq. (6.6) where $f(\epsilon_i) = \rho_{p,i}\varpi_{busy,i}\pi(R_{e,i}+\epsilon_i)^2$. Assuming switching overhead is negligible compared to the average OFF period, i.e., $T_{sw} \ll \lambda_{off}^{-1}$, u_i can be approximated as $u_i \approx 1 - \frac{\bar{v}\,T_{s,i}}{\epsilon_i}$. Then, the channel availability in the spatio-temporal domain can be expressed as:

$$\Lambda_i(\epsilon_i) \approx \gamma_i(\epsilon_i)u_i(\epsilon_i) \approx e^{-f(\epsilon_i)}\Big(1 - \frac{\bar{v}\,T_{s,i}}{\epsilon_i}\Big). \tag{6.15}$$

It can be easily shown that $\frac{\partial^2 \Lambda_i(\epsilon_i)}{\partial\, \epsilon_i^2} < 0$. By taking the first-order derivative of $\Lambda(\epsilon_i)$ and setting it to zero, we have:

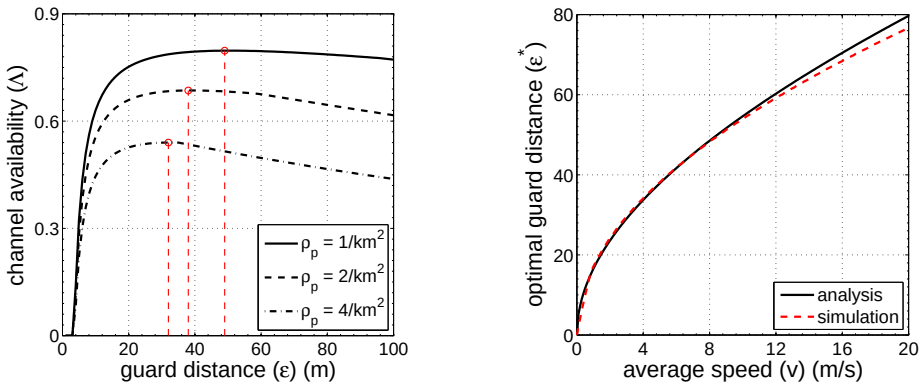
$$\frac{\partial \Lambda_i(\epsilon_i)}{\partial\, \epsilon_i} = e^{-f(\epsilon_i)}\bigg(- 2\rho_{p,i}\varpi_{busy,i}\pi(R_{e,i}+\epsilon_i)\Big(1 - \frac{\bar{v}\,T_{s,i}}{\epsilon_i}\Big) + \frac{\bar{v}\,T_{s,i}}{\epsilon_i^2}\bigg) = 0. \tag{6.16}$$

For mathematical simplicity, we assume that the term $2\rho_{p,i}\varpi_{busy,i}\pi\epsilon_i$ can be approximated as 0 in Eq. (6.16), which provides the following quadratic equation:

$$(R_{e,i} - \bar{v}\,T_{s,i})\epsilon_i^2 - R_{e,i}T_{s,i}\epsilon_i - \frac{\bar{v}\,T_{s,i}}{2\pi\rho_{p,i}\varpi_{busy,i}} = 0. \tag{6.17}$$

Then, by solving Eq. (6.17), the proposition follows.

Interestingly, Fig. 6.7(b) shows that the optimal guard distance increases as SUs' average speed increases. This results from balancing the tradeoff between temporal and spatial spectrum opportunities—i.e., it is better to increase the guard distance at the cost of reduced spatial spectrum opportunity, than to reduce the sensing interval. The figure shows that our analytical results closely match the exhaustive-search-based simulation results.



(a) Impact of guard distance (b) Optimal guard distance (ϵ^*)

Figure 6.7: Optimal guard distance (ϵ^*): (a) Channel availability Λ_i depends significantly on the design of guard distance, and (b) optimal guard distance differs for different SU mobilities. The parameters are set to $\rho_s = 10/\text{km}^2$, $\lambda_{idle,i} = 0.1$, $\varpi_{idle,i} = 0.4\,\forall i \in \mathcal{K}$, and $\rho_p = 2/\text{km}^2$ in (b).

6.5 Distributed Spectrum Access Strategy in Mobile CRNs

We now derive an optimal channel selection (access) strategy that maximizes each secondary link's throughput. In multi-user CRNs, it is important to consider the channel contention overhead, as it can affect the achievable throughput significantly. However, it may be infeasible for mobile SUs to estimate the interference on each channel in real time. Thus, we assume that all the SUs in the network follow the same channel access strategy, and derive the optimal strategy by taking into account SUs' mobility-dependent spectrum opportunity as well as channel access contention among SUs as follows.

Let us denote the mixed channel selection vector by $\mathbf{p} = [p_1, p_2, \ldots, p_{|\mathcal{K}|}]^T$ where $\sum_{i \in \mathcal{K}} p_i = 1$. Then, the total number of SUs selecting channel i in the network can be approximated as Np_i, where N is the total number of SUs in the network, which can be estimated as $N \approx \rho_s A$. A is the entire network coverage area and ρ_s is the average SU density. The probability that an arbitrarily-chosen SU on channel i has $m \in \mathbb{N}$ interfering neighbors, that have chosen the same channel, follows a Binomial distribution, i.e., $M_i \sim B(m; Np_i - 1, f_i)$. Here, $f_i = \frac{\pi R_{I,i}^2}{A}$ is the ratio of the SU's interference region to the total network area, where $R_{I,i}$ is the interference range of an SU on channel i.

The expected throughput of secondary link n can then be expressed as:

$$E[R_n] = \sum_{i=1}^{K} p_i \sum_{m=0}^{Np_i-1} \frac{\Lambda_i}{m+1} \binom{Np_i - 1}{m} f_i^m (1-f_i)^{Np_i-m-1}$$
$$= \frac{1}{N} \sum_{i=1}^{K} \Lambda_i \Big(\frac{1-(1-f_i)^{Np_i}}{f_i} \Big), \tag{6.18}$$

where $K = |\mathcal{K}|$ is the total number of licensed channels.

Then, the problem of finding an optimal channel selection strategy $\mathbf{p}^\star$ can be cast into the following optimization problem ($\mathbf{P1}$):

$$\text{minimize} \quad \mathbb{F}(\mathbf{p}) = -\sum_{i=1}^{K} \Lambda_i \Big(\frac{1-\bar{f}_i^{Np_i}}{f_i} \Big)$$
$$\text{subject to} \quad \sum_{i=1}^{K} p_i = 1 \quad \text{and} \quad \mathbf{p} \succeq 0,$$

where $\bar{f}_i = 1 - f_i$ for brevity.

To find the optimal sensing strategy $\mathbf{p}^\star$, we first show the convexity of $\mathbb{F}(\mathbf{p})$ by examining the second-order derivative of $\mathbb{F}(\mathbf{p})$ w.r.t. p_i, i.e.,

$$\frac{\partial^2 \mathbb{F}(p_i)}{\partial p_i^2} = \bar{f}_i^{Np_i} (\ln(\bar{f}_i^N))^2 > 0. \tag{6.19}$$

The inequality in Eq. (6.19) is straightforward. Hence, $\mathbb{F}(p)$ is convex in $\mathbf{p} \in [0,1]^K$.

Since the objective function is convex and constraints are affine, we now have a convex optimization problem. The Lagrangian with multipliers $\lambda \in \mathbb{R}^K$ and $\nu \in \mathbb{R}$ is given as:

$$L(\mathbf{p}, \lambda, \nu) = \sum_{i=1}^{K} \Lambda_i (\bar{f}_i^{Np_i} \ln(\bar{f}_i^N))) - \sum_{i=1}^{K} \lambda_i p_i + \nu (\sum_{i=1}^{K} p_i - 1)$$
$$= -\sum_{i=1}^{K} ((\lambda_i - \nu) p_i - \Lambda_i (\bar{f}_i^{Np_i} \ln(\bar{f}_i^N)) - \nu,$$

where $\lambda \succeq 0$ and $\nu = 0$.

Then, the Lagrange dual function, i.e., the minimum value of the Lagrangian over $\mathbf{p}$, is

Algorithm 8 OPTIMAL CHANNEL SELECTION ALGORITHM

// **Initialization**
$\mathbf{p} \leftarrow [\frac{1}{K}, \ldots, \frac{1}{K}]^T$ // $\mathbf{p}$ is channel selection probability
$\mathbf{p}_{prev} \leftarrow \mathbf{p}$
$\Delta \leftarrow \infty$
$\varepsilon \leftarrow 0.01$ // condition for the convergence
while ($\Delta > \varepsilon$) **do**
    Update the SU density on each channel $\rho_{s,i} \leftarrow \rho_s p_i$
    Update the keep-out radius $R_{e,i}$ using Eq. (6.3)
    Update the optimal guard distance ϵ_i^* using Eq. (6.14)
    Update the spatio-temporal channel availability $\Lambda_i(\epsilon_i^*)$
    Update the channel selection vector $\mathbf{p}$ using Eq. (6.23)
    $\Delta \leftarrow \mathbf{p} - \mathbf{p}_{prev}$
    $\mathbf{p}_{prev} \leftarrow \mathbf{p}$
end while
return $\mathbf{p}$

given as:

$$\begin{aligned} g(\lambda, \nu) &= \inf_{\mathbf{p}} L(\mathbf{p}, \lambda, \nu) \\ &= \sum_{i=1}^{K} \inf_{\mathbf{p}} (-(\lambda_i - \nu) p_i + \Lambda_i (\bar{f}_i^{N p_i} \ln(\bar{f}_i^N))) - \nu. \end{aligned}$$

It can be easily shown that there exists $\mathbf{p}$ such that the constraints hold with strict inequality, i.e., $p_i > 0 \, \forall i \in \mathcal{K}$ and $\sum_{i=1}^{K} p_i = 1$. Therefore, according to Slater's condition, strong duality holds with zero optimal duality gap.

The Karush-Kuhn-Tucker (KKT) conditions are given as:

$$\mathbf{p}^\star \succeq 0, \quad \sum_{i=1}^{K} p_i^\star = 1 \tag{6.20}$$

$$p_i^\star \left(\lambda^\star + \Lambda_i f_i^{-1} \bar{f}_i^{N p_i} \ln(\bar{f}_i^N) \right) = 0 \tag{6.21}$$

$$\lambda^\star + f_i^{-1} \bar{f}_i^{N p_i} \ln(\bar{f}_i^N) \geq 0. \tag{6.22}$$

By solving the above system of equations, we can derive the optimal channel selection strategy, $\mathbf{p}^\star$, as described in the following proposition.

Proposition 6.2 *(Optimal channel selection strategy) The optimal channel selection vector*

$\mathbf{p}^\star$ *that maximizes the expected secondary network throughput is:*

$$p_i^\star = \begin{cases} \left[\frac{-\ln(\Lambda_i)+\ln(f_i)+\ln(-N\ln(\bar{f}_i))-\ln(\lambda^\star)}{N\ln(\bar{f}_i)}\right]^+ & \textit{if } \varpi_{idle,i} > 0 \\ 0 & \textit{if } \varpi_{idle,i} = 0, \end{cases} \tag{6.23}$$

where $\Lambda_i = \gamma_i u_i \, \forall i \in \mathcal{K}$ *and* $\lambda^\star$ *is a constant s.t.* $\sum_{i=1}^{K} p_i = 1$.

Eq. (6.23) indicates that the channel selection probability p_i increases as the channel availability Λ_i increases, thus confirming our intuition. Interestingly, the optimal channel selection vector $\mathbf{p}^\star$ in Eq. (6.23) depends on SU density on each channel as the number of SUs affects the selection of guard distance (in Eq. (6.6)), influencing the amount of spatial spectrum opportunity. This coupling between channel selection strategy and spatial channel availability requires an iterative algorithm to find the optimal strategy, as described in Algorithm 8.

Proposition 6.2, however, provides the following counter-intuitive observation:

Corollary 6.1 *The optimal channel selection probability becomes more uniform as the number of SUs in the network increases, i.e.,* $\forall i \in \mathcal{K}$,

$$p_i^\star \to \frac{1}{K} \quad \textit{as} \quad N \to \infty, \tag{6.24}$$

where K *is the number of licensed channels, and* N *is the total number of SUs in the network.*

Corollary 6.1 indicates that the optimal channel selection probability becomes almost independent of spatio-temporal spectrum opportunities as SU density approaches infinity. The is because, when there exists a large number of SUs, the benefit from heterogeneous spatio-temporal spectrum opportunities becomes negligible due to the high level of interference among SUs.

6.6 Performance Evaluation

We evaluate the performance of the proposed spectrum sensing and distributed channel-selection schemes. We first describe the simulation setup, channel-selection schemes for performance comparisons, and performance metrics. Then, we present key evaluation results.

6.6.1 Simulation Setup

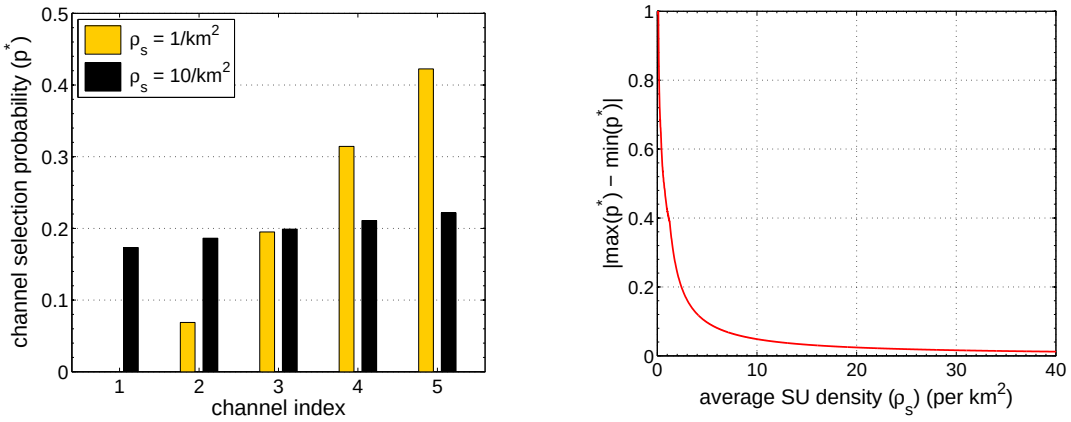
We consider a CRN in which mobile SUs coexist with PUs in a $5\,\text{km} \times 5\,\text{km}$ area. Throughout the simulation, we assume that there are 5 licensed channels,[5] and that the average channel `idle` probability is in the range of $[0.3, 0.7]$, unless specified otherwise. We also assume that λ_{idle} is 0.1 for all the channels and that average density of SUs ρ_s ranges in $[1, 10]/\text{km}^2$. We assume that the path-loss exponent α is 4, the SUs' transmit power P_o is $100\,\text{mW}$, the reference distance d_o is $1\,\text{m}$, the PUs' transmission range R_o is $250\,\text{m}$, the interference temperature limit (`ITL`) is $0.1\,\text{mW}$, and the sensing triggering threshold ξ is 0.3. We further assume that channel sensing and switching times are $T_s = 0.5\,\text{s}$ and $T_{sw} = 1\,\text{s}$, respectively.

To comparatively evaluate the efficacy of the proposed channel-selection scheme, we compare the following: (i) random channel selection (`RAND`), (ii) optimal channel selection strategy based only on PUs' temporal channel usage statistics (`OPT-T`), and (iii) optimal channel selection strategy based on PUs' spatio-temporal channel usage statistics (`OPT-ST`). In `RAND`, SUs randomly select a channel with an equal probability. In `OPT-T`, SUs use the channel-selection probability in Eq. (6.23) while setting $\gamma_i = \varpi_{idle,i}\ \forall i \in \mathcal{K}$ (thus eliminating the impact of heterogeneous PU density on channels). On the other hand, In `OPT-ST`, SUs fully exploit the spatio-temporal channel-usage characteristics of PUs.

To quantify the efficacy of the proposed algorithms, we use the following three main performance metrics:

- normalized secondary network throughput, i.e., $\frac{\sum_n R_n}{N}$,

[5] Although the number of available channels depends on wireless environments, we observed similar results for different numbers of channels.



(a) Impact of channel availability (b) Impact of SU density

Figure 6.8: Optimal channel-selection probability: (a) The optimal channel-selection strategy depends on the average channel availability (ϖ_{idle}), but (b) the effects of PU traffic statistics decreases as SU density increases. The parameters are set to $\boldsymbol{\varpi}_{idle} = [0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7]$, $\bar{v}$ is fixed at 4 m/s, and $\rho_{p,i} = 1/\text{km}^2\ \forall i \in \mathcal{K}$.

- throughput fairness (Jain's index [75]), i.e., $\frac{(\sum_n R_n)^2}{N \sum_n R_n^2}$, and
- normalized energy consumption in spectrum sensing, i.e., the fraction of time a CR device spent on sensing during channel access,

where R_n is the throughput of secondary link n, and N is the total number of secondary links in the network.

6.6.2 Optimal Channel Selection

6.6.2.1 Impact of Temporal Channel Availability

We first study the impact of PUs' temporal channel-usage statistics on the optimal channel-selection strategy. For this, we fix the PU density at $\rho_{p,i} = 1/\text{km}^2\ \forall i \in \mathcal{K}$ and set different channel `idle` probabilities, i.e., $\boldsymbol{\varpi}_{idle} = [0.3, 0.4, 0.5, 0.6, 0.7]$ (ϖ_{idle} increases with increasing channel index).

Fig. 6.8(a) shows SUs' preference to access channels with a higher average channel `idle` probability, i.e., $p_i > p_j$ when $\varpi_{idle,i} > \varpi_{idle,j}$. Interestingly, when SUs are densely populated, i.e., $\rho_s = 10/\text{km}^2$, the impact of PUs' temporal channel-usage statistics on the channel-selection strategy decreases. This is clearly shown in Fig. 6.8(b) where the largest difference in the channel-selection probability (i.e., $|\max(\mathbf{p}^\star) - \min(\mathbf{p}^\star)|$) decreases with

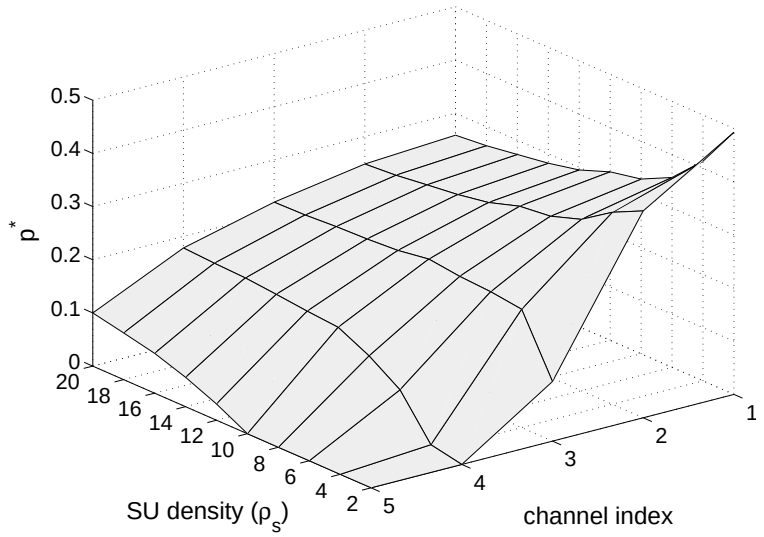


Figure 6.9: Impact of PU density on $\mathbf{p}^\star$: Spatial distribution of PUs affects the optimal channel-selection probability. $\boldsymbol{\rho}_p = [0.1,\ 0.2,\ 0.5,\ 1,\ 2]/\text{km}^2$ (p_i increases with increasing channel index), $\varpi_{idle,i} = 0.4$ and $\lambda_{idle,i} = 0.1\ \forall i \in \mathcal{K}$.

