A quantitative approach to wireless spectrum regulation

by

Kate Lee Harrison

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering — Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Anant Sahai, Chair
Professor Jean Walrand
Assistant Professor Sylvia Ratnasamy
Professor John Chuang

Spring 2015

A quantitative approach to wireless spectrum regulation

Copyright 2015

by

Kate Lee Harrison

Abstract

A quantitative approach to wireless spectrum regulation

by

Kate Lee Harrison

Doctor of Philosophy in Engineering — Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Anant Sahai, Chair

Wireless spectrum regulation is an area of increasing interest, complexity, and importance. After decades of single-purpose, exclusive spectrum allocations, the Federal Communications Commission (FCC) brought about the era of dynamically shared spectrum with their landmark ruling in 2008 [1]. Although unlicensed spectrum—"free for all" spectrum such as the 2.4 GHz band—is universally agreed to be critical to economic development, there is scarcely enough to meet current demands. WiFi could never have succeeded without sufficient unlicensed spectrum, yet its success has "filled up" much of the available spectrum. In an effort to address the shortage, the FCC ruled that unlicensed devices are now allowed to operate in portions of the television bands [1].

These "unused" portions are called the television whitespaces (TVWS). Due to the excellent propagation characteristics of the TV band frequencies, a commonly discussed application for TVWS is rural broadband, i.e. using wireless to provide last-mile access to high-speed Internet. Typical TVWS trial deployments have demonstrated the use of TVWS spectrum for backhaul, e.g. to connect school buildings in South Africa or for remote video feeds at a zoo. Others envision Internet-of-things (IoT) applications such as garbage can sensors. The critical piece, though, is that TVWS regulations have to have sufficient flexibility to encourage a wide variety of applications and innovation.

This thesis largely approaches spectrum regulations from a quantitative perspective. It begins by quantifying the current TVWS opportunity, finding that rural spectrum is abundant while urban spectrum is somewhat scarce. But spectrum alone does not tell the whole story so we next look at the data rates achievable with TVWS under various scenarios. The results of this exploration highlighted a shortcoming of the current regulations: lack of sufficient transmit power in rural areas. This led us to explore dynamic power limits for TVWS, showing that intelligent choices for these limits can dramatically improve the utility of whitespaces.

Next, we compare single-use spectrum allocation with whitespaces in the TV band context. We show that allowing the use of whitespaces is superior to repurposing the band from the incumbent's

perspective. We also use a novel data set from the FCC to further compare various spectrum allocation schemes.

In addition to dynamic power limits, we explore other changes to the whitespace ecosystem that we believe will increase utility and promote innovation among whitespace devices. In particular, we carefully decompose the current whitespace architecture into its constituent parts. The decomposition makes testing and certification easier while at the same time allowing more flexible composition of the components. For example, our proposed architecture allows a smartphone to provide a whitespace device with the information necessary to begin transmission in the TVWS; under the current paradigm, only another whitespace device can provide this support. We also demonstrate how the proposed architecture can enable the use of whitespace devices indoors by allowing a variety and combination of location information sources rather than depending solely on GPS which typically does not work indoors.

Whitespaces are not unique to the TV bands. In fact, much of our spectrum lays dormant, particularly in rural areas. While the FCC is currently looking into using whitespace in the 3.5 and 5 GHz bands [2,3], we explore the use of whitespaces in the cellular GSM bands. This work dovetails with existing work on community cellular networks (CCNs) [4,5], community-run cell phone networks in hyper-rural areas, which chiefly lack spectrum in which to legally operate. Our work on GSM whitespaces details how allowing CCNs to use carrier-owned spectrum in rural areas is a win-win-win for regulators, wireless carriers, and citizens alike. We use an existing CCN implementation to demonstrate and evaluate our ideas.

One of the most important commonalities in the aforementioned work has been the *quantitative* approach. Prior work principally explores the space either in a purely theoretical way or with limited real-world data. In contrast, the work in this thesis is almost entirely based on analysis of real-world data, complementing the theoretical work. In an effort to encourage others to explore spectrum regulation from a quantitative perspective, I have made my source code publicly available [6,7]. Part VI discusses the overall architecture of the Matlab and Python code, as well as some of the design decisions that were made.

Since whitespaces present both technical and political problems, we have had the opportunity to submit various official comments to both the FCC and Ofcom. In the spirit of transparency and completeness, we include these comments in Appendix A.