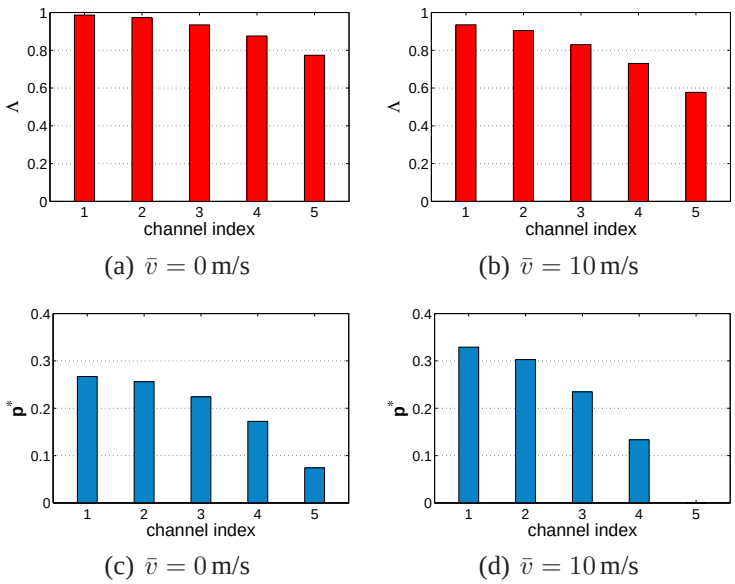
increasing SU density. Intuitively, as the number of SUs in the network increases, their channel access time decreases due to the need for sharing the channel. Thus, as the density tends to infinity, the achievable throughput of SUs becomes close to 0, regardless of the PUs' channel usage statistics.

6.6.2.2 Impact of Spatial Channel Availability

Fig. 6.9 shows the impact of PU density on the optimal channel-selection strategy. In the simulation, we assume a different PU density on each channel, while assuming that temporal channel usage statistics, i.e., ϖ_{idle}, are the same for all channels. The figure indicates that, the lower the PU density (channel index), the higher the channel-selection probability. However, PU density becomes less influential as the average SU density increases, similar to the case in Fig. 6.8(b).

6.6.2.3 Impact of SUs' Speed

Fig. 6.10 shows the impact of SUs' average speed on spatio-temporal channel availability Λ_i (in Figs. 6.10(a)-(b)), and on the optimal channel-selection strategy $\mathbf{p}^\star$ (Figs. 6.10 (c)-(d)). As shown in the figures, the SUs' speed has different consequences on channel



(a) $\bar{v} = 0\,\text{m/s}$ (b) $\bar{v} = 10\,\text{m/s}$

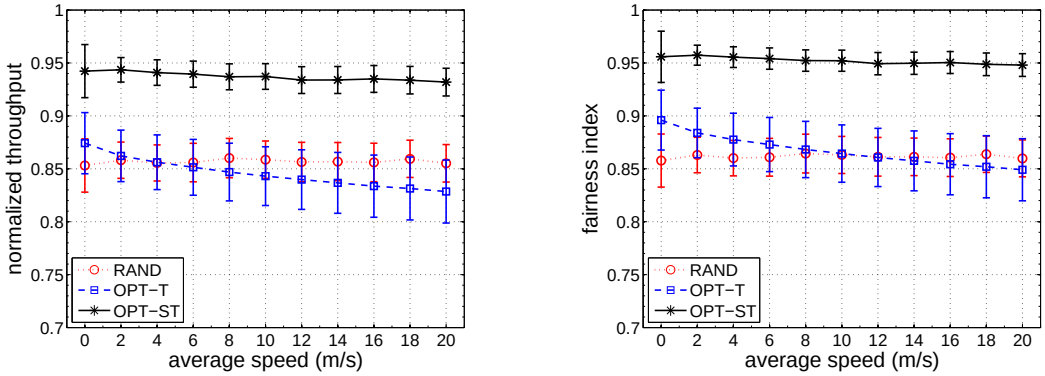
(c) $\bar{v} = 0\,\text{m/s}$ (d) $\bar{v} = 10\,\text{m/s}$

Figure 6.10: Impact of SUs' speed on Λ and $\mathbf{p}^\star$: The spatio-temporal channel availability depends on the SUs' speed, thus affecting the optimal channel-selection strategy $\mathbf{p}^\star$. The parameters are set to $\boldsymbol{\rho}_p = [0.1,\ 0.2,\ 0.5,\ 1,\ 2]/\text{km}^2$, $\rho_s = 10/\text{km}^2$, and $\varpi_{idle,i} = 0.4\,\forall i \in \mathcal{K}$.

availability (Λ), depending on the density of PUs on each channel; Λ decreases faster when PU density is high. As a result, the SUs' preference to access channels with a low PU density increases as their speed increases. The simulation settings are described in Fig. 6.10.

6.6.3 Performance Comparison

Next, we compare the performance of the three channel-selection schemes (i.e., `RAND`, `OPT-T`, and `OPT-ST`) in terms of throughput, fairness, and energy-efficiency. In the simulations, we set the average PU density on each channel to $\boldsymbol{\rho}_p = [0.1, 0.2, 0.5, 1, 2]/\text{km}^2$. The channel idle probabilities ϖ_{idle} are randomly selected in $[0, 1]$ such that $\sum_{i\in\mathcal{K}} \varpi_{idle,i} = 1$ for each network topology. The results are obtained from simulation runs over 10^3 randomly-generated topologies. Figs. 6.11 and 6.12 plot the average and $\pm 0.25\,\sigma$ intervals of throughput and fairness, under various SUs' speed and density.



(a) Throughput (b) Fairness

Figure 6.11: Performance of the proposed distributed channel-selection algorithm: OPT-ST outperforms other channel-selection schemes in terms of (a) network throughput and (b) fairness (Jain's index), under all simulated scenarios. In the simulation, the average SU density was fixed at $\rho_s = 1/\text{km}^2$.

6.6.3.1 Throughput and Fairness

Fig. 6.11(a) shows that the proposed OPT-ST outperforms the other channel-selection schemes (i.e., OPT-T and RAND) under all simulated scenarios, thanks to its ability to optimally select channels by exploiting the heterogeneous spatial/temporal spectrum opportunities of each channel. On the other hand, the performance of OPT-T decreases as SU speed increases, because the spatial spectrum opportunity becomes more diverse with higher SU mobility (see Fig. 6.10), which is not considered in OPT-T. Fig. 6.11(b) indicates that OPT-ST achieves the highest fairness among the three channel-selection schemes, as it correctly incorporates the impact of heterogeneous spectrum opportunities and channel access contention among SUs in the optimal channel selection strategy.

Fig. 6.12 shows the impact of SU density on throughput performance. As shown in the figure, the throughput degrades as SU density increases, mainly because of the increased level of SUs' contention for channel access. In addition, the performance of OPT-ST becomes close to RAND's as the density increases, since the optimal channel-selection strategy tends to become similar to a uniform distribution, which can be seen in RAND, in a dense network, as observed in Fig. 6.9.

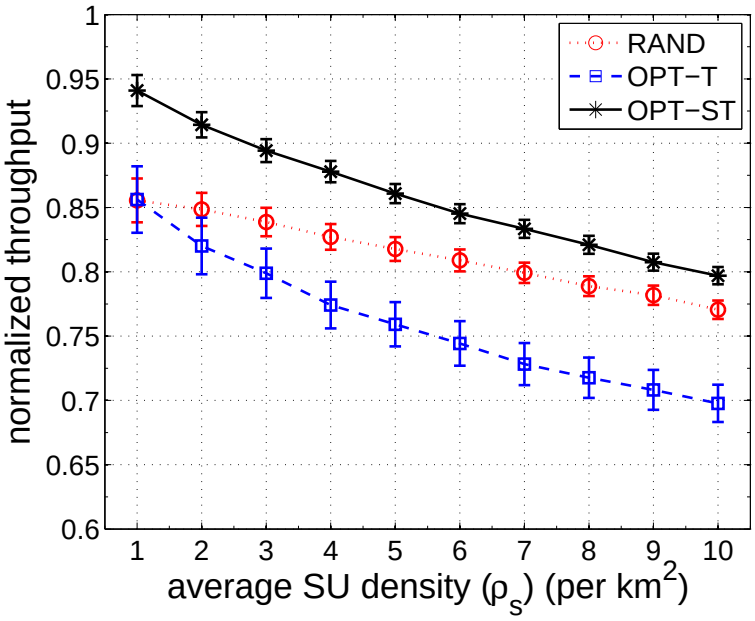


Figure 6.12: Impact of SU density on throughput performance: The performance of `OPT-ST` decreases as the average SU density increases. In the simulation, the SUs' speed is fixed at 4 m/s.

6.6.3.2 Energy Saving in Spectrum Sensing

Finally, we study the energy-saving perspective in spectrum sensing. Frequent spectrum sensing can consume a considerable amount of energy, especially in battery-powered mobile CR devices. Fig. 6.13 plots the CR's normalized energy consumption in different settings, including use of a fixed guard distance (i.e., $\epsilon = 20, 40$ m) and use of the optimal guard distance (ϵ^*). The figure indicates that energy consumption due to spectrum sensing in mobile CR devices can be reduced by up to 74 % while ensuring primary protection.

6.7 Related Work

Spectrum sensing has been studied extensively as a key technology for primary detection and protection [23, 67, 71, 88, 100, 134, 149]. Most existing work, however, focuses on optimizing the sensing interval based on PUs' *temporal* channel-usage statistics. To validate such channel models, Wellens *et al.* [149] studied the impact of channel-occupancy statistics obtained from extensive measurements on the performance of MAC-layer sensing schemes. They showed that the channels with longer `busy`/`idle` periods follow ex-

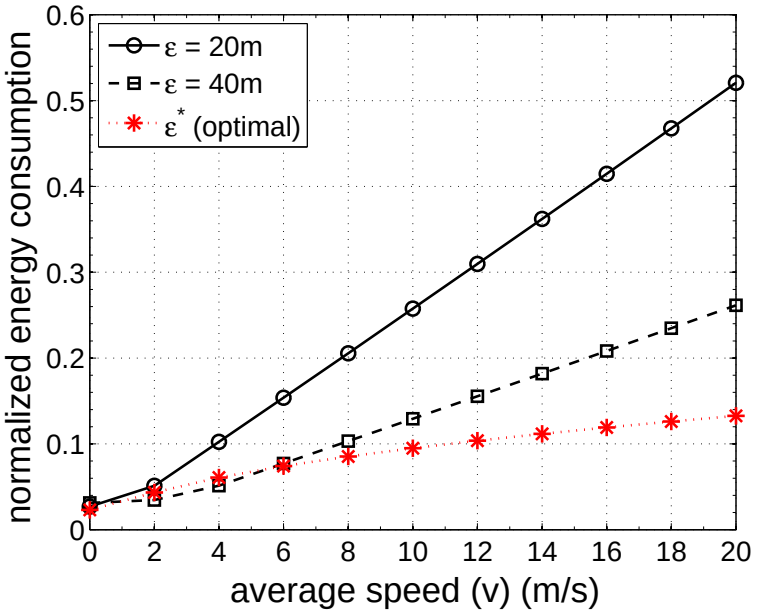


Figure 6.13: Energy savings via the use of optimal guard distance: SUs can save energy significantly due to spectrum sensing via the optimal guard distance, while meeting the primary interference constraints.

ponential distributions and that spectrum sensing and access strategies designed under the assumption of exponentially-distributed PU traffic are highly efficient. However, such models hinge on the assumption of *stationary* CRNs, in which both PUs and SUs are stationary. Thus, they may not be suitable for *mobile* CRNs, in which channel availability depends on dynamically changing SUs' locations. By contrast, we model channel availability from a mobile SU's perspective by incorporating the impact of SU mobility (e.g., speed).

Despite its practical importance, the problem of allowing mobile SUs in CRNs has received little attention. The IEEE 802.22 standard draft provides a two-stage sensing (TSS) mechanism [22], but it is designed exclusively for the detection of a stationary TV transmitter, and does not specify any efficient mechanisms for spectrum sensing for portable/mobile CRs. Recently, the FCC [49] imposed a minimum sensing interval of 60 seconds for TV band devices (TVBD). However, this may not be sufficient to protect PUs from interference induced by SU mobility. Moreover, while most previous work focused on either scheduling spectrum sensing [100] or spatial CR deployment [64, 145] for primary protection, we *jointly* exploit guard distance and the sensing interval to maximize spatio-temporal spectrum opportunities for mobile SUs.

6.8 Conclusion

Taking mobility into consideration is vitally important for full realization of the benefits of DSA in CRNs. In this chapter, we considered the case of a CRN with mobile SUs. We identified and addressed the three fundamental challenges in maximizing spectrum efficiency in mobile CRNs. In particular, we presented a novel channel-availability model, a mobility-aware spectrum-sensing strategy, and an optimal distributed channel-selection (or access) strategy tailored to mobile CRNs. Our evaluation results verified the correctness of our channel-availability model under various SU mobility patterns. Our performance comparison study has also shown that the channel-access strategy improves the throughput and fairness of mobile SUs significantly over the conventional strategy that relies solely on PUs' temporal channel-usage statistics.

CHAPTER 7

Optimal Spectrum Pricing in Dynamic Spectrum Market

7.1 Introduction

In this chapter, we propose a new spectrum-pricing model in a dynamic spectrum market (DSM), where, in order to maximize their profits, wireless service providers (WSPs) compete with *heterogeneous* spectrum resources—channels with disparate center frequencies and propagation profiles. In our model, we assume the availability of a wide range of heterogeneous bands in the spectrum plane, and analyze the spectrum pricing-demand relationship between WSPs (in the service plane) and SUs (in the user plane). In the user plane, SUs sublease and *share* the spectrum that provides the maximum utility. These features—spectrum heterogeneity and spectrum sharing—are essential for us to understand the WSPs' pricing competition in a DSM, but have not been explored well.

Based on a realistic price-demand model, we formulate WSPs' pricing competition as a non-cooperative game, taking into account the SUs' desire to maximize their utility. Here "utility" refers to spectrum consumers' judgements about the tradeoff between achievable capacity and spectrum leasing cost. We examine the existence and uniqueness of the spectrum price Nash equilibrium (NE), which depends upon SU density (i.e., total spectrum demand[1]) and spectrum heterogeneity. Our investigation into the effects of three essential features—(i) spectrum heterogeneity, (ii) spectrum sharing among SUs, and (iii) total spectrum demand (i.e., SU density)—provides useful insights and practical guidelines for

[1]We refer to "spectrum demand" as the number of SUs in a DSM rather than the SUs' bandwidth demand.

designing spectrum pricing and purchase strategies in DSM.

7.1.1 Contributions

In summary, this chapter makes the following contributions:

- Introduction of a new DSM model where WSPs with heterogeneous spectrum resources compete for a higher market share. We demonstrate the impact of spectrum heterogeneity via in-depth measurements on a GNURadio/USRP testbed. To the best of our knowledge, this is the first attempt to analyze the impact of spectrum heterogeneity in a DSM.
- Investigation of a new spectrum price-demand model based on the desire of SUs to maximize their own utility, by evaluating the impact of spectrum heterogeneity, spatial spectrum sharing, and total spectrum demand.
- Derivation of SUs' optimal WSP selection strategy based on a mean-field approach to study how spectrum heterogeneity affects market equilibrium. Our mean-field approach simplifies the market model using a set of differential equations, and is shown to effectively approximate an exact model using large-dimension Markov chains.
- Modeling of the pricing strategies among WSPs as a non-cooperative game and identification of the key factors that influence the NE points, taking into account the price-demand relation caused by the utility maximizing behavior of SUs.

7.1.2 Organization

The remainder of this chapter is organized as follows. Section 7.2 describes the duopoly DSM model and formulates the pricing game among WSPs as a non-cooperative game. Section 7.3 shows the impact of spectrum heterogeneity via in-depth measurements on a software-defined radio testbed. Section 7.4 analyzes the impact of SU density on their achievable utility by analyzing mutual interference among SUs. Section 7.5 studies the SUs' optimal WSP selection strategies that maximize achievable utility. Section 7.6 derives the WSPs' optimal spectrum pricing strategy based on a realistic price-demand func-

tion. Section 7.7 reviews existing work for spectrum pricing in DSM. Finally, Section 7.8 concludes this chapter.

7.2 System Model and Assumptions

In this section, we present a signal propagation model for our analysis which will be used throughout this chapter. We then define the DSM model, utility functions of SUs and WSPs, and formulate the pricing competition of WSPs as a non-cooperative game.

7.2.1 A Dynamic Spectrum Market (DSM) Model

We consider a duopoly DSM where two WSPs compete in the same geographical area, as illustrated in Fig. 7.1. Each WSP is assumed to have long-term access rights for a licensed channel with a different center frequency, obtained from primary spectrum owners, for example, via auction [80, 165]. WSPs then grant access rights to their channels to multiple SUs by advertising the spectrum price, either via database query or direct broadcasting over a dedicated control channel. WSPs have access to complete information about customer population (i.e., SU density) and their preferences (i.e., SUs' utility).[2] Each WSP possesses a single channel for leasing, and we focus on the case where the WSPs' leased channels have considerably different center frequencies, thus exhibiting disparate wireless signal propagation characteristics.

Interactions among DSM participants can be modeled as a 3-tier structure [25, 76] (see Fig. 7.1) consisting of: (i) the *spectrum plane*, where licensed spectrums are auctioned and sold to wireless service providers (WSPs), (ii) the *service plane*, where WSPs sublease the spectrum by enticing SUs with competitive prices and good spectrum quality, and (iii) the *user plane*, where SUs choose the WSP that maximizes their utility. Although spectrum pricing competition in DSM has been studied extensively [45, 73, 113, 114], most existing work has not considered spectrum-heterogeneity as a primary factor in establishing the pricing strategy (except [154]).

[2] Learning mechanisms can be used to infer such information when it is not available [154].

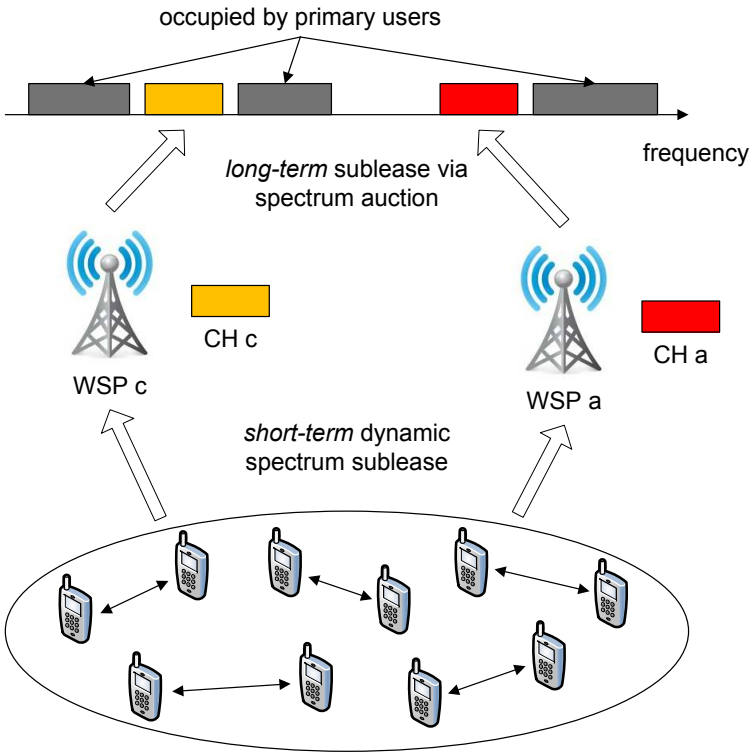


Figure 7.1: A duopoly dynamic spectrum market model: WSPs compete with heterogeneous channels (leased from the primary spectrum owners) to entice more SUs in the same geographical area in order to maximize profit.

For the user plane, we consider an ad-hoc secondary network consisting of a set, $\mathcal{N}$, of transmitter-receiver pairs, referred to as SUs. Each pair constitutes a basic unit for spectrum leasing; in essence, SUs purchase short-term rights to access the channels from a WSP at a fixed spectrum price set by the WSP. We assume that SUs are SDR devices (e.g., USRP [3]) with CR-capability. By exploiting the ability to access a wide range of spectrum bands, SUs aim to maximize their utility (i.e., the difference between the channel capacity and spectrum leasing cost in Eq. (7.2)) by choosing the "best" WSP. SUs are randomly deployed in areas following a point Poisson process [17, 32] with average density ρ, i.e., the distribution of the number of active links within the deployment area, A, is $n_A \sim Poisson(n; \rho|A|)$. Note that although we consider an ad hoc secondary network, our analysis can also be applied to an infrastructure-based network model, where communication between an access point (or base station) and its associated clients is one-to-one at any given time.

7.2.2 Signal Propagation and Spectrum Reuse Model

Signal propagation is known to depend on the center frequency of each channel: the lower the frequency band, the better the signal propagation characteristics. For ease of analysis without losing key insights to be gained from spectrum heterogeneity, we consider the following simple signal propagation model that reflects the impact of spectrum heterogeneity [59]:

$$P_R = P_o\, g_c(r) = P_o \left(\frac{c_o}{f_c}\right)^{\alpha} r^{-\alpha}, \quad \text{(Watts)} \tag{7.1}$$

where P_R is the received signal power, P_o the transmission power, c_o the speed of light, i.e., $c_o = 3 \times 10^8$ m/s, f_c the center frequency of the channel c, r the distance between the transmitter and receiver, and $\alpha\,(> 2)$ the path-loss exponent.[3] We assume that all the SUs in the network use the same fixed transmission power level P_o. While we use a simple signal propagation model, more realistic models (e.g., [74]) could be used for specific wireless environments (e.g., indoor or outdoor) at the cost of complexity of analysis. Since shadow or multi-path fading is shown to not affect average interference significantly [69], we do not consider it in our model.

Buddhikot *et al.* [19] suggested three different models for spectrum sharing, which are referred to as *exclusive use*, *shared use*, and *commons* models. These models overcome the limitations of the traditional *command-and-control* model. In order to focus on the impacts of spectrum heterogeneity in a DSM, in this chapter we consider the *exclusive use* model, in which primary spectrum owners grant their exclusive spectrum access rights to a third party (e.g., WSPs). This exclusive model is suitable for spectrum bands with relatively long ON/OFF primary activity periods, e.g., DTV channels. Besides, this model can provide high quality-of-service (QoS) and reliability because it does not require frequent performance of spectrum sensing by SUs, or frequent service interruptions due to primary activities.

[3] We assume that the path-loss exponent is $\alpha > 2$ so that the cumulative interference does not diverge as the network size grows.

7.2.3 Utility-Maximizing Spectrum Demand and User Preference

One of our main contributions is to derive a realistic price-demand function in the DSM, driven by SUs' desire to maximize their utility. Specifically, the utility function of SU $i \in \mathcal{N}$, which is associated with WSP (channel) c,[4] is defined as the difference between the SUs' achievable link capacity and spectrum price:

$$\mathcal{U}_i(c) = B \log\left(1 + \frac{P_o g_{c,i}}{I_{c,i} + N_o}\right) - p_c, \tag{7.2}$$

where B is channel bandwidth, $g_{c,i}$ the channel gain between the secondary transmitter and receiver, N_o the noise power level, and P_o the transmit power. (Per FCC regulation, there is a cap on transmit-power levels for SUs.) The average of cumulative interference power caused by the SUs on channel c at the receiver of link i is denoted by $I_{c,i}$, and p_c denotes the spectrum price (per unit time). To simplify the analysis, we assume that all the secondary transmitter-receiver pairs are separated by the same distance, and thus the channel gain $g_{c,i}$ only depends on channel frequency, i.e., $g_{c,i} = g_c \ \forall i$. For the similar reason, we assume $I_{c,i} = I_c \ \forall i$. Henceforth, we omit the subscript i for brevity. We consider a fixed (unit) bandwidth demand from SUs, i.e., $B = 1$ for all channels.

Let $\mathcal{C} = \{c, a\}$ denote the set of WSPs (channels) in a DSM. Based on the utility function in Eq. (7.2), SU i selects the channel $c_i \in \mathcal{C}$ that maximizes expected utility, i.e.,

$$c_i^* = \arg\max_{c \in \mathcal{C}} \mathcal{U}_i(c), \tag{7.3}$$

where $\mathcal{C}$ is the set of channels available at WSPs, e.g., $\mathcal{C} = \{a, c\}$ for a duopoly DSM.

7.2.4 Spectrum Pricing Game among WSPs

The main objective of WSPs is to maximize their profit by leasing the licensed channel to multiple SUs at the highest possible leasing price. Therefore, WSPs play a pricing game

[4] We equate a WSP with its channel(s).

to compete for market share. The payoff (profit) function of a WSP $c \in \mathcal{C}$ is defined as:[5]

$$\mathcal{V}_c(p_c, p_{-c}) = N_c(p_c, p_{-c}) \cdot p_c - b_c, \tag{7.4}$$

where N_c is the number of SUs associated with WSP c (spectrum demand), p_c the spectrum leasing price, and b_c the fixed investment cost, i.e., the fee paid to the primary spectrum owner for the long-term spectrum lease (per unit time).

Note that analyzing the price-demand relationship, i.e., $N_c(p_c, p_{-c})$, is not straightforward. Traditional economic models tends to assume a known relation between WSPs' price and SUs' demand. However, in our model, the spectrum demand N_c, i.e., the number of SUs on channel c, depends not only on WSPs' spectrum leasing prices $\{p_c, p_{-c}\}$, but also on the channel quality (capacity) determined by the frequency-dependent co-channel interference, as shown in Eq. (7.2). SUs can freely choose the WSP that maximizes their payoff. Thus, WSPs must consider spectrum heterogeneity in devising an optimal spectrum pricing strategy that maximizes profit.

Based on the WSPs' utility in Eq. (7.4), the spectrum pricing game among WSPs can be defined as shown below.

Definition 1 (*Spectrum pricing game between WSPs*) A spectrum pricing game between the WSPs can be formalized as a strategic choice:

$$p_c^* = \arg\max_{p_c \in \mathbb{R}} \mathcal{V}_c(p_c, p_{-c}), \tag{7.5}$$

where p_{-c} denotes the price chosen by competing WSPs.

In what follows, we first demonstrate the impact of spectrum heterogeneity in Section 7.3, analyze SU utility in Section 7.4, and derive the optimal WSP selection and spectrum pricing strategies in Sections 7.5 and 7.6, respectively.

[5]Let the subscript $-c$ denote the competitor of WSP c.

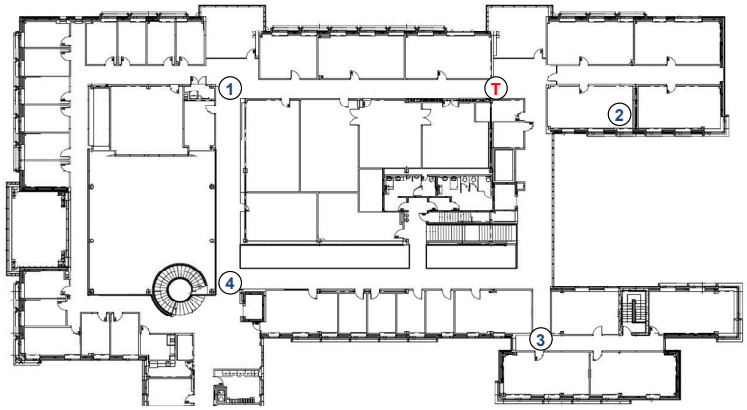


Figure 7.2: Software-defined radio testbed: GNURadio/USRP2 nodes are placed at different locations on the fourth floor of the CSE Building at the University of Michigan.

7.3 Characteristics of Spectrum Heterogeneity

In this section, we demonstrate the effects of spectrum heterogeneity on received signal strength (RSS) via measurements on our software-defined radio testbed. We first describe our experimental setup and then present the measurement results.

7.3.1 Experimental Setup

To evaluate the impact of spectrum heterogeneity, we constructed a GNUradio/USRP2 [6] testbed on the fourth floor of the Computer Science and Engineering (CSE) Building at the University of Michigan. This floor has multiple offices and conference rooms and relatively straight corridors, which allow us to evaluate the impact of spectrum heterogeneity under both line-of-sight (LOS) and non-line-of-sight (NLOS) settings.

We deployed 5 USRP2 nodes in the topology shown in Fig. 7.2. We placed the transmitter at a fixed location in the corridor (denoted as circled T in the figure), and purposely placed 4 receiver nodes at different locations (e.g., corridors and offices, denoted as 1-4 in the figure) to test various signal-propagation environments. The measurements were done at night to minimize the effects of environmental changes, such as moving people/obstacles and interference from other networks. This allows us to focus on evaluating the impact of spectrum heterogeneity on network performance without the need to deal with all the transient network dynamics, e.g., the fluctuations in RSS due to moving obstacles.

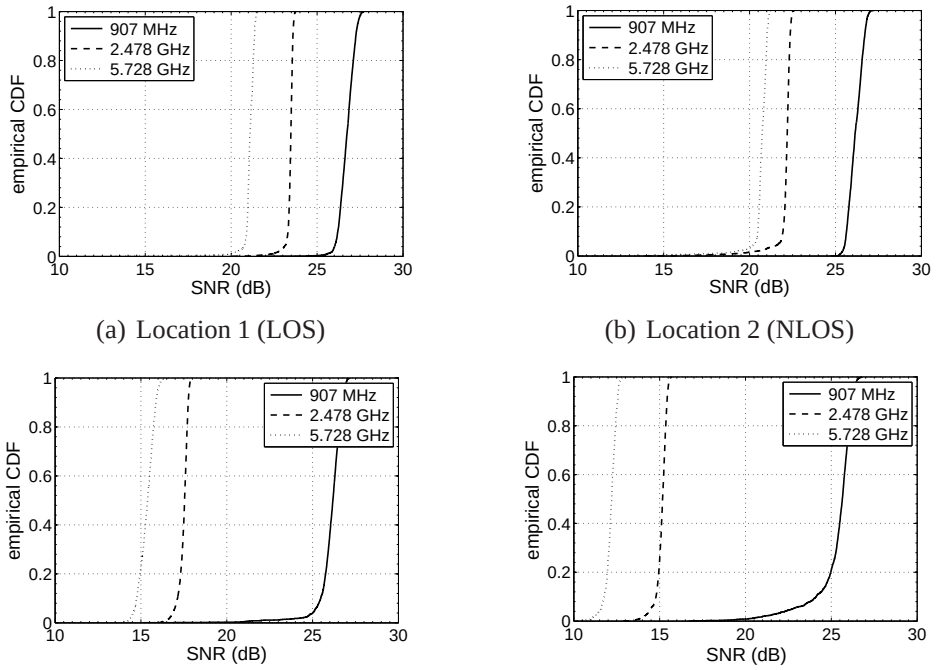
We equipped the USRP2 nodes with two different sets of daughterboards and antennae that operate on different spectrum bands. For high-frequency spectrum, we mounted the VERT2450 (dual Band 2400–2480 MHz and 4.9–5.9 GHz omnidirectional antenna) on a XCVR2450 board (2.4–2.5 GHz and 4.9–5.85 GHz dual-band daughterboard). For low-frequency bands, we mounted the VERT900 (824-960 MHz omnidirectional antenna) on a WBX board (50 MHz to 2.2 GHz daughterboard). Both the XCVR2450 and WBX have the same transmit power level (20 dBm).

We use the benchmark dbpsk encoding/decoding module in GNURadio to test the signal quality on different spectrum bands. The bit rate is set to 0.1 Mbps and each BPSK symbol goes through a raised-root-cosine filter with 8 taps, resulting in a signal bandwidth of 50 KHz. Through experiments, we found that the transmit power of the testbed increases linearly with transmit gain. Therefore, we set the transmission gain of both XCVR2450 and WBX to the maximum, to ensure that they have the same output power.

7.3.2 Experimental Results

To evaluate the effect of spectrum heterogeneity, we measured the signal-to-noise ratio (SNR) of a transmitted signal on three different frequency bands, i.e., 907 MHZ, 2.478 GHz, 5.728 GHz, at four different receiver locations. Receiver location 1 is LOS setting, and the rest are NLOS settings. The measurement lasted 5 minutes for each experiment. Note that the USRP RF circuits have different gains for different frequency bands. Hence we first calibrate the output power for different frequency bands, so that they may have comparable SNRs at short distances. In this way, the hardware artifacts are isolated and for each link, the signal quality only depends on its frequency.

Fig. 7.3 plots the empirical cumulative distribution function (c.d.f.) of the measured SNR. The figure clearly indicates the impact of spectrum heterogeneity: the lower the frequency, the higher the SNR, regardless of the receiver locations. Fig. 7.3(d) shows that, when the receiver is in the NLOS setting, high frequency bands, i.e., 2.478 GHz and 5.728 GHz, suffer from significant deterioration in signal strength because of the obstacles (i.e., the walls between the transmitter and receiver). On the other hand, the low frequency



(a) Location 1 (LOS) (b) Location 2 (NLOS)

(c) Location 3 (NLOS) (d) Location 4 (NLOS)

Figure 7.3: Impact of spectrum heterogeneity: The distribution of measured SNR depends significantly on the center frequency of the channel; the lower the frequency, the higher the SNR due to the better signal propagation characteristics.

band, i.e., $907\,$MHz, achieves a relatively high SNR thanks to its good wall-penetration characteristics.

Next, we study the signal propagation characteristics of different spectrum bands by measuring the RSS (in dB). We place the transmitter at a fixed location and vary the transmitter-receiver separation from $15\,$m to $45\,$m in an indoor, LOS setting. Fig. 7.4 illustrates that low frequency band shows consistent advantage for all the distance settings. In addition, RSS linearly decreases when the logarithmic distance, i.e., $10\log_{10}(d)$, between the transmitter-receiver pair increases, regardless of the center frequency. This again verifies the trend predicted by the empirical propagation model in Eq. (7.1).

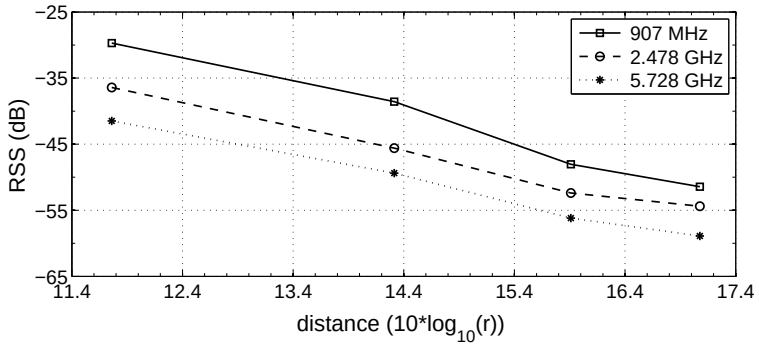


Figure 7.4: Signal propagation over heterogeneous spectrum: RSS decreases almost linearly as the logarithmic distance between the transmitter and receiver increases.

7.4 Analysis of Secondary Utility under Spectrum Heterogeneity

In this section, we characterize co-channel interference (i.e., $I_{c,i}$ in Eq. (7.2)) among SUs to capture the effects of spectrum heterogeneity and spectrum sharing on the achievable capacity of SUs.

For spectrum sharing among co-channel SUs, we consider the physical model [62] where all the SUs can transmit at the same time. We consider the physical model rather than the protocol model [62]. [6] Note that, although we consider the physical model, the main insights would not be different for the protocol model. We approximate the distribution of co-channel interference, I_c, on channel $c \in \mathcal{C}$, by quantifying the interference from SUs located inside and outside the interference range, R_I^c, which is defined as $R_I^c \triangleq \sup\{r \in \mathbb{R} \,\big|\, P_o\, g_c(r) > \eta\}$ where η is a predefined threshold that depends on the desired data rate, modulation scheme, etc.