To Adele Sohn Harrison
October 1, 1922 - December 24, 2014

Contents

V Future whitespaces 153

9 GSM whitespaces 154

VI Methods 180

10 MATLAB code 182

11 WEST: Whitespace Evaluation SofTware 194

VIConclusions 218

12 Conclusions 219

13 Future work 222

Bibliography 227

Appendices 239

A Comments to regulatory bodies 240

B Background on repacking and the incentive auctions 271

List of Figures

List of Tables

Acknowledgments

First and foremost, I would like to thank my advisor, Professor Anant Sahai. Without his constant support and encouragement, I would not be submitting this dissertation. It was an incredible stroke of luck to be assigned him as my temporary advisor when I entered Berkeley. He soon suggested that I try working on a small project "for about two weeks;" little did I know that it would form the basis of my work for the next six years.

He allowed me an incredible amount of freedom, not only in my choice of work but also in how I carried it out. He was always willing to entertain — or suggest — an off-the-wall idea, and he always managed to see the nugget of gold at the center. He constantly pushed me to improve myself, but at the same time reminded me that he believed in me. Some of the most enjoyable moments were our long conversations about research, teaching, life at Berkeley, and much more.

Thanks to my other collaborators who helped develop the results in this dissertation: Mubaraq Mishra, Shaddi Hasan, Kurtis Heimerl, Kashif Ali, Sean Roberts, Eric Brewer, Vidya Muthukumar, Angel Daruna, and Vijay Kamble. It was a pleasure working with all of you. I've especially enjoyed working with Vidya, who had a critical role on two papers in her first year of graduate school. It has been exciting watching her grow so much in such a short time and I'm grateful that I got to be a part of it.

Thank you also to my committee, Prof. Jean Walrand, Prof. Sylvia Ratnasamy, and Prof. John Chuang, for all their help and input along the way.

I also want to thank the students of Wireless Foundations that made it such a fun place to work. In particular, Gireeja Ranade, Nihar Shah, and Vasuki Narasimha Swamy contributed significantly to the fun and friendly environment that I will sorely miss.

My fellow TAs and readers for CS 70 in the spring terms of 2013 and 2014 helped form some of my favorite memories at Berkeley. In particular, Chung-Wei Lin, Sibi Venkatesan, Chenya Zhao, Claire Baek, Govind Ramnarayan, Joseph Hui, Sahaana Suri, and Khoa Tran made it an incredibly enjoyable and rewarding experience. I thank Professor Sahai for putting together such a stellar team and for inspirin g us to do an excellent job and have fun at the same time.

Shirley Salanio has my gratitude for her speed, reliability, and cheerful attitude in helping me navigate through the departmental work. My family deserves the biggest thanks of all. They taught me the importance of education from a young age and provided me with the inspiration and opportunity to pursue higher education. They have never wavered in their support and belief in me, and I cherish all of our time together.

Part I

Introduction

Chapter 1

Introduction

1.1   A short history of spectrum regulation

In the 1920s, wireless systems were suffering heavily from excessive amounts of interference due
to a lack of coordination [15]. This interference caused unnecessary performance degradation
because transmitters were operating on overlapping frequencies. Congress responded by passing
the Federal Radio Act of 1927 [16], later replaced by the Federal Communications Act of 1934
[17]. These acts created the Federal Communications Commission (FCC) to oversee and facilitate
spectrum allocation. They primarily use exclusive-use allocations to keep inter-system interference
to a minimum [18].

But now we are at the dawning of a new age, one in which we appear to have all but exhausted this
natural resource[1]. For example, examine Figure 1.1 which shows the frequency allocations in the
United States. Anyone seeking more spectrum—new entrants and incumbents alike—would need
to either buy out an existing company (so as to acquire its license, a common approach) or wait for
a license to expire and bid at the subsequent auction. Both of these are extremely costly options.
As an example, the 2015 AWS-3 auction in the United States sold 65 MHz of spectrum for over
$40 billion [19].

At the same time, we see the massive success of products such as wireless routers. These devices
utilize the Industrial, Scientific, and Medical (ISM) bands which are available for use by unlicensed
devices[2]. Although these bands are free to use, there is still a price to pay: congestion and short
range. The success of WiFi is so great that it has overwhelmed these bands, showing a need for
more unlicensed spectrum. However, the range of WiFi is limited to several hundred meters at best
and is therefore suitable for only short-range purposes. This precludes, for example, the use of the
ISM bands for wireless data backhaul.

---

[1] Or at least our ability to utilize it: one can climb ever higher in frequency but reliable and reasonable-range
transmission becomes ever harder.

[2] These devices are still subject to FCC certification but they need not acquire a license to use the band.
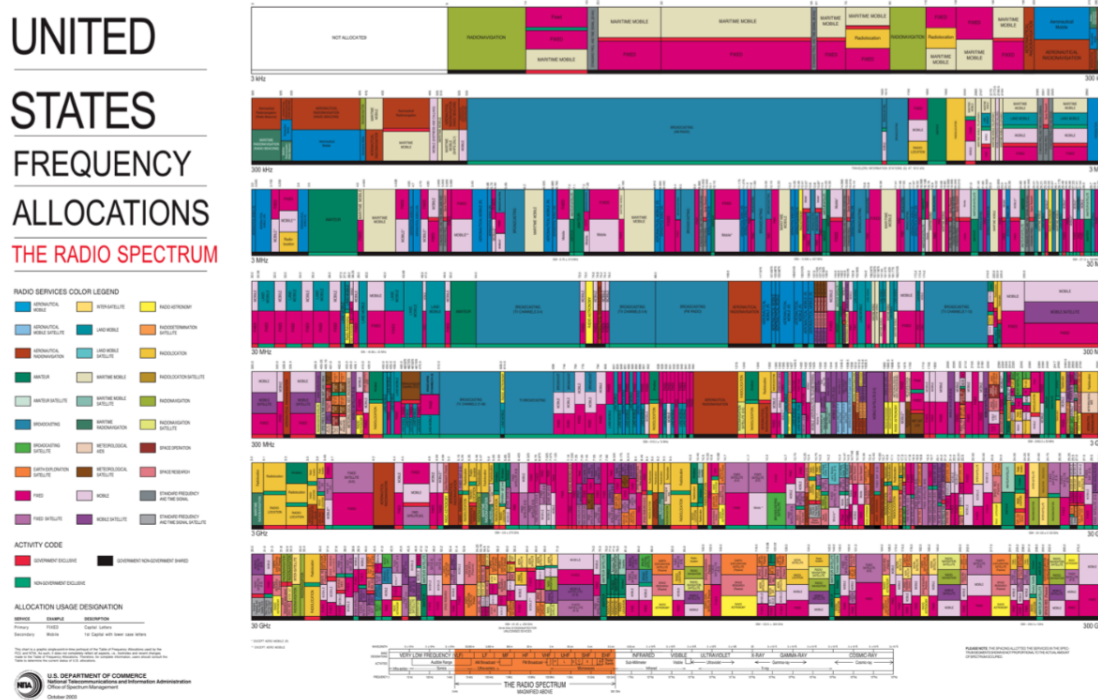
Figure 1.1: Frequency allocations in the United States [8].

Also in play is the country's need for pervasive broadband access, as described in the National Broadband Plan [20]. This piece of legislation, passed in 2010, dictates that 100 million American households should have access to broadband by the year 2020. However, the population distribution of the United States makes this an expensive undertaking, since many Americans live in rural areas and it is well-known that the "last mile" of access is the most expensive to provide. This makes wireless solutions more attractive than traditional physical solutions such as fiber. However, the cost of acquiring the spectrum to do so can quickly tip the scales.

We have recently seen an evolution in seemingly-unrelated technologies. For example, it was demonstrated (via actual deployment) that radar systems and wireless LANs could actually intelligently coexist in the 5 GHz band without causing deleterious effects on one another. This showed that technology has given us a potential to move away from the strict time-space-and-frequency boxes that the FCC and other regulators used to describe allocations and assignments.

Finally, a variety of studies (e.g. [21]), and in some cases common knowledge, have shown that although spectrum allocations are typically made on a nationwide basis, they are not used nationwide. For example, a 2004 study at the Berkeley Wireless Research Center showed that up to 70% of spectrum below 2.5 GHz remains unused *even in an area as populated as Berkeley* [22].

In summary:

- We appear to have run out of usable spectrum.

- Licensed spectrum is expensive to acquire, both in time and dollars.

- We need more unlicensed spectrum in order to spur economic growth.

- We need more spectrum in order to provide Internet access to rural areas.

- Intelligent coexistence in the same band has been demonstrated to work.

- There is room in existing spectrum usage for other entities.