We first approximate the sum of co-channel interference caused by SUs located inside the interference range as a Gaussian random variable. In practice, secondary systems maintain a certain distance between them to avoid interference, so we assume that the minimum distance between secondary transmitters is sufficiently large (e.g., $> 10\,\text{m}$). The total interference at a fixed point in a uniformly-distributed wireless network can be accurately

[6] The physical and the protocol models [62] are most widely used for modeling wireless interference. In the former, SUs can transmit data concurrently but share the channel via a non-orthogonal multiplexing protocol (e.g., CDMA). In the latter, SUs multiplex the channels using an orthogonal scheme (e.g., OFDMA), and the per-user capacity is inversely proportional to the number of interfering neighbors.

approximated as a Gaussian random variable [82].[7]

Let $\mathcal{G}_{in,c} = \sum_{S_c} g_c(r)$ denote the *normalized* interference (i.e., sum of channel gains) from a set S_c of co-channel SUs located inside the interference range. Then, the probability density function (p.d.f.) of Gaussian random variable $\mathcal{G}_{in,c} \sim N(\mu_c, \sigma_c^2)$ is given as:

$$\mathcal{G}_{in,c}(x) = \frac{1}{\sqrt{2\pi}} \exp\left(- \frac{(x-\mu_c)^2}{2\,\sigma_c^2} \right), \tag{7.6}$$

where the mean (m_1) and variance (m_2) of the interference $\mathcal{G}_{in,c}$ is given as [99]:

$$\begin{aligned} m_k(\rho, c) &= \rho_c \pi ((R_I^c)^2 - \epsilon^2) \int_{\epsilon}^{R_I^c} \frac{2r}{((R_I^c)^2 - \epsilon^2)} (g_c(r))^k \mathrm{d}r \\ &= \frac{2\rho_c \pi}{(k\alpha - 2)} \left(\frac{c_o}{f_c} \right)^{\alpha k} \left(\frac{1}{\epsilon^{k\alpha-2}} - \frac{1}{(R_I^c)^{k\alpha-2}} \right), \end{aligned} \tag{7.7}$$

where ϵ is the minimum separation distance from the receiver.

We now quantify the total interference caused by SUs located outside the interference range.

Lemma 7.1 The total interference caused by SUs on channel c located outside the interference region (i.e., unit disk of radius R_I^c centered at the receiver) can be approximated as:

$$I_{out,c} = 2\pi\, P_o \left(\frac{c_o}{f_c} \right)^{\alpha} \frac{\rho_c\, (R_I^c)^{2-\alpha}}{\alpha - 2}. \tag{7.8}$$

Proof See Appendix B.

Finally, based on Eqs. (7.7) and (7.8), the interference caused by SUs on channel c can be approximated as $I_c \sim \mathcal{N}(\mu_c, \sigma_c^2)$ where

$$\mu_c = E[I_{in,c}] + I_{out,c} = \left(\frac{c_o}{f_c} \right)^{\alpha} \frac{2\pi\rho_c P_o}{\alpha - 2} \left(\frac{1}{\epsilon^{\alpha-2}} \right) \tag{7.9}$$

$$\sigma_c^2 = \frac{\rho_c \pi P_o}{\alpha - 1} \left(\frac{c_o}{f_c} \right)^{2\alpha} \left(\frac{1}{\epsilon^{2\alpha-2}} - \frac{1}{(R_I^c)^{2\alpha-2}} \right), \tag{7.10}$$

[7] Note that we can ignore the impact of multi-path fading on channel selection because the time scale of channel switching is much larger than the average fading duration.

where ρ_c is the density of SUs (links) on channel c. Similarly, we can derive interference for channel a (i.e., μ_a and σ_a^2). Eq. (7.9) indicates that the total interference linearly increases with SU density on channel c, ρ_c, which can be approximated as $\rho_c \approx \frac{N_c}{|A|}$ where A is the entire network area.

The interference distribution in Eqs. (7.9) and (7.10) is a function of center frequency f_c, which serves as the basis for developing an optimal WSP selection strategy and for analyzing the pricing game among WSPs with heterogeneous spectrum bands.

7.5 Optimal WSP Selection Strategy via Mean-Field Approach

In this section, we derive the optimal WSP selection strategy for SUs using a mean-field approach, assuming that the WSPs possess different spectrum bands. We begin with a mean-field approximation of the spectrum market. We then prove its convergence, and derive the optimal WSP selection strategy in the mean-field regime.

7.5.1 A Mean-Field Model for Spectrum Market

The mean field method [16] is a simple and effective way of analyzing the state evolution of a large number of interacting objects. In particular, it is suitable for analyzing how the local behavior of individual nodes affects the global properties of a large-scale network. In our problem, an SU's behavior is described by its type (i.e., its preferred WSP), and the global properties are the steady-state distribution of SU types.

Our mean-field approach uses differential equations to approximate the evolution of the market, whose state converges to the fixed point of the equation (namely, the *mean-field limit*) under certain conditions [16]. In what follows, we first use a mean-field model to describe how the DSM evolves, and then justify the convergence of the market to its mean-field.

7.5.2 Evolution and Convergence of the Market

We first provide the following key definitions:

- A *link* is defined as a connected transmitter-receiver pair with *active* traffic. Therefore, a link can be considered "newly joined" if it has just switched from an idle period to a period of bursty transmission.

- Let N be the number of active links. Links can "join" and "depart" according to a Poisson distribution. However, we assume that the link population evolves to a steady state, such that the departure rate equals the arrival rate, and the total number of links remains roughly constant.

- Let λ be the traffic rate of a link. We also assume that the ON-OFF traffic pattern of a link is bursty, following a Poisson distribution with rate λ.

- Let $N_c(t)$ denote the total number of active links using channel c at time t. Links are classified according to the channel that they use, i.e., a link i is of *type* c, if it selects channel $c \in \mathcal{C}$.

We study the evolution of the spectrum market within a short period of time, Δt. The number of newly joined secondary links within this period is $N\,\lambda\,\Delta t$. This is also the number of departed links within Δt, since we focus on a steady state of the SU population when the departure rate equals the arrival rate. Each newly joined link leases a channel from a WSP with a short-term contract. Note that active links that have already leased a channel are in transmitting/receiving mode, and must maintain their current channel (WSP) selection.

Let P_c be the probability that, for a randomly selected link i, channel c provides the maximum utility, i.e.,

$$P_c = Pr\big\{c = \arg\max_{c^* \in \mathcal{C}} \mathcal{U}_i(c^*)\big\}, \quad \forall c \in \mathcal{C}, \tag{7.11}$$

where the utility $\mathcal{U}_i(c)$ is defined in Eq. (7.2). Then, among the newly joined links within Δt, the number of links selecting channel c is $N\,\lambda\,\Delta t\,P_c$.

The total number of channel c SUs, i.e., links using channel c, in the network at time $(t + \Delta t)$ is:

$$N_c(t + \Delta t) = N_c(t) + N\,\lambda\,\Delta t\,P_c - N_c(t)\,\lambda\,\Delta t. \tag{7.12}$$

Eq. (7.12) describes the evolution of a market. The market equilibrium can be defined as a fixed point of the market evolution:

$$\begin{aligned} &\frac{\partial N_c(t)}{\partial t} = \frac{N_c(t + \Delta t) - N_c(t)}{\Delta t} = N\,\lambda\,P_c - N_c(t)\,\lambda = 0 \\ &\iff P_c = \frac{N_c(t)}{N}. \end{aligned} \tag{7.13}$$

Eq. (7.13) indicates that the probability that an SU selects WSP c is equivalent to the fraction of SUs using channel c, which is referred to as the *channel occupancy measure*, i.e., $\Pi_c(t) = N_c(t)/N$. Intuitively, the occupancy measure, $\Pi_c(t)$, reflects the market share of WSP c at time t.

Proposition 7.1 (Convergence of channel occupancy) The channel-occupancy measure $\mathbf{\Pi} = \{\Pi_a, \Pi_c\}$ converges to a deterministic process in the continuous-time domain.

Proof See Appendix C.

From now on, we will focus on deriving the channel (WSP) selection probability P_c in the mean-field model of Eq. (7.12), which depends primarily on three key factors: (i) amount of interference I_c on channel c, (ii) spectrum leasing prices p_c, and (iii) total spectrum demand ρ. Note that the interference intensity I_c depends on the occupancy measure of channel c, which, in turn, affects the channel-selection probability P_c. This circular dependency eventually converges to a fixed point, i.e., the mean-field limit of market dynamics.

7.5.3 SUs' Optimal Selection of WSPs

We now analyze the SUs' optimal channel (WSP) selection strategy, assuming that each SU is a rational market entity that selects a WSP to maximize his utility. In making a strategic choice, each SU takes into account the achievable capacity and leasing cost, but

$$P_c = Q\left(\frac{(\frac{c_o}{f_a})^\alpha \frac{2\pi\rho P_o}{\alpha-2}(\frac{1}{\epsilon^{\alpha-2}})(e^{p_c-p_a}P_c - P_a) + N_o(e^{p_c-p_a}(\frac{f_c}{f_a})^\alpha - 1)}{(\frac{c_o}{f_a})^\alpha \sqrt{\frac{\pi\rho P_o}{\alpha-1}} \sqrt{P_a(\frac{1}{\epsilon^{2\alpha-2}} - \frac{1}{(R_I^a)^{2\alpha-2}}) + e^{2(p_c-p_a)}P_c(\frac{1}{\epsilon^{2\alpha-2}} - \frac{1}{(R_I^c)^{2\alpha-2}})}} \right) \quad (7.14)$$

$$P_a = 1 - P_c. \quad (7.15)$$

cannot directly affect the price set by the WSPs. This model mirrors a real-world market economy where customers are obedient price-takers, but the joint effect of their choices causes the sellers to compete and reach an equilibrium price.

We derive the optimal WSP selection strategy in a mean-field regime for given spectrum prices $\boldsymbol{p} = \{p_a, p_c\}$. For an arbitrarily-chosen SU in a DSM, the probability that channel c provides better utility is:

$$\begin{aligned} P_c &= Pr\big(\mathcal{U}_c - \mathcal{U}_a > 0\big) \\ &= Pr\Big(\log\Big(\frac{P_o g_c}{I_c + N_o}\Big) - \log\Big(\frac{P_o g_a}{I_a + N_o}\Big) > p_c - p_a \Big) \\ &= Pr\Big(\log\Big(\frac{I_a + N_o}{I_c + N_o}\Big) > p_c - p_a - \alpha \log\Big(\frac{f_a}{f_c}\Big) \Big) \\ &= Pr\Big(I_a + N_o - e^{p_c - p_a}\Big(\frac{f_c}{f_a}\Big)^\alpha (I_c + N_o) > 0 \Big), \end{aligned} \quad (7.16)$$

where p_c (p_a) and f_c (f_a) are the price and center frequency of channel c (a), respectively.

Note that a more commonly used approach for analyzing the equilibrium state is to equate the user's utility, i.e., $U_i(c) = U_i(a)$ (e.g., [114]). However, such an equilibrium state may not be reached depending on the network environment, as will be shown in Section 7.6.4.

For given prices, the channel-selection probability P_c depends solely on the interference statistics on channel c. In Eq. (7.16), the interference power on each channel can be approximated as a normal random variable as derived in Section 7.4.

Let $I_{ca} = I_a + N_o - \gamma_{ca}(I_c + N_o)$ where $\gamma_{ca} = e^{p_c - p_a}(\frac{f_c}{f_a})^\alpha$. Note that N_o and γ_{ca} are constants, and I_{ca} is thus the difference between the two Gaussian random variables, which

is also Gaussian. Then, $I_{ca} \sim \mathcal{N}(\mu_{ca}, \sigma_{ca}^2)$ where

$$\mu_{ca} = \mu_a + N_o - \gamma_{ca}(\mu_c + N_o) \tag{7.17}$$

$$\sigma_{ca}^2 = \sigma_a^2 + \gamma_{ca}^2 \sigma_c^2, \tag{7.18}$$

where the mean and variance of the interference are shown in Eqs. (7.9) and (7.10). Then, the channel-selection probability is:

$$P_c = Pr(I_{ca} > 0) = Pr\Big(\frac{I_{ca} - \mu_{ca}}{\sigma_{ca}} > \frac{-\mu_{ca}}{\sigma_{ca}}\Big) = Q\Big(\frac{-\mu_{ca}}{\sigma_{ca}}\Big), \tag{7.19}$$

where $Q(x) = \frac{1}{2\pi} \int_x^\infty e^{-\frac{t^2}{2}} \mathrm{d}t$. Using Eqs. (7.9), (7.10), (7.18) and (7.19), one can derive the channel-selection probabilities.

Proposition 7.2 (*WSP selection strategy*) For the case with two WSPs (channels) c and a, the mean-field limit of the channel-selection strategy P_c and P_a follows Eqs. (7.14) and (7.15).

Proposition 7.2 indicates that the mean-field limit of the WSP selection strategy is influenced not only by the spectrum prices, but also by the channel heterogeneity reflected by interference ranges (R_I^c,R_I^a) and center frequencies (f_c,f_a). This clearly indicates that spectrum heterogeneity can affect the optimal spectrum pricing that maximizes the WSP's profit. Proposition 7.2, however, shows that SUs' traffic intensity λ does not affect the system's steady-state.

Proposition 7.3 (*Asymptotic behavior of WSP-selection strategy*) The optimal WSP-selection probability becomes more uniform as SU density increases, i.e.,

$$P_c \to 0.5 \quad \text{as} \quad \rho \to \infty, \tag{7.20}$$

where ρ is the average SU density, which can be approximated as $\rho \approx \frac{N}{|A|}$.

Proof As $\rho \to \infty$, the WSP-selection probability P_c in Eq. (7.14) reduces to:

$$\lim_{\rho \to \infty} P_c = Q\big(+ \infty \, (e^{p_c - p_a} P_c - P_a)\big). \tag{7.21}$$

Then, we have:

$$\lim_{\rho\to\infty} P_c = \begin{cases} 1 & P_c < \frac{P_a}{e^{p_a - p_a}} \\ 0.5 & P_c = \frac{P_a}{e^{p_a - p_a}} \\ 0 & P_c > \frac{P_a}{e^{p_a - p_a}}. \end{cases} \tag{7.22}$$

In Eq. (7.22), there exists a unique solution, i.e., $\lim_{\rho\to\infty} P_c = 0.5$ when $p_c = p_a$. On the other hand, when $p_c \neq p_a$, there is no solution because $p_c = p_a$ is the unique NE point under the condition $\rho \to \infty$. We will detail the price NE in Section 7.6.

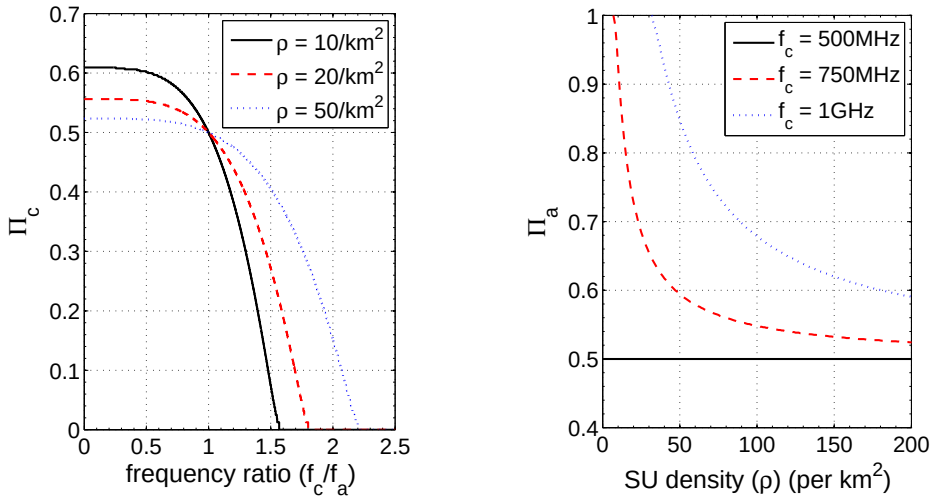
Proposition 7.3 indicates that the WSP-selection probability becomes independent of spectrum heterogeneity when the number of SUs in the network, N, approaches infinity. This is because, when there exist a large number of interferers, interference power dominates noise power, i.e., $I_c \gg N_o$, and as a result, the benefit from low frequency becomes negligible.

7.5.4 Numerical Results

Here we present numerical results that show the behavior of the channel-occupancy measure under different DSM settings.

Fig. 7.5(a) shows the impact of heterogeneous channel frequencies on the channel occupancy, Π_a and Π_c. In the simulations, we fix the center frequency of channel a at $f_a = 500\,\text{MHz}$ and increase the frequency f_c up to $2.5\,\text{GHz}$. We set spectrum prices to $p_a = p_c = 1$, to eliminate the effect of prices on channel occupancy. The figure shows that, when $f_c < f_a$, $\Pi_c > 0.5$, due to the favorable signal-propagation characteristics of channel c; on the other hand, when $f_c > f_a$, $\Pi_c < 0.5$ for the same reason. Interestingly, channel occupancy depends on average secondary network density (i.e., total spectrum demand) ρ. This is because, in a dense network where interference power exceeds noise power, i.e., $N_o \ll I_c$, the benefit of favorable signal-propagation characteristics diminishes. As a result, the channel-occupancy curve becomes flatter. Note that when $f_c = f_a$, $\Pi_c = \Pi_a = 0.5$, regardless of SU density.

Fig. 7.5(b) shows the channel-occupancy measure while varying average SU density in the range $\rho \in [0, 200]/\text{km}^2$. Here we fix the center frequencies at $f_a = 500\,\text{MHz}$ and assume



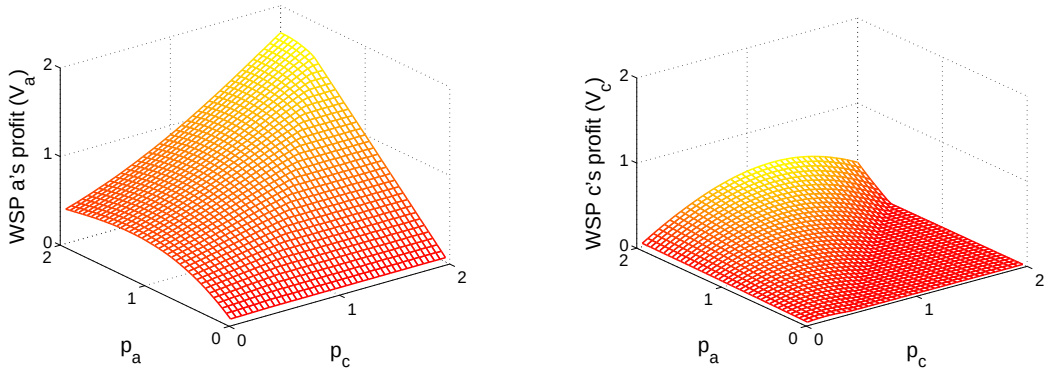
(a) Impact of spectrum heterogeneity (b) Impact of SU density

Figure 7.5: Characterization of channel-occupancy measure: (a) The occupancy of channel c, Π_c, increases as the frequency ratio $\frac{f_c}{f_a}$ decreases, and (b) channel occupancy becomes less sensitive to spectrum heterogeneity as the network density increases. The parameters are set to $\epsilon = 100$ m, and $P_o = 100$ mW, and spectrum prices are fixed at $p_a = p_c = 1$.