1.2   The beginning of whitespaces

All of these pieces and many others contributed to an increased interest in dynamic spectrum sharing. One of the landmark moments was the FCC's publication of a Memorandum Opinion and Order (MO&O) in 2008 that made spectrum sharing legal in the over-the-air television bands [1]. These rules were subsequently updated in 2010 [9] and 2012 [23] but many of the main ideas have remained constant.

Some of the other key regulatory moments are shown in Figure 1.2, including the 2015 adoption of TV whitespaces by Canada [24] and the United Kingdom [25]. Among the events not pictured are the whitespace efforts for other bands happening in the United States, e.g. the 2014 5 GHz Report and Order [26] and the 2015 3.6 GHz Report and Order [27].
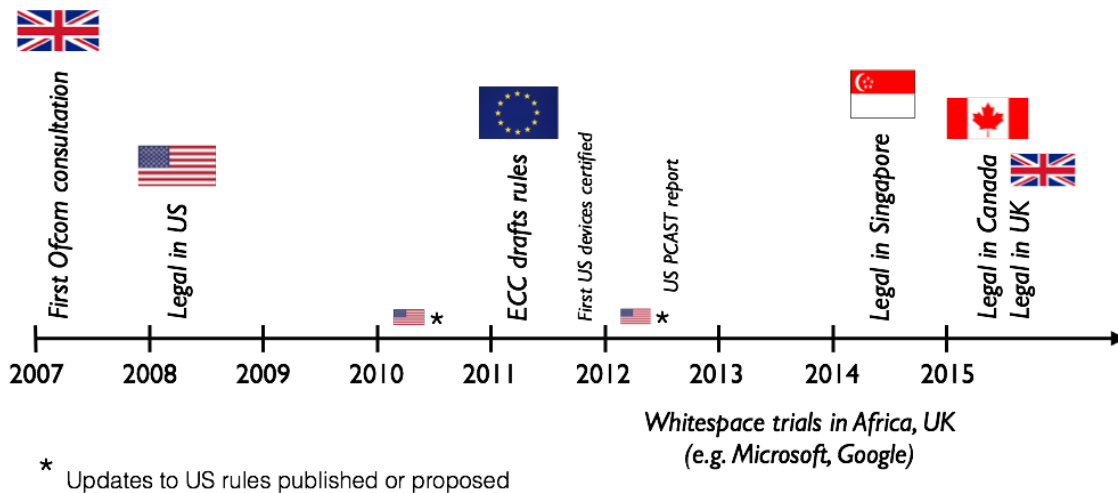
Figure 1.2: Timeline of TV whitespace activity around the world.

The term *spectrum whitespaces* refers to unused (time, space, frequency) resources, as illustrated in Figure 1.3. Whitespaces are found outside of the service areas of incumbents and on unused frequencies.

(a) Abstract representation of spectrum whitespace (b) Spatial variation of TV whitespace on a single channel in the United States
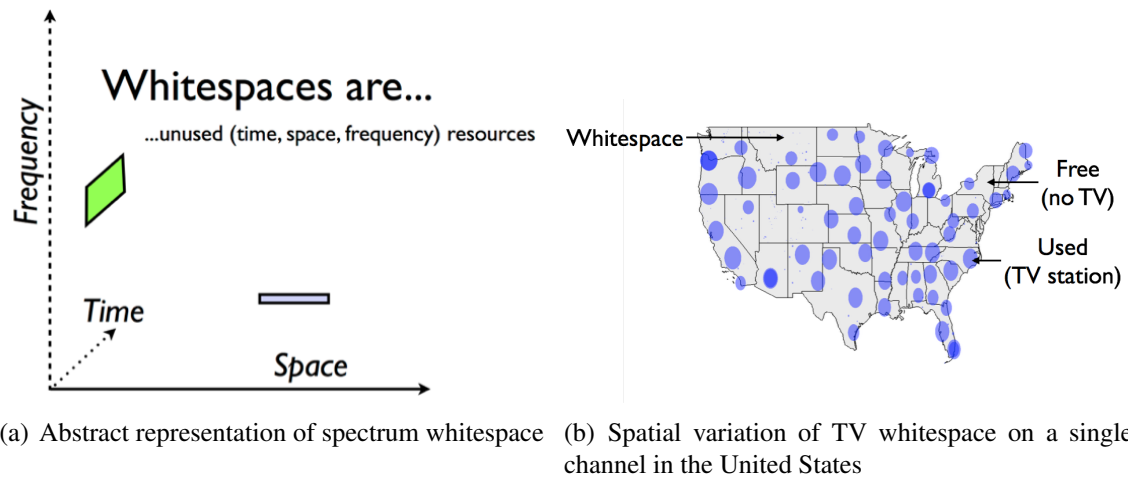
Figure 1.3: Illustration of whitespaces

We briefly review the relevant portions of the United States whitespace regulations. Regulations in other regions are similar in spirit and hence the details are omitted for the sake of clarity.

- The incumbents, television operators, are referred to as *primaries*. This name reflects their priority within the band: their service area and quality of service must be (roughly) preserved. In an ideal world, the primary would be unaware that it is actually sharing the band.

- *Secondary* devices are allowed to operate in the same bands, but they must follow various rules which are intended to provide protection to the primaries. This may mean sacrificing its quality of service to avoid interfering with the primary.

- The service area of each television station is protected in two dimensions: in space and in frequency. This is illustrated in Figure 1.4.

  - Spatially, a buffer zone consisting of the first N km beyond the service area[3] is created (shown in red). This "no man's land" guarantees a minimum distance between the primary's users and the secondary's transmitters.

  - In frequency, a buffer zone of one adjacent channel in each direction is imposed for "fixed" white space devices (i.e. devices which are professionally installed) operating above a certain power. This helps guard against secondary transmitters with poor emissions masks and protects primaries with poor receiver masks.

  Thus a single primary's footprint, as viewed by a secondary device, is a slightly-enlarged version of the service area which is up to three-channels "wide" in frequency (i.e. all of

---

[3]In practice the value of N varies by device type, height, and region. It may be as small as 4 km but up to 32 km in the US.

the colored circles in Figure 1.4). The area outside of these footprints is what is termed "television whitespaces."

- Secondary devices have a fixed power limit based on their device type (fixed vs. personal/portable).

- Secondary devices must have a way of determining which channels are safe to use (more on this later).
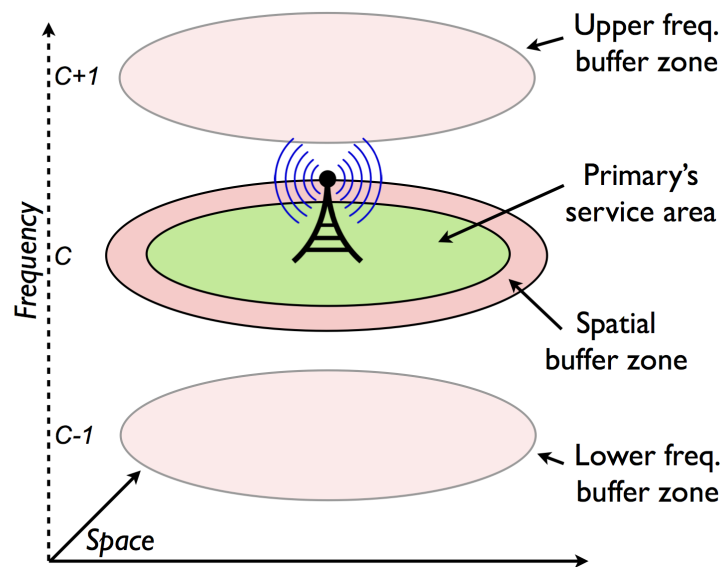
Figure 1.4: *Exclusion zone illustration.* Secondary devices are not allowed to transmit inside of any of the colored circles. The primary transmitter's natural noise-limited service area is shown in green. A spatial buffer zone (a "no man's land" of sorts; shown in dark red) guarantees physical separation between the primary users and the secondary transmitters. Buffer zones in frequency protect primary users with poor receiver masks from secondary transmitters with lenient emissions masks.

1.3 Contributions

1.3.1 Quantifying whitespaces

My initial work, shown in Chapter 2, focuses on quantifying the opportunity opened up by the FCC. I built on the work of Mubaraq Mishra, one of my advisor's former students, who had started this line of investigation by gathering the necessary data on TV tower allocations [28] and US population [10] along with a propagation model [29]. With this data, I was able to answer the following questions:

1. Based on the FCC's TV whitespace regulations, how many of the 6-MHz TV channels are available for secondary use? The results are shown as a heat map in Figure 1.5.