$f_c \in \{500\,\text{MHz}, 750\,\text{MHz}, 1\,\text{GHz}\}$. The figure indicates that the channel occupancy Π_a is always greater than or equal to 0.5 due to its favorable signal-propagation characteristics. When SU density is low, the channel occupancy Π_a is close to 1 as most SUs tend to enjoy the favorable signal-propagation characteristics of channel a without worrying about mutual interference. Under these conditions, the DSM behaves monopolistically. However, as SU density increases, the channel-occupancy measure Π_a decreases because, in such a high interference regime, it becomes harder for SUs to exploit the benefits of favorable signal-propagation characteristics. Thus, the DSM behaves like a duopoly. The figure also shows that the occupancy measure approaches 0.5 in all the tested cases, confirming the correctness of Proposition 7.3.

7.6 Equilibrium of the Spectrum-Pricing Game

In this section, we study the impact of spectrum price on the WSP's profit as defined in Eq. (7.4), and characterize the Nash equilibrium (NE) points of pricing strategies.



(a) Profit of WSP a (V_a) (b) Profit of WSP c (V_c)

Figure 7.6: Profit of WSPs: The achievable profit of WSPs depends on spectrum leasing prices $\mathbf{p} = (p_a, p_c)$ and spectrum heterogeneity (i.e., channel frequency). We fix the center frequencies at $f_a = 500$ MHz and $f_c = 1$ GHz, and set SU density to $\rho = 50/\text{km}^2$.

7.6.1 Impact of Spectrum Price on WSP's profit

Here we evaluate the impact of the WSPs' spectrum-pricing strategy $\mathbf{p} = (p_a, p_c)$ on their achievable profits. Without loss of generality, we set the heterogeneous spectrum bands at $f_a = 500$ MHz and $f_c = 1$ GHz. However, we observed similar patterns for different frequency bands. We fixed SU density at $\rho = 50/\text{km}^2$, and set the investment costs in Eq. (7.4) at $b_a = b_c = 0$ to eliminate their impact on WSPs' profit, which will be studied separately in Section 7.6.6.

Fig. 7.6 shows that WSP a always achieves a higher profit than WSP c, i.e., $V_a > V_c$, thanks to its favorable spectrum profile. Fig. 7.6(a) shows that the profit of WSP a (i.e., V_a) monotonically increases as competing WSP c increases its price p_c. This is because WSP a tends to entice more customers due to channel a's better signal-propagation characteristics. The advantage becomes more pronounced when the competitor WSP c sets a higher price and loses part of its market share. In contrast, as shown in Fig. 7.6(b), when WSP c quotes a higher price than that of WSP a, its achievable profit remains 0, i.e., WSP a monopolizes the market. This indicates that channel c is not competitive unless the price of channel a rises above a certain threshold.

Fig. 7.7 shows the impact of relative price, $\frac{p_c}{p_a}$, on WSPs' profit with respect to SU density ρ (i.e., spectrum demand). We set price $p_a = 1$ and vary the price p_c from 0 to

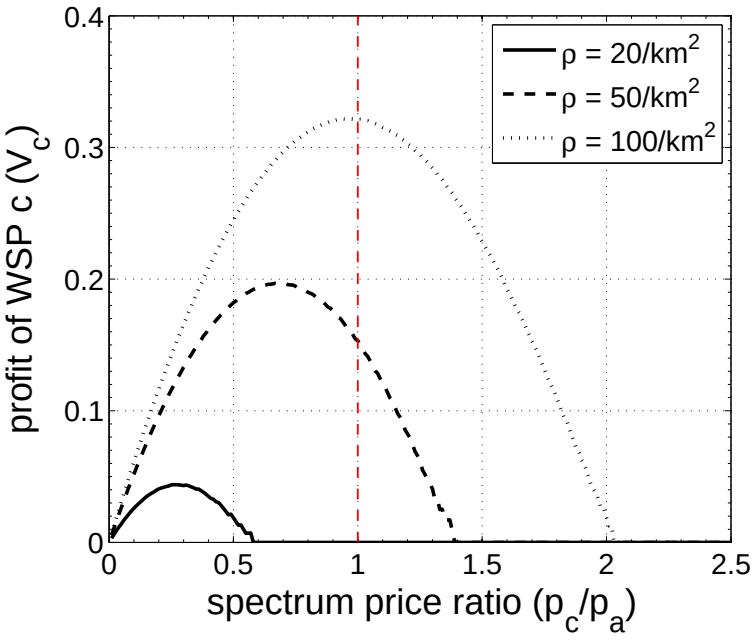


Figure 7.7: Impact of price ratio: There exists an optimal pricing ratio that maximizes profit, and the effect of pricing is coupled with the SUs' density.

2.5. Here we have made three observations. First, the relationship between price ratio and profit is a concave function as shown in Fig. 7.7. When the price ratio is relatively low, WSP c's profit decreases as the price ratio further decreases. Despite the fact that a lower price attracts more SUs, the advantage is limited by increased interference among them. When the ratio is relatively high, the profit also decreases as the ratio further increases, due to the significant decrease in customers. Second, when the price ratio, $\frac{p_c}{p_a}$, is above a certain threshold, the profit V_c becomes 0 (i.e., the profit curve becomes flat) since the high price makes channel c unattractive to customers. However, such a threshold increases with an increasing SU density (i.e., spectrum demand) where WSPs can take advantage of a large number of customers. Third, in a sparse network with low SU density, profit V_c is maximized when $p_c \ll p_a$ because the interference on channel a remains negligible, even when most users are associated with WSP a. In contrast, in a dense network, V_c is maximized when $p_c \approx p_a$ because all SUs will suffer from high interference regardless of the channel characteristics. Therefore, WSP a loses its competitive advantage of superior signal propagation characteristics. Note that this corresponds to our findings in Proposition 7.3 in Section 7.5.3.

7.6.2 Nash Equilibrium for Pricing Game

In a DSM, WSPs must carefully set the spectrum price, since too high a price results in loss of market share, while too low a price will limit their achievable profits. We capture this tradeoff with the notion of Nash equilibrium (NE).

Definition 2 (Spectrum price Nash equilibrium) An NE in the duopoly game is defined as a strategy set $\{p_c^*, p_a^*\}$ that satisfies:

$$p_c^* = \arg\max_{p_c} \mathcal{V}_c(p_c, p_a^*) \tag{7.23}$$

$$p_a^* = \arg\max_{p_a} \mathcal{V}_c(p_c^*, p_a). \tag{7.24}$$

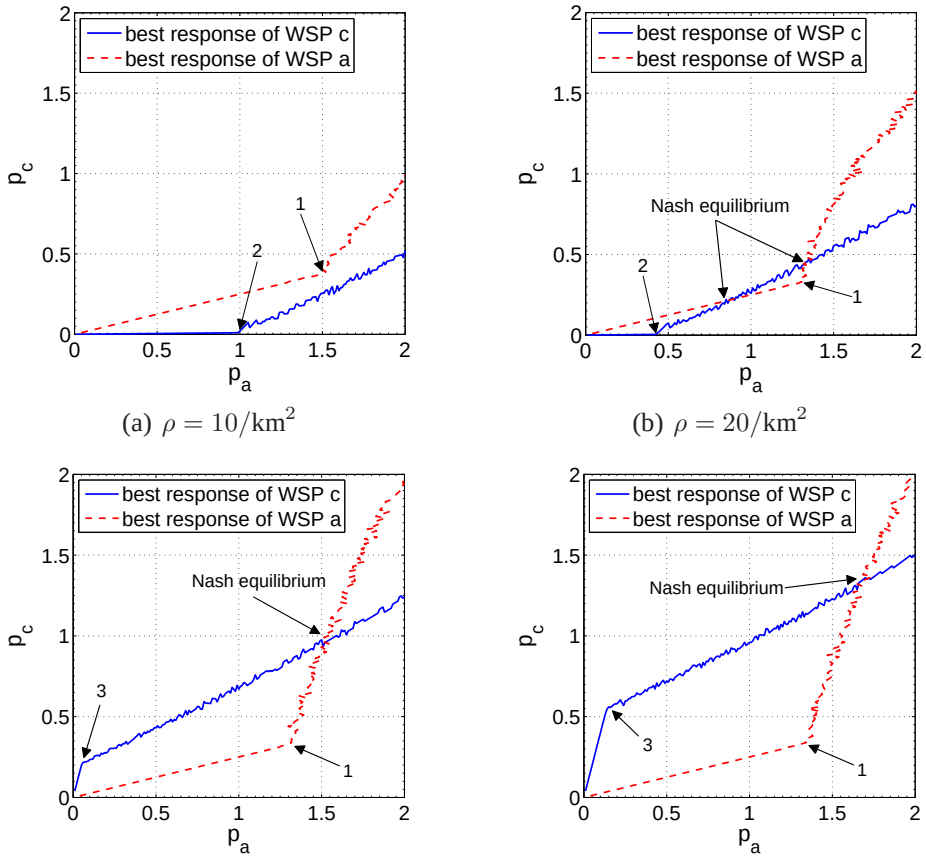
Intuitively, an NE strategy set implies that no player can increase its profit by unilaterally adjusting the price. With the above definition, we can derive the NE of the duopoly game. Unfortunately, it is difficult to find a closed-form expression for the NE. Hence, we numerically solve Eqs. (7.23) and (7.24) using a simple iterative search algorithm to obtain the NE price.

7.6.3 Existence and Uniqueness of Nash Equilibrium

Based on the above definition of NE, we examine the existence and uniqueness of the NE points when SU density changes, which is equivalent to changing the spectrum demand over the entire network. In the simulation, we consider a representative scenario in which the frequency of WSP a is lower than that of WSP c, i.e., $f_a = 500\,\text{MHz}$ and $f_c = 1\,\text{GHz}$, and thus, we expect the NE points to be formed such that $p_a^* > p_c^*$.[8]

Fig. 7.8 shows the best responses of WSPs under different SU densities. We have made three key observations. First, the WSP c's best response (solid lines) increases as the spectrum price p_a increases, and vice versa. This is because the WSPs compete over the same pool of customers in a given network coverage area, and hence, WSPs' optimal spectrum pricing is always relative to the competitors' spectrum prices. That is, if WSP a

[8] Although we presented the NEs for a specific set of frequencies, we observed from simulations a similar behavior for other frequency bands.



(a) $\rho = 10/\text{km}^2$ (b) $\rho = 20/\text{km}^2$

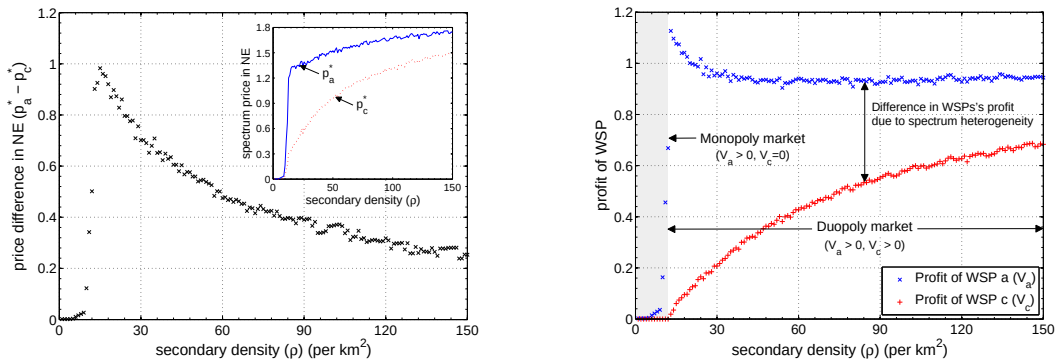
(c) $\rho = 50/\text{km}^2$ (d) $\rho = 100/\text{km}^2$

Figure 7.8: Best response functions for the WSPs: The existence and uniqueness of the NE depends on spectrum heterogeneity as well as the secondary network density. In the simulation, we set $f_a = 500\,\text{MHz}$, $f_c = 1\,\text{GHz}$, and $b_a = b_c = 0$.

quotes a high spectrum price, then the SUs' achievable utility from WSP a will decrease, changing their preference to the competitor, i.e., WSP c. This will allow WSP c to increase its price p_c to reach an equilibrium point.

Note that this relative behavior of spectrum pricing provides an economic incentive to WSPs for collusion. However, such a collusion can be prevented in practice for the following reasons. There will be alternative technologies to access the wireless spectrum, e.g., IEEE 802.11, and hence, WSPs will lose their competitiveness as they advertise unreasonably high prices. Moreover, rational SUs would not purchase the spectrum if their achievable utility (i.e., difference between capacity and price) is too low, i.e., less than 0. Therefore, WSPs cannot set spectrum prices arbitrarily to increase their profit.

Second, when SU density is low, i.e., $\rho = 10/\text{km}^2$, the price NE does not exist because



(a) Difference in NE prices (b) Profit of WSPs

Figure 7.9: Impact of secondary network density under heterogeneous and homogeneous spectrum bands: (a) The difference in NE prices, i.e., $p_a^* - p_c^*$, decreases with increasing SU density where $f_a = 500$ MHz and $f_c = 1$ GHz, and (b) the NE prices increase with SU density. We assume zero investment cost, i.e., $b_a = b_c = 0$, in the simulation.

total spectrum demand is not high enough for WSPs to make a profit. Although $p_c = p_a = 0$ can also be considered as an NE point, the WSPs will avoid this strategy since this NE point will provide a negative revenue to both WSPs. That is, to attract customers, WSPs have to lower their prices until they reach 0, and thus, there is no economic incentive for WSPs to participate in the market. In contrast, with high secondary density, i.e., $\rho > 20/\text{km}^2$, the NEs are formed at some positive values, thus providing economic incentives to WSPs.

Third, Fig. 7.8 indicates that the best responses exhibit *phase transitions* (the transition thresholds denoted as 1,2,3), resulting in a different number of NEs depending on market settings. For example, the figures show that the growing rate of the best responses of WSP a (dashed lines) changes at certain thresholds (denoted as 1). This is because when p_a remains below the threshold, it is optimal for WSP a to increase the price p_a at a higher pace than p_c, i.e., $\frac{\Delta p_a}{\Delta p_c} > 1$, to take advantage of channel a's superior spectrum characteristics. However, when p_a increases beyond the threshold, the high spectrum price limits the growth of the utility of SUs. As a result, channel c becomes more attractive than channel a, and thus, $\frac{\Delta p_a}{\Delta p_c} < 1$. Similarly, the best response of WSP c has the threshold property denoted as 2 and 3 in the figures.

One interesting observation is that, in dense networks, i.e., $\rho = 50, 100/\text{km}^2$, the price p_c increases faster than p_a until p_a reaches the threshold 3. This is because, despite channel

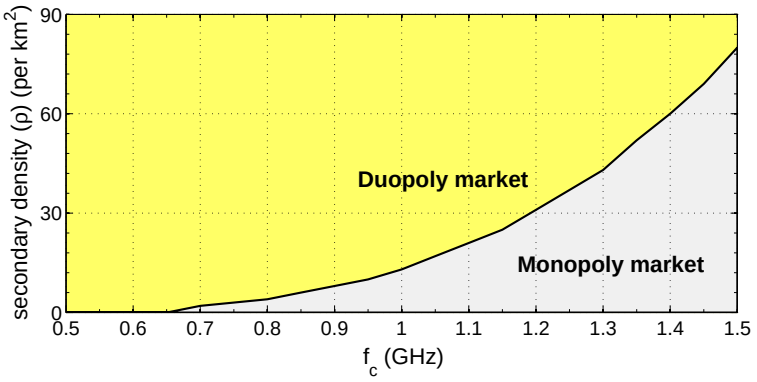


Figure 7.10: Monopoly vs. duopoly DSM: The market can be monopolized (gray area) by WSP a when the channel frequency f_c of the competitor, WSP c, is relatively higher than f_a and the SU density is low. We assume that f_a is fixed at 500 MHz.

a's higher quality, when the price p_a is too low compared to the NE price, WSP c can quote a higher price, i.e., $p_c > p_a$ to maximize its own profit, benefiting from a large number of customers.

7.6.4 Market Dynamics under Various SU Densities

As we observed in Section 7.5, SU density (or spectrum demand) is a critical factor in WSPs' pricing competition. Here we investigate the impact of SU density on market dynamics by examining the NE prices, WSPs' profit, and SUs' utility.

Fig. 7.9(a) shows the difference between the NE prices, i.e., $p_a^* - p_c^*$, as a function of SU density. When the density is low, i.e., $\rho < 10/\text{km}^2$, NE does not exist as we observed in Fig. 7.8(a), and WSPs cannot make a profit because the market (spectrum demand) is too small. As the density increases, however, the NE price of channel a (p_a^*) grows drastically, whereas the price p_c^* remains 0 due to its inferior spectrum profile. This means WSP c cannot make profit if they quote a price greater than $p_c > 0$. As a result, WSP a *monopolizes* the market, as more clearly shown in Fig. 7.9(b) (shaded region). As the density further increases, however, WSP c starts to share the market, i.e., *duopoly* because the SUs on channel a begin to suffer from co-channel interference.