2. Given that the "stray" TV signals must be treated as noise by the secondary systems, what is the theoretical capacity (using Claude Shannon's formula) for a single transceiver pair? This can be calculated for a variety of communication distances.

3. Clearly the whitespaces will not be used exclusively by a single pair of transceivers. What is the theoretical capacity given realistic models for self-interference among secondaries? My work includes both a MAC exclusion model as well as a cellular model for secondary self-interference.
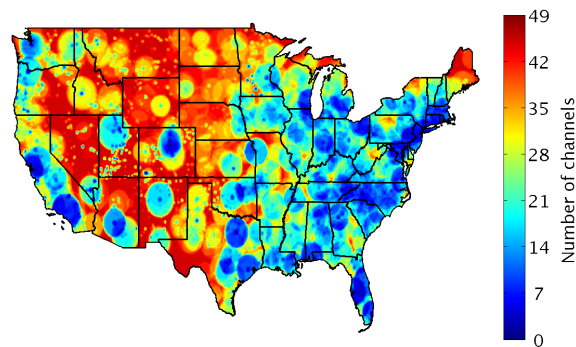
Figure 1.5: Number of 6-MHz TV channels available for secondary use in the United States.

My work along with Mubaraq's helped shape the way the community studies whitespaces. Prior to our papers, there were no nationwide assessments of whitespace opportunities and regulations: most researchers focused either on specific examples at a few locations or on much more general models. Our work showed that data-driven nationwide studies are not only possible but *essential* to whitespace-related research. Since our papers were published, in particular [30], many other researchers have followed our lead [31–35]. In one case, an Australian regulator specifically suggested that whitespace proponents conduct a study similar to ours in [30] for Australia[4] [36]. Recently Ofcom has produced similar maps for the United Kingdom [37].

This thesis has a clear message: *a data-based and quantified analysis of regulatory decisions is essential*. As evidenced by this thesis, simulations using real-world data have a large role to play in the future of wireless spectrum regulations. In [38] and in this thesis[5], we begin by simply calculating *how much* whitespace is available and estimating its utility via achievable data

---

[4]We have since conducted this study and the results are shown in Chapter 11.

[5]For completeness, this thesis also contains the work presented in my MS thesis, [38].

rates under various system deployment scenarios. We continue by exploring the ways in which geographic variation and a desire for fairness affect potential regulatory changes aimed to increase this utility. We show that regulations or management schemes for a varied environment need to be cognizant of this variation in order to maximize utility.

Next, we compare whitespaces to the more traditional spectrum reallocation in Part III. We show that whitespaces are the only rational option if the incumbent service is to be preserved to any reasonable degree. We also utilize datasets released in conjunction with the FCC's upcoming incentive auction to explore a different aspect of this tradeoff in which incumbents can be "repacked" (i.e. assigned to a different channel) rather than simply removed. We show that efficient repacking of incumbents essentially trades whitespace for cleared spectrum and that even the most efficient repackings leave plenty of whitespace spectrum.

1.3.2 Improving the whitespace ecosystem

In Part IV, we move up to a higher-level discussion about the fundamental components of the whitespace ecosystem. These components include the elements necessary to determine a whitespace device's location and to contact a whitespace database. We argue that these components are already naturally separate from the mission of the device itself, and that thinking of them as such rather than as intertwined systems promotes innovation and less expensive devices, improves certification, and helps to address the indoor localization problem[6]. We also tackle another aspect of the localization problem: supporting weakly-localized devices. While some devices have GPS built in, others may not. Regulators have different approaches to supporting these weakly-localized devices and while some are too lax, others are too stringent. We propose a simple scheme which allows regulators to be appropriately careful without placing a large burden on whitespace devices or significantly degrading the availability of spectrum. One of our key observations is that a whitespace device only really needs to know a set of safe operating parameters, *not its location*, which is a fundamentally simpler problem. Another key observation is that the whitespace database only needs to ensure that the parameters it gives to a device are *safe*, not that they are as permissive as possible. This allows database providers and regulators to trade off efficiency of spectrum utilization with flexibility of implementation.

While the television whitespaces are the most studied whitespaces to date, there is a wealth of whitespace in the rest of the spectrum. We paired with researchers at UC Berkeley's TIER (Technology and Infrastructure for Emerging Regions) to study GSM whitespaces, i.e. pieces of the GSM (cellular) spectrum which are not being utilized. In Chapter 9 and our award-winning paper [39], we show that opening up the GSM whitespaces can be a win-win-win for regulators, incumbent cellular providers, and rural citizens alike. This work, in conjunction with TIER's extensive work on Community Cellular Networks, led to the formation of the startup Endaga which aims to help bring GSM cellular coverage to rural areas throughout the world.

---

[6]Traditional whitespace devices use GPS to determine their locations and GPS notoriously does not work well indoors. Our proposal would allow a whitespace device to get its location from a variety of sources.

1.3.3 Higher-level goals

The common thread for much of our work has been our *quantitative* approach. Because we strongly believe that research should be collaborative and open, we have made our research code freely available online from the very beginning. Many of the results presented here were generated with Matlab and we discuss the structure, assumptions, and capabilities of the associated code in Chapter 10. After several years of experience developing and using our Matlab code base, we decided to refactor it and switch to Python. Refactoring allowed us to rewrite the code with flexibility in mind, drawing on years of experience and with an eye toward the future. We chose Python because it is more readily available than Matlab and accessibility is critical to promoting adoption. Chapter 11 describes the design of WEST, our Python framework for quantifying whitespace availability and utility.

Finally, we feel that an important part of our work is sharing our findings with regulators. In the past two years, we have submitted four official comments to the Federal Communications Commission (US) and Ofcom (UK). These submissions are a matter of public record but for completeness can also be found in Appendix A. We strongly encourage other researchers in our field to follow this example and engage regulators as much as possible.

1.4 Research vision

I have several beliefs which have driven and continue to drive much of my work. I will briefly describe them below so that the commonalities and motivations behind my work become clear.

1. *Regulations should not unnecessarily constrain device or system designs.* Light-handed and flexible regulations should be preferred, as they are easier to enforce and otherwise beneficial to both parties. I will expand on this in Chapters 7 and 4.

2. *Regulatory decisions should be data-driven.* In an age where data is often easy to come by and computing resources are incredibly powerful and cheap, we no longer need to use heuristics and hope for the best. For example, many simulations to estimate aggregate interference, secondary utility, etc. could be used to calculate the optimal size of the spatial buffer zone. By contrast, the current numbers are derived from a back-of-the-envelope calculation[7] with no subsequent analysis of the consequences.

3. *Data means nationwide data.* It is insufficient to study only toy models or a small geographic region when making nationwide regulations. As Figure 1.5 shows, there is a large spatial variation in the whitespaces which must be taken into account.

---

[7]From paragraph 181 of the 2008 FCC rules: "For example, using the FCC curves in Section 73.699 of the rules and the D/U protection ratios specified above, we find that a transmit antenna at a height of 30 meters transmitting with 4 watts EIRP could cause co-channel interference to a TV receiver with an antenna 10 meters above ground at a distance of 14.4 kilometers and adjacent channel interference at 0.74 kilometers." [1, ¶181] (note: the cochannel buffer zone was previously 14.4km).

4. *Research tools, code, and data should be open-source.* Researchers should strive to make their published work easily reproducible by others. This has the following benefits:

   - Researchers have a clear and easy way to compare their results with prior work.

   - When researchers are able to directly compare their work to others', it is easier to understand the studied phenomena which can no longer be explained away by differences in methodology.

   - As certain tools become more popular, useful baselines and metrics will emerge. This enhances the quality of research discussions.

   - Researchers save time by not needing to "reinvent the wheel" before working on something interesting.

5. Comments to regulatory bodies are currently given in a sort of haphazard way: each entity creates its own experimental platform (using mathematics, computer code, and/or real hardware) and uses data from that to back up its opinions. However, some decisions, such as the size of the spatial buffer zone, can be very cleanly analyzed in a standard way. I believe that *there should exist a standard FCC-approved set of software that can be used to frame comments*. This will decrease overhead within the regulator and generally improve the discourse.

Some of these are admittedly long-term goals which will not be achieved by a single person. However, I hope that my work will help guide the community in these directions, and so I try to lead by example with my own research.

Part II

Whitespace utility

Chapter 2

Evaluating whitespaces

Some or all of the work in this chapter appeared in the author's Master's thesis, [38].

This chapter quantifies the opportunity available to secondaries obeying the FCC's regulations in the TV whitespaces for both single-link and cellular applications. It also determines the relative impact of various effects: the no-talk regions necessary for primary operation, noise from primaries, self-interference from secondaries, height, and range.