Fig. 7.9(b) shows WSPs' profit defined in Eq. (7.4) for various SU densities. As we discussed, when density is low, WSP a dominates (monopolizes) the market, i.e., $V_a > 0$

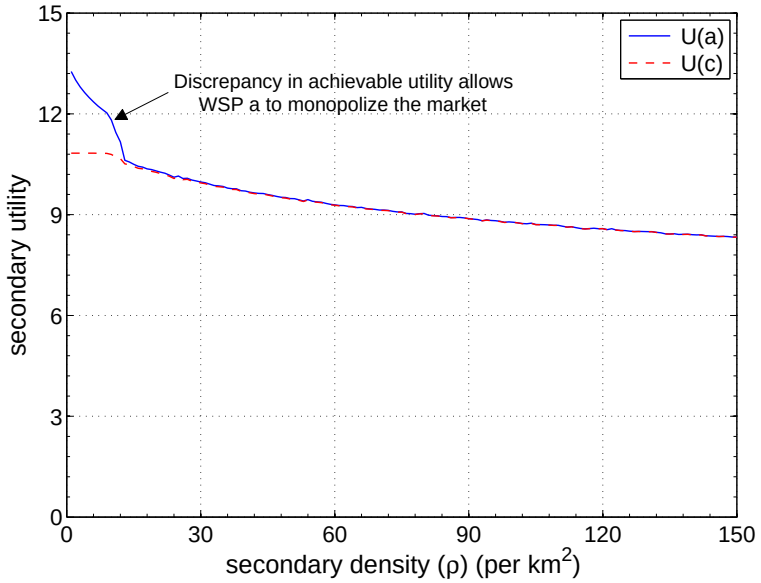


Figure 7.11: Behavior of secondary utility: The achievable secondary utility differs only when SU density is low (i.e., monopoly market), and remains the same in the duopoly market.

and $V_c = 0$, thanks to its superior spectrum profile. As the SU density increases beyond a certain density threshold (i.e., $\rho = 12/\text{km}^2$), the market becomes *duopoly* and the difference in achievable profit decreases as the size of the market grows. Such a threshold density depends on spectrum heterogeneity. Fig. 7.10 clearly shows that the range of SU density below which WSP a monopolizes the market increases as the center frequency of channel c increases. For example, when $f_c = 2f_a$, WSP a will dominate the market until SU density becomes larger than $\rho = 13/\text{km}^2$. In addition, such a boundary of SU density increases super-linearly, partly because of the relationship between received signal strength and center frequency: $P_R \propto f_c^{-\alpha}$, as indicated in Eq. (7.1).

Fig. 7.11 shows SUs' achievable utilities on each channel, i.e., U_a and U_c. The figure shows that, when $\rho < 13/\text{km}^2$, the utility on channel a exceeds that of channel c, i.e., $U_a > U_c$, thus forming the monopoly market. On the other hand, in the duopoly market, there is no difference in achievable utilities, and thus the market is stabilized.

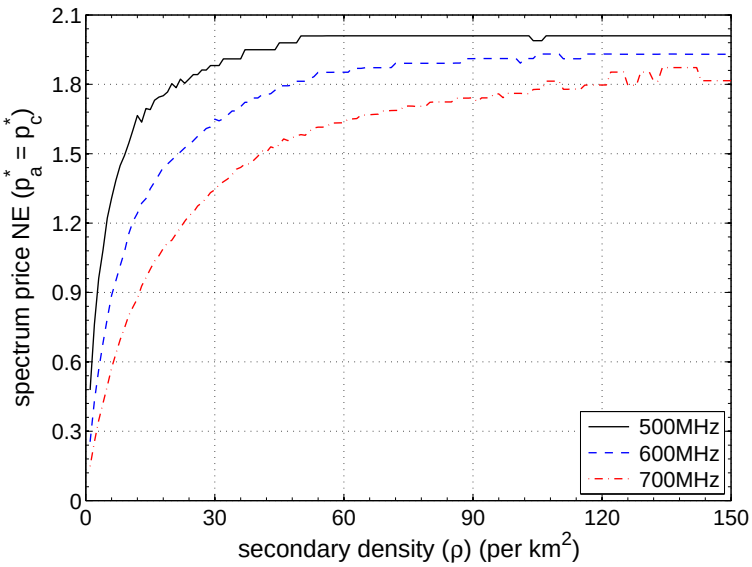


Figure 7.12: The behavior of NE prices in a DSM with homogeneous spectrum bands: The equilibrium prices increase with SU density and converge at different rates; the lower the center frequency, the faster the convergence due to their large interference range.

7.6.5 Price NE under Spectrum Homogeneity

To demonstrate the impact of secondary density, while separating it from spectrum heterogeneity, we consider three *homogeneous* spectrum bands, i.e., $f_a, f_c \in \{500\,\text{MHz}, 600\,\text{MHz}, 700\,\text{MHz}\}$, and plot the corresponding NE points in Fig. 7.12. Due to spectrum homogeneity, the NE prices are equal, i.e., $p_a^* = p_c^*$, regardless of the secondary density. We set the leasing cost $b_a = b_c = 0$ to eliminate its impacts on NE prices. From Fig. 7.12, we have two main observations. First, the equilibrium price increases with increasing secondary density (i.e., total spectrum demand) due to the increasing number of customers. In addition, the lower the frequency band, the higher the price for any given secondary density, since low-frequency bands return higher utility (i.e., capacity minus spectrum price) to the SUs. Second, the equilibrium price converges faster with low frequency bands due mainly to the large interference power (range) of low frequency bands. This is because the potential benefit of using low frequency bands (i.e., a longer transmission range) diminishes faster with secondary density due to their large interference range.

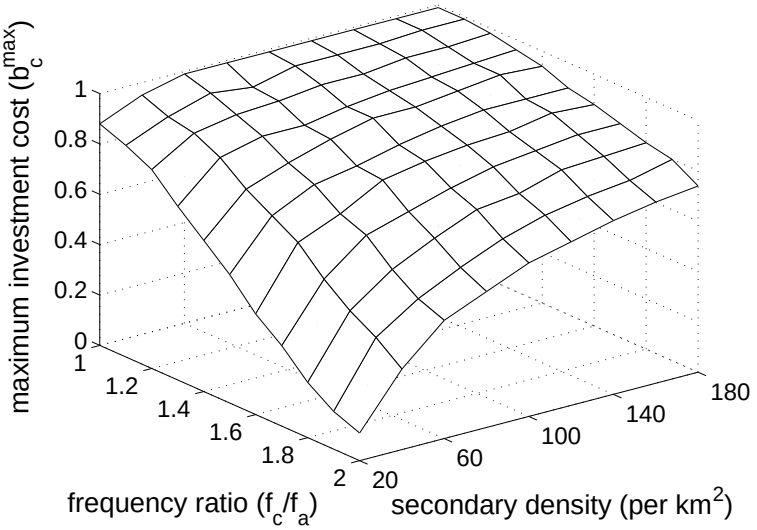


Figure 7.13: Maximum investment cost: WSP c's maximum investment cost b_c for making profit in the market is determined by the channel frequency f_c and secondary density. We assume $f_a = 500$ MHz and $b_a = 1$ for the competitor (i.e., WSP a).

7.6.6 Impact of Spectrum Investment Cost

Our analysis on WSPs' pricing game can provide a practical guideline on WSPs' spectrum investment decisions, such as a purchasing strategy from the spectrum owners (e.g., via auction) in the spectrum plane, as shown in Fig. 7.1. Let us consider a spectrum market where WSP a operates with a channel at frequency $f_a = 500$ MHz, which is obtained at cost $b_a = 1$. Then, WSP c ponders whether to join the market by purchasing a channel with f_c from legacy spectrum owners at price b_c, which we refer to as *spectrum investment cost*.

Fig. 7.13 shows the maximum investment cost b_c^{max}, beyond which the profit becomes negative, i.e., WSP c cannot make a profit in the market. The maximum investment cost depends on spectrum heterogeneity as well as SU density. The figure indicates that the maximum investment cost b_c^{max} is always lower than $b_a = 1$ due to channel c's inferior spectrum profile, but it approaches b_a as the SUs density increases in the market or channel c has a better spectrum profile, i.e., a lower value of f_c.

7.7 Related Work

The problem of optimal spectrum pricing in spectrum market has been studied extensively, and we discuss some of the work closely related to ours. Niyato *et al.* [113] analyzed spectrum pricing competition in cognitive radio networks with multiple primary service providers. Inaltekin *et al.* [73] considered heterogeneous channel conditions due to nodes' physical distances from the base station in wireless IP networks. Jia *et al.* [76] studied the duopoly wireless spectrum market where two WSPs compete for bandwidth and price to maximize their profit. Duan *et al.* [45] studied WSPs' investment and pricing mechanisms by considering SUs' physical-layer wireless characteristics. In [44], they also studied WSPs' optimal spectrum investment and pricing decisions in cognitive radio networks where spectrum availability dynamically changes due to the unpredictability of PUs' channel usage patterns. Gajić *et al.* [54] studied pricing competition among WSPs via a two-stage multi-leader-follower game. Mutlu *et al.* [111] studied measurement-based on-line pricing for secondary spectrum access and developed a pricing framework for an unknown demand function and call-length durations. However, none of the above studies considered the heterogeneity of a wide range of available spectrum bands in the spectrum market and spectrum sharing among co-located SUs in accessing the leased spectrum resources.

The closest to our study is [80] which considered two CR-specific features: (i) bandwidth (supply) uncertainty due to PUs' activities, and (ii) spatial reuse of wireless spectrum. They studied an interesting market scenario where multiple WSPs compete with each other by jointly optimizing the spectrum price based on time and location-dependent spectrum availability. Such fine-grained coordination, however, might not be suitable for a highly dynamic wireless environment due to its high computation and communication overhead. In contrast, we assume a decentralized DSM where individual spectrum consumers purchase the payoff-maximizing spectrum, just as in a real-world market economy. Spectrum price stabilizes when multiple WSPs competing for market share reach a Nash equilibrium.

7.8 Conclusion

The dynamic spectrum market (DSM) is a promising paradigm to provide economic incentives that facilitate DSA. In this chapter, we identified two key factors in a DSM—spectrum heterogeneity and spectrum sharing among SUs—and studied their impact on price competition among wireless service providers (WSPs) in a three-step approach. We first observed that SUs must share the wireless spectrum in the spatial domain, and established the effect of SU density (spectrum demand) on achievable utility when they are associated with the same WSP. We then derived the SUs' optimal WSP selection strategy that maximizes the utility, for given spectrum profile and leasing prices. Finally, we formulated WSPs' spectrum pricing as a non-cooperative game and identified its Nash equilibrium points. Our analysis demonstrates that spectrum heterogeneity significantly influences WSPs' spectrum pricing, especially in a sparse network. In a dense network, the benefit of a lower-frequency band diminishes due to severe co-channel interference, and thus, spectrum heterogeneity has less impact on spectrum pricing.

In the future, we would like to investigate the impact of spectrum heterogeneity on WSPs' auction strategy in the spectrum plane. It would also be interesting to extend the analytical framework to a DSM with multiple WSPs. Moreover, we plan to study the dependency of an optimal spectrum price on other system parameters, e.g., maximum transmission power.

CHAPTER 8

Conclusions and Future Directions

This thesis focused on improving efficiency and robustness in spectrum management for DSA. Despite the promises of improved spectrum efficiency that can mitigate the imminent spectrum-scarcity problem, several practical challenges must be addressed in order to realize the benefits of DSA. This thesis has first identified the two important and fundamental challenges that hinder the realization of DSA—i.e., PUs' fear of interference from SUs' communications and lack of economic incentive for sharing their spectrum resources. Then, the thesis has presented a comprehensive spectrum management system for CRNs, which provides novel approaches and enabling techniques in various aspects of DSA, including energy-efficiency, attack-tolerance, mobility, and spectrum pricing. In what follows, the main research contributions of the thesis are summarized, and future research directions are discussed.

8.1 Primary Contributions

This thesis makes the following contributions toward efficient and robust spectrum management for CRNs.

- **Efficient Detection of Large-Scale Primary Users**: The detection of large-scale PUs is challenging due to the stringent detection requirements imposed by regulatory entities like the FCC, which cannot be met by a single sensor with one-time sensing. To improve detection performance while minimizing service disruptions by spectrum

sensing, we proposed a joint sensor selection and sensing scheduling framework that exploits the spatio-temporal variations in received primary signal strength by constructing a spatial RSS-profile for an incumbent signal. Specifically, we formulated the problem of sensing scheduling as a sequential hypothesis test and presented an optimal sensor-selection algorithm that minimizes the average sensing overhead. The evaluation results have shown that the proposed sensing algorithms reduce sensing overhead by up to $94\,\%$ for practical scenarios.

- **Secure Detection of Large-Scale Primary Users**: In CRNs, sensors (or SUs) can be compromised by an attacker or malfunctioning due to hardware/software defects, and thus their reports cannot be fully trusted at the fusion center. Such manipulated or erroneous sensing reports can cause either waste of spectrum opportunities or excessive interference to PUs' communications. To remedy this problem, we developed a novel attack-tolerant distributed sensing protocol, called ADSP, that selectively filters out abnormal sensor reports or penalizes them, and thus maintains the accuracy of incumbent detection. ADSP exploits shadow fading correlation in the measured primary signal strength at nearby sensors. To realize this idea, we proposed a sensor clustering method, and designed filters and data-fusion rules based on the correlation analysis of sensing reports. The evaluation results in realistic shadow-fading environments have shown that ADSP can meet the detection requirements, even in the presence of attacks.

- **Efficient Detection of Small-Scale Primary Users**: The detection of small-scale PUs is more challenging than that of large-scale PUs due to their unique features, such as unpredictable channel usage patterns, small signal footprint, and mobility. To meet this challenge, we proposed a sensing algorithm tailored to small-scale PU detection, called DeLOC that iterates between cooperative sensing and location/transmit-power estimation to further improve sensing performance. In particular, we developed a novel spatio-temporal fusion scheme that exploits (i) spatial diversity by cooperative sensing with an optimal fusion range, and (ii) temporal diversity by scheduling a series of sensing stages with an optimal stopping time. Our

evaluation results have shown that DeLOC reduces detection delay significantly while achieving high detection performance.

- **Robust Tracking of Mobile Small-Scale Primary Users**: In CRNs, SUs must be able to accurately and reliably track the location of small-scale mobile PUs to better utilize *spatial* spectrum opportunities, while protecting primary communications. However, it is challenging to accurately track the location of mobile PUs because the tracking process must rely solely on the reported sensing results, which can be easily compromised by malicious sensors (or attackers). To remedy this problem, we presented a framework, called SOLID, for accurate, attack/fault-tolerant tracking of small-scale mobile PUs, such as wireless microphones. Our evaluation results have shown that SOLID lowers the localization error significantly regardless of the presence or absence of attackers.

- **Opportunistic Spectrum Access for Mobile Cognitive Radios**: It is important for CRNs to incorporate mobility of SUs to fully realize the benefits of DSA as various future mobile devices are expected to incorporate CR functionality. In this thesis, we identified and addressed three fundamental challenges posed by mobile SUs that do not exist in the case of stationary CRNs. We showed via analysis that the channel availability experienced by a mobile SU can be modeled as a two-state Markov chain, and introduced guard distance in the space domain for efficient spectrum reuse. To further enable efficient spectrum sharing, we derived an optimal distributed channel-selection (access) strategy that maximizes the secondary network throughput performance. The evaluation results have shown that the proposed spectrum sensing and distributed channel access schemes significantly improve network throughput and fairness, while reducing the SUs' energy consumption for spectrum sensing.

- **Optimal Spectrum Pricing in Dynamic Spectrum Market**: The dynamic spectrum market (DSM) is a key economic means for realizing the opportunistic spectrum access that will mitigate the anticipated spectrum-scarcity problem. In DSM, determining the optimal spectrum leasing price is an important, challenging problem that requires a comprehensive understanding of market participants' interests

and interactions. In this thesis, we studied spectrum pricing competition in a DSM characterized by a duopoly, where two wireless service providers (WSPs) lease spectrum access rights and secondary users (SUs) purchase the spectrum use to maximize their utility. We identified two essential, but previously-overlooked, properties of DSM: (i) heterogeneous spectrum resources at WSPs and (ii) spectrum sharing among SUs. The proposed analytical framework has important implications for the impact of spectrum heterogeneity in a real-world DSM, and provides practical guidelines for WSPs' pricing strategies.

8.2 Future Research Directions

This section describes several additional research issues that are related to the extension and application of CR technologies in future wireless networks.

- **Mobile Wireless Systems with Cognitive Radio Capabilities**: CR technologies can be widely used in future wireless systems to further improve wireless spectrum efficiency. One interesting direction is to apply the CR sensing technologies in heterogeneous wireless networks. For example, due to the pervasive usage of mobile portable devices, such as smartphones, mobile data usage over cellular networks has increased dramatically over the past several years. Therefore, off-loading such mobile data traffic to Wi-Fi networks is a viable option for wireless service providers to provide high quality wireless Internet access without the need for purchasing expensive bandwidth resources. Advanced spectrum sensing techniques, such as sensing scheduling algorithms, can be applied in the design of energy-efficient mechanisms for Wi-Fi opportunity discovery for mobile smartphone users.

- **Leveraging Spectrum Heterogeneity in Future Wireless Networks**: With the current trend of spectrum deregulation, we envision that a wide range of spectrum resources will be available in future wireless environments. Regarding spectrum heterogeneity, in Chapter 7, we studied the problem of optimal spectrum pricing in a DSM, in which WSPs compete for a market share with heterogeneous spectrum re-

sources. Such spectrum heterogeneity can be leveraged for improving network performance in various aspects, such as resource allocation (e.g., channel and transmit-power), protocol design, and network optimization. For example, the IEEE 802.11af Working Group aims at the operation of Wi-Fi networks on TV white spaces. However, the current MAC-layer protocol design, e.g., CSMA/CA, is tailored to operate optimally on high frequency bands, i.e., 2.4 GHz and 5.7 GHz ISM bands. Therefore, it is of practical importance to study the impact of spectrum heterogeneity on the performance of MAC and upper layer protocol stacks and incorporate them into the design of future wireless systems.

APPENDICES

APPENDIX A

Proof of Proposition 4.1

To minimize the average number of sensing rounds N in Eq. (4.8), we need to find the data-fusion range that maximizes the expected test statistic, i.e., $\mathbb{E}[\lambda \,|\, \mathcal{H}_k]$. Note that $\mathbb{E}[\Lambda_N]$ is a function of desired false-alarm and mis-detection probability, and does not depend on the fusion range, as indicated in Eq. (4.10). Let $\mathcal{J}(R_f) \triangleq \mathbb{E}[\lambda \,|\, \mathcal{H}_1]$ in Eq. (4.19). Then, we have:

$$\begin{aligned}\mathcal{J}(R_f) &\triangleq \mathbb{E}[\lambda \,|\, \mathcal{H}_1] \\ &= \frac{1}{2}(\phi_0^2 - \phi_1^2)\rho\pi R_f^2 + \frac{\phi_1 - \phi_0}{\sigma_n}(P_o\, e^{\frac{1}{2}\sigma^2} m_1 + N_o \rho\pi R_f^2) \\ &= a_1 R_f^{2-\alpha} + a_2 R_f^2 + a_3, \end{aligned} \tag{A.1}$$

where

$$a_1 = \frac{2\,(\phi_1 - \phi_0) P_o\, e^{\frac{1}{2}\sigma^2} \rho\pi d_o^\alpha}{\sigma_n(2-\alpha)}, \tag{A.2}$$

$$a_2 = \frac{1}{2}(\phi_0^2 - \phi_1^2)\rho\pi + \frac{(\phi_1 - \phi_0) N_o \rho\pi}{\sigma_n}, \tag{A.3}$$

$$a_3 = \frac{2\,(\phi_1 - \phi_o) P_o\, e^{\frac{1}{2}\sigma^2} \rho\pi d_o^\alpha}{\sigma_n(\alpha-2)\epsilon^{\alpha-2}}. \tag{A.4}$$

To find an optimal value of R_f, we need to show the concavity of $\mathcal{J}(R_f)$ w.r.t. R_f. The

first order derivative of $\mathcal{J}(R_f)$ is given as:

$$\frac{\partial \mathcal{J}(R_f)}{\partial R_f} = a_1(2-\alpha)R_f^{1-\alpha} + 2a_2 R_f. \tag{A.5}$$

Then, the second order derivative is given as:

$$\begin{aligned}
&\frac{\partial^2 \mathcal{J}(R_f)}{\partial R_f^2} = a_1(2-\alpha)(1-\alpha)R_f^{-\alpha} + 2a_2 \\
&= \frac{4(\phi_1-\phi_0)\rho\pi}{\sigma_n}\Big((1-\alpha)P_o\, e^{\frac{1}{2}\sigma^2}\Big(\frac{d_o}{R_f}\Big)^{\alpha} + N_o - \frac{(\phi_1+\phi_0)\sigma_n}{2}\Big) \\
&\approx 2\rho\pi\big(1-2\alpha\big)\Big(\frac{P_R}{\sigma_n}\Big)^2.
\end{aligned} \tag{A.6}$$

In practice, $\alpha > \frac{1}{2}$, so it is easy to show that $\frac{\partial^2 \mathcal{J}(R_f)}{\partial R_f^2} < 0$ and hence $\mathcal{J}(R_f)$ is concave. Therefore, the optimal fusion range R_f can be derived from Eq. (A.5) as:

$$R_f^* = \arg\max_{R_f} \mathcal{J}(R_f) = R_f\Big|_{\frac{\partial \mathcal{J}(R_f)}{\partial R_f}=0} = \left(\frac{a_1(\alpha-2)}{2a_2}\right)^{\frac{1}{\alpha}}. \tag{A.7}$$

Thus, the proposition holds. $\square$

APPENDIX B

Proof of Lemma 7.1

Since the interfering nodes are distributed uniformly outside the interference range S, the nodes inside a differential element area $r\,\mathrm{d}r\,\mathrm{d}\theta$ generates the following amount of interference:

$$\mathrm{d}I_{out}(\rho, c) = P_o\, g(r)\, \rho_c\, r\, \mathrm{d}r\, \mathrm{d}\theta = P_o \left(\frac{c_o}{f_c}\right)^{\alpha} r^{1-\alpha}\, \rho_c\, \mathrm{d}r\, \mathrm{d}\theta. \tag{B.1}$$

Then, based on Eq. (B.1), the total interference cause by nodes located outside the interference region S, i.e., unit disk of radius R_I^c centered at the receiver, can be calculated as:

$$\begin{aligned} I_{out}(\rho, c) &= \int_{\text{outside the disk}} \mathrm{d}I_{out}(\rho, c) \\ &= \int_0^{2\pi} \int_{A-S} P_o \left(\frac{c_o}{f_c}\right)^{\alpha} r^{1-\alpha}\, \rho_c\, \mathrm{d}r\, \mathrm{d}\theta \\ &\overset{(a)}{\leq} \int_0^{2\pi} \int_{R_I^c}^{+\infty} P_o \left(\frac{c_o}{f_c}\right)^{\alpha} r^{1-\alpha}\, \rho_c\, \mathrm{d}r\, \mathrm{d}\theta \\ &= P_o \left(\frac{c_o}{f_c}\right)^{\alpha} \rho_c \int_0^{2\pi} \frac{r^{2-\alpha}}{2-\alpha}\bigg|_{R_I^c}^{+\infty} \mathrm{d}\theta \\ &= 2\pi\, P_o \left(\frac{c_o}{f_c}\right)^{\alpha} \frac{\rho_c\, (R_I^c)^{2-\alpha}}{\alpha - 2}. \end{aligned} \tag{B.2}$$

Eq. (B.2) indicates that I_{out} can be upper-bounded by a constant value shown in the last line. Note that the inequality (a) is based on the fact that the amount of interference in a finite network is lower than or equal to that of the network with infinite size.

APPENDIX C

Proof of Proposition 7.1

We prove that the *occupancy measure* $\Pi_c(t)$ converges to a deterministic process in the continuous-time domain, i.e., the mean-field limit of the Markov process. We first rescale the original discrete-time Markov process, and define a new continuous-time process $\Pi'_c(t)$, such that $\Pi'_c(\frac{t}{N}) = \Pi_c(t)$.

In [16], Benaïm *et al.* established the sufficient conditions for the convergence of an N-dimensional Markov process towards the mean field limit, for a large N. Specifically, five conditions need to be satisfied for the convergence. Specific to our model, the conditions can be interpreted as follows.

C1 The system resource does not scale with N.

C2 Intensity vanishes at rate $o(N)$.[1]

C3 Second moment of intensity must be bounded.

C4 Transition matrix of the resource is a smooth function of $\frac{1}{N}$ and the mean field limit.

C5 Intensity is a smooth function of $\frac{1}{N}$ and the mean field limit.

Note that C1 and C4 can be easily satisfied since no explicit channel resource is defined within our model. In fact, the links' state transitions implicitly change the channel status. So, we proceed to prove that C2, C3, and C5 are satisfied in our model.

[1] In this paper, $f(x) = o(g(x))$ if $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 0$.

Proof for C2. Let N_n be the number of nodes in the network. Each node can be a data source, following a Poisson distribution with rate λ. Consider a time slot $[t, t+\delta]$ where $\delta = \frac{\Delta t}{N}$. Define the number of arrivals within $[t, t+\delta]$ as $\widetilde{N}(t, t+\delta)$, which is the number of state transitions of each source node, within δ. According to the Poisson property,

$$Pr\big(\widetilde{N}(t, t+\delta) = 1\big) = \lambda\, \delta\, e^{-\lambda\delta}. \tag{C.1}$$

Therefore, the intensity—i.e., the total number of state transitions in the system—is:

$$\begin{aligned} N_n \cdot Pr(\widetilde{N}(t, t+\delta) = 1) &= N_n \cdot \lambda\, \delta\, e^{-\lambda\delta} \\ &= N_n \cdot \lambda \frac{\Delta t}{N} e^{-\lambda \frac{\Delta t}{N}} \\ &\overset{(a)}{\leq} \sqrt{\frac{AN}{\pi R^2}} \cdot \lambda \frac{\Delta t}{N} e^{-\lambda \frac{\Delta t}{N}} \\ &= o(N), \end{aligned} \tag{C.2}$$

where the inequality (a) is due to $N \leq N_n^2\big(\frac{\pi R^2}{A}\big)$ where R is the transmission range of SUs.

Proof for C3. Let $W^N(t)$ denote the number of source nodes that make a transition at time slot t. The number of links that make a transition is bounded by the number of source nodes, hence $W^N(t) = N_n$.

$$\mathbb{E}\big[W^N(t)^2\big] = N_n^2 \leq N^2 \frac{A}{N\pi R^2} = N^2 \epsilon(N)^2. \tag{C.3}$$

Proof for C5. Given the occupancy measure $\Pi_c(t)$ at time slot t, the drift can be obtained directly from Eq. (7.12). Specifically, given that $N_c(t) = m$, then the drift is given as:

$$\mathbb{E}\big[N_c(t+1) - N_c(t) \,|\, N_c(t) = m\big] = N\,\lambda\, P_c - m\,\lambda,$$

which is obviously a smooth function of $\frac{1}{N}$.

BIBLIOGRAPHY

[1] IEEE 802.22 Working Group on Wireless Regional Area Networks, http://www.ieee802.org/22/.

[2] IEEE 802.22 Working Group on Wireless Local Area Networks, http://www.ieee802.org/11/.

[3] USRP: Universal Software Radio Peripheral, http://www.ettus.com.

[4] http://en.wikipedia.org/wiki/Mobile_virtual_network_operator/.

[5] http://www.specex.com.

[6] GNU Software Radio Project, http://www.gnu.org/software/gnuradio/.

[7] A. Dejonghe *et al.* Versatile spectrum sensing on mobile devices? In *Proc. IEEE DySPAN*, April 2010.

[8] Ian F. Akyildiz, Brandon F. Lo, and Ravikumar Balakrishnan. Cooperative spectrum sensing in cognitive radio networks: A survey. *Elsevier Physical Communications*, 4(1):40–62, March 2011.

[9] Siavash M. Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, October 1998.

[10] Albert Algans, Klaus Ingemann Pedersen, and Preben Elgaard Mogensen. Experimental analysis of the joint statistical properties of azimuth spread, delay spread, and shadow fading. *IEEE Journal on Selected Areas in Communications*, 20(3):523–531, April 2002.

[11] S. Anand, Z. Jin, and K. P. Subbalakshmi. An analytical model for primary user emulation attacks in cognitive radio networks. In *Proc. IEEE DySPAN*, October 2008.

[12] Moataz M.H. El Ayadi, Mohamed S. Kamel, and Fakhri Karray. Toward a tight upper bound for the error probability of the binary Gaussian classification problem. *Pattern Recognition*, 41(6):2120–2132, June 2008.

[13] Aline Baggio and Koen Langendoen. Monte-Carlo localization for mobile wireless sensor networks. In *Proc. MSN*, December 2006.

[14] Paramvir Bahl, Ranveer Chandra, Thomas Moscibroda, Rohan Murty, and Matt Welsh. White space networking with Wi-Fi like connectivity. In *Proc. ACM SIGCOMM*, August 2009.

[15] Paramvir Bahl and Venkata N. Padmanabhan. RADAR: An In-Building RF-based User Location and Tracking System. In *Proc. IEEE INFOCOM*, March 2000.

[16] Michel Benaim and Jean-Yves Le Boudec. A class of mean field interaction models for computer and communication systems. *Performance Evaluation*, 65(11-12):823–838, April 2008.

[17] Christian Bettstetter. On the minimum node degree and connectivity of a wireless multihop network. In *Proc. ACM MobiHoc*, June 2002.

[18] Gregory J. Buchwald, Stephen L. Kuffner, Lawrence M. Ecklund, Monique Brown, and Jr Adgar H. Callaway. The design and pperation of the IEEE 802.22.1 disabling beacon for the protection of TV whitespace incumbents. In *Proc. IEEE DySPAN*, October 2008.

[19] Milind M. Buddhikot. Understanding dynamic spectrum access: Models, taxonomy and challenges. In *Proc. IEEE DySPAN*, April 2007.

[20] Danijela Cabric, Shridhar Mubaraq Mishra, and Robert W. Brodersen. Implementation issues in spectrum sensing for cognitive radios. In *Asilomar conference on Signals, Systems and Computers*, November 2004.

[21] Winston Caldwell. Geolocation with database requirement summary. IEEE 802.22-06/0140r0, July 2006.

[22] Dave Cavalcanti and Carlos Cordeiro. Proposed resolution for comments related to the superframe and sensing. IEEE 802.22-07/0176r0, April 2007.

[23] Nicholas B. Chang and Mingyan Liu. Optimal channel probing and transmission scheduling in a multichannel system. In *Proc. ACM MobiCom*, September 2007.

[24] Chang Joo Kim *et al.* Fractional BW usage for WRAN systems. IEEE 802.22-06/0117r0, July 2006.

[25] John M. Chapin and William H. Lehr. The path to market success for dynamic spectrum access technology. *IEEE Communications Magazine*, 45(5):96–103, May 2007.

[26] Hou-Shin Chen, Wen Gao, and David G. Daut. Spectrum sensing for wireless microphone signals. In *Proc. IEEE SECON*, June 2008.

[27] Qing Chen, Fei Ye, and Zhisheng Niu. Graph theoretical analysis of opportunistic scheduling policy for wireless ad hoc networks. In *Proc. IEEE Globecom*, November 2007.

[28] Ruiliang Chen, Jung-Min Park, and Kaigui Bian. Robust distributed spectrum sensing in cognitive radio networks. In *Proc. IEEE INFOCOM*, April 2008.

[29] Ruiliang Chen, Jung-Min Park, and Jeffrey H. Reed. Defense against primary user emulation attacks in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 26(1):25–37, January 2008.

[30] Wei-Peng Chen, Jennifer C. Hou, and Lui Sha. Dynamic clustering for acoustic target tracking in wireless sensor networks. *IEEE Transactions on Mobile Computing*, 3(3):258–271, July–September 2004.

[31] Kae Won Choi, Wha Sook Jeon, and Dong Geun Jeong. Sequential detection of cyclostationary signal for cognitive radio systems. *IEEE Transactions on Wireless Communications*, 8(9):4480–4485, September 2009.

[32] Youngkyu Choi and Sunghyun Choi. Service charge and energy-aware vertical handoff in integrated IEEE 802.16e/802.11 networks. In *Proc. IEEE INFOCOM*, May 2007.

[33] Gerald Chouinard. WRAN reference model. IEEE 802.22-04/0002r12, September 2005.

[34] Gerald Chouinard. Wireless microphone Sensing. IEEE 802.22-07/0530r1, November 2007.

[35] Gerald Chouinard. Use of collaborative sensing to reduce false positive results. IEEE 802.22-08/0118r0, May 2008.

[36] Gerald Chouinard. Sensing performance with the 802.22.1 wireless microphone beacon. IEEE 802.22-09/0068r1, March 2009.

[37] Kaushik R. Chowdhury, Marco Di Felice, and Ian F. Akyildiz. TP-CRAHN: A transport protocol for cognitive radio ad-hoc networks. In *Proc. IEEE INFOCOM*, April 2009.

[38] Carlos Cordeiro, Kiran Challapali, Dagnachew Birru, and Sai Shankar. IEEE 802.22: An introduction to the first wireless standard based on cognitive radio. *Journal of Communications*, 1(1):38–47, April 2006.

[39] Carlos Cordeiro, Kiran Challapali, and Monisha Ghosh. Cognitive PHY and MAC layers for dynamic spectrum access and sharing of TV bands. In *Proc. ACM TAPAS*, August 2006.

[40] Fadel F. Digham, Mohamed-Slim Alouini, and Marvin K. Simon. On the energy detection of unknown signals over fading channels. In *Proc. IEEE ICC*, May 2003.

[41] Fadel F. Digham, Mohamed-Slim Alouini, and Marvin K. Simon. On the energy detection of unknown signals over fading channels. *IEEE Transactionson on Communications*, 55(1):21–24, January 2007.

[42] Min Ding and Xiuzhen Cheng. Fault tolerant target tracking in sensor networks. In *Proc. ACM MobiHoc*, May 2009.

[43] Wenliang Du, Jing Deng, Yunghsiang S. Han, and Pramod K. Varshney. A witness-based approach for data fusion assurance in wireless sensor networks. In *Proc. IEEE Globecom*, December 2003.

[44] Lingjie Duan, Jianwei Huang, and Biying Shou. Cognitive mobile virtual network operator: Investment and pricing with supply uncertainty. In *Proc. IEEE INFOCOM*, March 2010.

[45] Lingjie Duan, Jianwei Huang, and Biying Shou. Competition with dynamic spectrum leasing. In *Proc. IEEE DySPAN*, April 2010.

[46] Vinko Erceg, Larry J. Greenstein, Sony Y. Tjandra, Seth R. Parkoff, Ajay Gupta, Boris Kulic, Arthur A. Julius, and Renee Bianchi. An empirically based path loss model for wireless channels in suburban environments. *IEEE Journal on Selected Areas in Communications*, 17(7):1205–1211, July 1999.

[47] FCC. Facilitating opportunisties for flexible, efficient, and reliable spectrum use employing spectrum agile radio technologies. Rep. ET Docket No. 03-108, December 2003.

[48] FCC. Notice of proposed rulemaking and order. FCC 08-188, August 2008.

[49] FCC. Second Report and Order. FCC 08-260, Nov 2008.

[50] FCC. Second Memorandum Opinion and Order. FCC 10-174, September 2010.

[51] A. Fehske, J. Gaeddert, and J. H. Reed. A new approach to signal classification using spectral correlation and neural networks. In *Proc. IEEE DySPAN*, November 2005.

[52] Lawrence F. Fenton. The sum of log-normal probability distributions in scatter transmission systems. *IRE Transactions on Communications Systems*, 8(1):57–67, March 1960.

[53] Ingo Forkel, Marc Schinnenburg, and Markus Ang. Generation of two-dimensional correlated shadowing for mobile radio network simulation. In *Proc. WPMC*, September 2004.

[54] Vojislav Gajić, Jianwei Huang, and Bixio Rimoldi. Competition of wireless providers for atomic users: Equilibrium and social optimality. In *Proc. Allerton*, September 2009.

[55] Amir Ghasemi and Elvino S. Sousa. Collaborative spectrum sensing for opportunistic access in fading environments. In *Proc. IEEE DySPAN*, November 2005.

[56] Amir Ghasemi and Elvino S. Sousa. Asymptotic performance of collaborative spectrum sensing under correlated log-normal shadowing. *IEEE Communications Letter*, 11(1):34–36, January 2007.

[57] Amir Ghasemi and Elvino S. Sousa. Opportunistic spectrum access in fading channels through collaborative sensing. *Journal of Cummunications*, 2(2):71–82, March 2007.

[58] Amir Ghasemi and Elvino S. Sousa. Interference aggregation in spectrum-sensing cognitive wireless networks. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):41–56, February 2008.

[59] Andrea Goldsmith. *Wireless Communications*. Cambridge University Press, 2005.

[60] M. Gudmundson. Correlation model for shadow fading in mobile radio systems. *Electronic Letters*, 27(23):2145–2146, November 1991.

[61] Chao Gui and Prasant Mohapatra. Power conservation and quality of surveillance in target tracking sensor networks. In *Proc. ACM MobiCom*, September 2004.

[62] Piyush Gupta and P. R. Kumar. The capacity of wireless networks. *Transactions on Information Theory*, 46(2):388–404, March 2000.

[63] David Gurney, Greg Buchwald, Larry Ecklund, Steve Kuffner, and John Grosspietsch. Geo-location database techniques for incumbent protection in the TV white space. In *Proc. IEEE DySPAN*, October 2008.

[64] Muhammad Fainan Hanif, Mansoor Shafi, Peter J. Smith, and P. Dmochowski. Interference and deployment issues for cognitive radio systems in shadowing environments. In *Proc. IEEE ICC*, June 2009.

[65] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer, 2001.

[66] Simon Haykin. *Adaptive Filter Theory*. 2nd ed. Prentice Hall, 1991.

[67] Anh Tuan Hoang and Ying-Chang Liang. Adaptive scheduling of spectrum sensing periods in cognitive radio networks. In *Proc. IEEE Globecom*, November 2007.

[68] Anh Tuan Hoang, Ying-Chang Liang, David Tung Chong Wong, and Rui Zhang. Opportunistic spectrum access for energy-constrained cognitive radios. *IEEE Transactions on Wireless Communications*, 8(3):1206–1211, March 2009.

[69] Xuemin Hong, Cheng-Xiang Wang, and John Thompson. Interference modeling of cognitive radio networks. In *Proc. IEEE VTC-spring*, May 2008.

[70] Lingxuan Hu and David Evans. Localization for mobile sensor networks. In *Proc. ACM MobiCom*, September 2004.

[71] Senhua Huang, Xin Liu, and Zhi Ding. Optimal Sensing-Transmission Structure for Dynamic Spectrum Access. In *Proc. IEEE INFOCOM*, April 2009.

[72] Joseph Mitola III and Jr. Gerald Q. Maguire. Cognitive radio: Making software radios more personal. *IEEE Personal Communications*, 6(4):13–18, August 1999.

[73] Hazar Inaltekin, Tom Wexler, and Stephen B. Wicker. A duopoly pricing game for wireless IP services. In *Proc. IEEE SECON*, June 2007.

[74] ITU-R Recommendation P.1546-3. Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3000 MHz, 2007.

[75] Rajendra K. Jain, Dah-Ming W. Chiu, and William R. Hawe. A quantitative measure of fairness and discrimination for resource allocation in shared computer systems. Technical Report DEC-TR-301, Digital Equipment Corporation, Sep 1984.

[76] Juncheng Jia and Qian Zhang. Competitions and dynamics of duopoly wireless service providers in dynamic spectrum market. In *Proc. ACM MobiHoc*, May 2008.

[77] Jianfeng Wang *et al.* First cognitive radio networking standard for personal/portable devices in TV white spaces. In *Proc. IEEE DySPAN*, April 2010.

[78] John Benko *et al.* Draft PHY/MAC specification for IEEE 802.22. IEEE 802.22-06/0069r1, May 2006.

[79] Praveen Kaligineedi, Majid Khabbazian, and Vijay K. Bharava. Secure cooperative sensing techniques for cognitive radio systems. In *Proc. IEEE ICC*, May 2008.

[80] Gaurav S. Kasbekar and Saswati Sarkar. Spectrum pricing games with bandwidth uncertainty and spatial reuse in cognitive radio networks. In *Proc. ACM MobiHoc*, September 2010.

[81] Hyoil Kim and Kang G. Shin. In-band spectrum sensing in cognitive radio networks: Energy detection or feature detection? In *Proc. ACM MobiCom*, September 2008.

[82] Klein S. Gilhousen *et al.* On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40(2):303–312, May 1991.

[83] Solomon Kullback. *Information Theory and Statistics*. John Wiley and Sons, 1959.

[84] Kun Tan *et al.* Sora: High performance software radio using general purpose multi-core processors. In *Proc. USENIX NSDI*, April 2009.

[85] Lifeng Lai, Yijia Fan, and H. Vincent Poor. Quickest detection in cognitive radio: A sequential change detection framework. In *Proc. IEEE Globecom*, Nov 2008.

[86] Zhongding Lei and Francois Chin. A reliable and power efficient beacon structure for cognitive radio systems. In *Proc. IEEE ICC*, May 2008.

[87] Khaled Ben Letaief and Wei Zhang. Cooperative communications for cognitive radio networks. *IEEE Proceedings*, 97(5):878–893, May 2009.

[88] Ying-Chang Liang, Yonghong Zeng, Edward C. Y. Peh, and Anh Tuan Hoang. Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communications*, 7:1326–1337, April 2008.

[89] Benyuan Liu, Peter Brass, Olivier Dousse, Philippe Nain, and Don Towsley. Mobility improves coverage of sensor networks. In *Proc. ACM MobiHoc*, May 2005.

[90] Fang Liu, Xiuzhen Cheng, and Dechang Chen. Insider attacker detection in wireless sensor networks. In *Proc. IEEE INFOCOM*, May 2007.

[91] Song Liu, Yingying Chen, Wade Trappe, and Larry J. Greenstein. ALDO: An anomaly detection framework for dynamic sptectrum access networks. In *Proc. IEEE INFOCOM*, April 2009.

[92] Song Liu, Yingying Chen, Wade Trappe, and Larry J. Greenstein. Non-interactive localization of cognitive radios based on dynamic signal strength mapping. In *Proc. IEEE/IFIP WONS*, February 2009.

[93] Yao Liu, Peng Ning, and Huaiyu Dai. Authenticating primary users' signals in cognitive radio networks via integrated cryptographic and wireless link signatures. In *Proc. IEEE Symposium on Secureity and Privacy*, May 2010.

[94] J. Lundén, V. Koivunen, A. Huttunen, and H. V. Poor. Collaborative cyclostationary spectrum sensing for cognitive radio systems. *IEEE Transactions on Signal Processing*, 57(11):4182–4195, November 2009.

[95] Jun Ma, Guodong Zhao, and Ye (Geoffrey) Li. Soft combination and detection for cooperative spectrum sensing in cognitive radio networks. *IEEE Transactions on Wireless Communications*, 7(11):4502–4507, November 2008.

[96] Zhiyao Ma, Wei Chen, Khaled B. Letaief, and Zhigang Cao. A semi range-based iterative localization algorithm for cognitive radio networks. *IEEE Transactions on Vehicular Technology*, 59:704–717, February 2010.

[97] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory*. Oxford University Press, 1995.

[98] Mark A. McHenry. NSF spectrum occupancy measurements project summary. Shared Spectrum Company Report, August 2005.

[99] R. Menon, R. M. Buehrer, and J. H. Reed. On the impact of dynamic spectrum sharing techniques on legacy radio systems. *IEEE Transactions on Wireless Communications*, 7(11):4198–4207, November 2008.

[100] Alexander W. Min and Kang G. Shin. An Optimal Sensing Framework Based on Spatial RSS-profile in Cognitive Radio Networks. In *Proc. IEEE SECON*, June 2009.

[101] Alexander W. Min and Kang G. Shin. Impact of mobility on spectrum sensing in cognitive radio networks. In *Proc. ACM CoRoNet*, September 2009.

[102] Alexander W. Min and Kang G. Shin. On sensing-access tradeoff in cognitive radio networks. In *Proc. IEEE DySPAN*, April 2010.

[103] Alexander W. Min, Kang G. Shin, and Xin Hu. Attack-tolerant distributed sensing for dynamic spectrum access networks. In *Proc. IEEE ICNP*, October 2009.

[104] Alexander W. Min, Xinyu Zhang, and Kang G. Shin. Spatio-temporal fusion for small-scale primary detection in cognitive radio networks. In *Proc. IEEE INFOCOM*, March 2010.

[105] Alexander W. Min, Xinyu Zhang, and Kang G. Shin. Detection of small-scale primary users in cognitive radio networks. *IEEE Journal on Selected Areas in Communications*, 29(2):349–361, February 2011.

[106] Shridhar Mubaraq Mishra, Anant Sahai, and Robert W. Brodersen. Cooperative sensing among cognitive radios. In *Proc. IEEE ICC*, June 2006.

[107] Shridhar Mubaraq Mishra, Rahul Tandra, and Anant Sahai. Coexistence with primary users of different scales. In *Proc. IEEE DySPAN*, April 2007.

[108] Apurva Mody, Tim Brown, Amita Sethi, Ranga Reddy, Thomas Kiernan, and Gerald Chouinard. Collaborative sensing for security. IEEE 802.22-08/0301r011, December 2008.

[109] Torsten Muetze, Patrick Stuedi, Fabian Kuhn, and Gustavo Alonso. Understanding radio irregularity in wireless networks. In *Proc. IEEE SECON*, June 2008.

[110] Rohan Murty, Ranveer Chandra, Thomas Moscibroda, and Paramvir Bahl. SenseLess: A database-driven white spaces networks. Technical Report MSR-TR-2010-127, MSR, September 2010.

[111] Huseyin Mutlu, Murat Alanyali, and David Starobinski. On-line pricing of secondary spectrum access with unknown demand function and call length distribution. In *Proc. IEEE INFOCOM*, March 2008.

[112] Jari Nieminen, Riku Jäntti, and Lijun Qian. Primary user detection in distributed cognitive radio networks under timing inaccuracy. In *Proc. IEEE DySPAN*, April 2010.

[113] Dusit Niyato and Ekram Hossain. Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion. *IEEE Journal on Selected Areas in Communications*, 26(1):192–202, January 2008.

[114] Dusit Niyato, Ekram Hossain, and Zhu Han. Dynamics of multiple-seller and multiple-buyer spectrum trading in cognitive radio networks: A game-theoretic

modeling approach. *IEEE Transactions on Mobile Computing*, 8(8):1009–1022, August 2009.

[115] OET Report FCC/OET 08-TR-1005. Evaluation of the performance of prototype TV-band white space devices phase II. Technical Research Branch Laboratory Division, Office of Engineering and Technology, Federal Communications Commission, October 2008.

[116] Taejoon Park and Kang G. Shin. Attack-tolerant localization via iterative verification of locations in sensor networks. *ACM Transactions on Embedded Computing Systems*, 8(1):2:1–2:24, December 2008.

[117] Neal Patwari and Piyush Agrawal. Effects of correlated shadowing: Connectivity, localization, and RF tomography. In *Proc. IEEE IPSN*, April 2008.

[118] Edgar Reihl. Wireless Microphone Characteristics. IEEE 802.22-06/0070r0, May 2006.

[119] Janne Riihijärvi, Petri Mähönen, Matthias Wellens, and Martin Gordziel. Characterization and modelling of spectrum for dynamic spectrum access with spatial statistics and random fields. In *Proc. IEEE PIMRC*, September 2008.

[120] William Rose. Enhanced protection for low power licensed devices operating in TV broadcast bands. IEEE 802.22-06/0073r2, May 2006.

[121] Masoomeh Rudafshani and Suprakash Datta. Localization in wireless sensor networks. In *Proc. IEEE IPSN*, April 2007.

[122] Walid Saad, Zhu Han, Merouane Debbah, Are Hjorungnes, and Tamer Basar. Coalition games for distributed collaborative spectrum sensing in cognitive radio networks. In *Proc. IEEE INFOCOM*, April 2009.

[123] Aysel Safak. Statistical analysis of the power sum of multiple correlated log-normal components. *IEEE Transactions on Vehicular Technology*, 42(1):58–61, February 1993.

[124] M. Sawada, Daniel Cossette, Barry Wellar, and Tolga Kurt. Analysis of the urban/rural broadband divide in Canada: Using GIS in planning terrestrial wireless deployment. *Government Information Quarterly*, 23(3-4):454–479, September 2006.

[125] Yngve Selén, Hugo Tullberg, and Jonas Kronander. Sensor selection for cooperative spectrum sensing. In *Proc. IEEE DySPAN*, October 2008.

[126] Senhua Huang and Xin Liu and Zhi Ding. Distributed power control for cognitive user access based on primary link control feedback. In *Proc. IEEE INFOCOM*, March 2010.

[127] Stephen Shellhammer, Sai Shankar, Rahul Tandra, and James Tomcik. Performance of power detector sensors of DTV signals in IEEE 802.22 WRANs. In *Proc. ACM TAPAS*, August 2006.

[128] Stephen J. Shellhammer. Spectrum Sensing in IEEE 802.22. In *IAPR Workshop on Cognitive Information Processing*, June 2008.

[129] Steve Shellhammer. Performance of the power detector. IEEE 802.22-06/0075r0, May 2006.

[130] Steve Shellhammer and Gerald Chouinard. Spectrum sensing requirements summary. IEEE 802.22-06/0089r5, July 2006.

[131] Steve Shellhammer and Rahul Tandra. An evaluation of DTV pilot power detection. IEEE 802.22-06/0188r0, September 2006.

[132] Steve Shellhammer and Rahul Tandra. Numerical spectrum sensing requirements. IEEE 802.22-06/0088r0, June 2006.

[133] Steve Shellhammer and Rahul Tandra. Performance of the power detector with noise uncertainty. IEEE 802.22-06/0134r0, July 2006.

[134] Tao Shu and Marwan Krunz. Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors. In *Proc. ACM MobiCom*, September 2009.

[135] Bernard Sklar. Rayleigh fading channel in mobile digital communication systems part I: Characterization. *IEEE Communications Magazine*, 35(7):90–100, July 2002.

[136] Adam Smith, Hari Balakrishnan, Michel Goraczko, and Nissanka Priyantha. Tracking moving devices with the cricket location system. In *Proc. ACM MobiSys*, June 2004.

[137] Steve Shellhammer et al. Spectrum sensing simulation model. IEEE 802.22-06/0028r10, September 2006.

[138] Carl R. Stevenson, Gerald Chouinard, Zhongding Lei, Wendong Hu, Stephen J. Shellhammer, and Winston Caldwell. IEEE 802.22: The first cognitive radio wireless regional area network standard. *IEEE Communications Magazine*, 47(1):130–138, January 2009.

[139] Carl R. Stevenson, Carlos Cordeiro, Eli Sofer, and Gerald Chouinard. RAN Requirements. IEEE 802.22-05/0007r46, Sep 2005.

[140] Dietrich Stoyan, Wilfrid S. Kendall, and Joseph Mecke. *Stochastic Geometry and Its Applications*. Chichester: John Wiley & Sons, 1986.

[141] A. Taherpour, Y. Norouzi, M. Masiri-Kenari, A. Jamshidi, and Z. Zeinalpour Yazdi. Asymptotically optimum detection of primary user in cognitive radio networks. *IET Communications*, 1(6):1138–1145, December 2007.

[142] Rahul Tandra and Anant Sahai. SNR walls for signal detection. *IEEE Journal on Selected Areas in Communications*, 2(1):4–17, February 2008.

[143] Jayakrishnan Unnikrishnan and Venugopal V. Veeravalli. Cooperative sensing for primary detection in cognitive radio. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):18–27, February 2008.

[144] E. Visotsky, S. Kuffner, and R. Peterson. On collaborative detection of TV transmissions in support of dynamic spectrum sharing. In *Proc. IEEE DySPAN*, pages 338–345, November 2005.

[145] Mai Vu, Natasha Devroye, and Vahid Tarokh. On the primary exclusive region of cognitive networks. *IEEE Transactions on Wireless Communications*, 8(7):3380–3385, July 2009.

[146] Mai Vu, Saeed S. Ghassemzadeh, and Vahid Tarokh. Interference in a Cognitive Network with Beacon. In *Proc. IEEE WCNC*, June 2008.

[147] Abraham Wald. *Sequential Analysis*. Dover Publications, 2004.

[148] Fan Wang, Marwan Krunz, and Shuguang Cui. Spectrum sharing in cognitive radio networks. In *Proc. IEEE INFOCOM*, April 2008.

[149] Matthias Wellens, Janne Riihijärvi, and Petri Mähöne. Evaluation of adaptive MAC-layer sensing in realistic spectrum occupancy scenarios. In *Proc. IEEE DySPAN*, April 2010.

[150] Kristen Ann Woyach, Anant Sahai, George Atia, and Venkatesh Saligrama. Crime and punishment for cognitive radios. In *Proc. Allerton*, Septmber 2008.

[151] Liang Xiao, Larry J. Greenstein, and Narayan B. Mandayam. Sensor-assisted localization in cellular systems. *IEEE Transactions on Wireless Communications*, 6(12):4244–4248, December 2007.

[152] Yan Xin, Honghai Zhang, and Sampath Rangarajan. SSCT: A simple sequential spectrum sensing scheme for cognitive radio. In *Proc. IEEE Globecom*, November 2009.

[153] Guoliang Xing, Rui Tan, Benyuan Liu, Jianping Wang, Xiaohua Jia, and Chih-Wei Yi. Data fusion improves the coverage of wireless sensor networks. In *Proc. ACM MobiCom*, 2009.

[154] Yiping Xing, R. Chandramouli, and Carlos Cordeiro. Price dynamics in competitive agile spectrum access markets. *IEEE Journal on Selected Areas in Communications*, 25(3):613–621, April 2007.

[155] Wenyuan Xu, Pandurang Kamat, and Wade Trappe. TRIESTE: A trusted radio infrastructure for enforcing specTrum etiquettes. In *Proc. Allerton*, September 2008.

[156] Yanyan Yang, Yunhuai Liu, Qian Zhang, and Lionel Ni. Cooperative Boundary Detection for Spectrum Sensing Using Dedicated Wireless Sensor Networks. In *Proc. IEEE INFOCOM*, April 2010.

[157] Yi Yang, Xinran Wang, Sencun Zhu, and Guohong Cao. SDAP: A secure hop-by-hop data aggregation protocol for sensor networks. In *Proc. ACM MobiHoc*, May 2006.

[158] Jungkeun Yoon, Mingyan Liu, and Brian Noble. Sound mobility models. In *Proc. ACM MobiCom*, September 2003.

[159] Murtaza Zafer, BongJun Ko, and Ivan Wang-Hei Ho. Cooperative transmit-power estimation under wireless fading. In *Proc. ACM MobiHoc*, May 2008.

[160] Fanzi Zeng, Chen Li, and Zhi Tian. Distributed compressed spectrum sensing in cooperative multihop cognitive networks. *IEEE Journal on Selected Topics in Signal Processing*, 5(1):37–48, February 2011.

[161] Yonghong Zeng and Ying-Chang Liang. Simulations for wireless microphone detection by eigenvalue and covariance based methods. IEEE 802.22-07/0325r0, July 2007.

[162] Pei Zhang and Margaret Martonosi. LOCALE: Collaborative localization estimation for sparse mobile sensor networks. In *Proc. ACM IPSN*, April 2008.

[163] Wei Zhang, Sajal K. Das, and Yonghe Liu. A trust based framework for secure data aggregation in wireless sensor networks. In *Proc. IEEE SECON*, September 2006.

[164] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen. Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework. *Journal on Selected Areas in Communications*, 25(3):589–600, April 2007.

[165] Xia Zhou, Sorabh Gandhi, Subhash Suri, and Haitao Zheng. eBay in the sky: Strategy-proof wireless spectrum auctions. In *Proc. ACM MobiCom*, September 2008.

[166] Qiyue Zou, Songfeng Zheng, and Ali H. Sayed. Cooperative sensing via sequential detection. *IEEE Transactions on Signal Processing*, 58(12):6266–6283, December 2010.