The FCC created exciting new prospects when it designed the TV whitespaces, the "vacant" areas of the TV bands in which unlicensed devices are permitted to operate. These devices are allowed to transmit when they are far enough from the TV towers. We first explore the amount of spectrum the FCC regulations provide for unlicensed use. However, this spectrum is not actually "white" due to pollution (TV signals which are essentially noise to secondaries), making it less valuable than clean spectrum. We will show using a simple single-link model that this pollution is actually a minor effect in comparison to the need to halt transmissions near the TV towers.

Along with environmental concerns such as TV signal noise, devices operating in the TV whitespaces must concern themselves with traditional system parameters, namely transmitter height and communication range. We will show that in general the range plays a significantly larger role than transmitter height. However, a poor choice of height may lead to a 50% degradation in link quality.

In addition to pollution, secondaries must also concern themselves with noise from other secondary devices called self-interference. We incorporate medium access control (MAC), a common method for mitigating self-interference [18]. In this scheme, devices take turns transmitting in order to increase their individual rate. Using this model we show that the single-link rates are overly optimistic.

Furthermore, although some systems operate fixed-range links, many do not. We argue that the range is often a function of the local population density rather than an endogenous variable. Under this model, we see that rural areas experience much lower rates despite having more available

channels due to large communication ranges. Conversely, urban areas do quite well despite having few channels.

Finally, we consider using a standard cellular model in addition to our point-to-point model. We show that, similar to the results for the MAC model, the opportunity has decreased relative to the single-link model.

2.1   Background: overview of current US regulations

In this section we will briefly review the relevant portions of the regulations set forth by the FCC in their regulations [1, 9, 23] for TV whitespace devices which utilize databases to find "unused" spectrum. These regulations are currently awaiting revision via the Part 15 Notice of Proposed Rulemaking (NPRM) [40]. No major changes are expected but a variety of small details will likely be altered.

Secondary devices are classified as depicted in Figure 2.1(a). Restrictions on a device are based on its classification, as shown in Figure 2.1(b). Restrictions generally fall into a few categories:

(a) Classification   (b) Restrictions
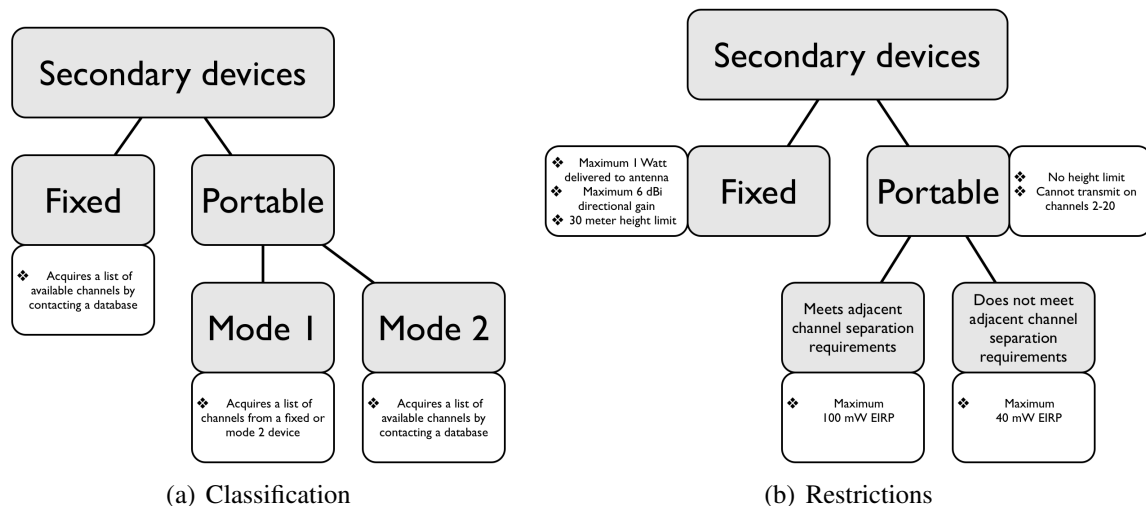
Figure 2.1: Device classification and restrictions for database-connected devices according to the 2010 FCC rules [9].

- **Power**: a limit on the effective isotropic radiated power (EIRP) or on amount of power delivered to the antenna and the directional gain.

- **Height**: a limit on the height above average terrain (HAAT) of a device. This is only imposed for fixed devices as portable devices have no way to measure their HAAT.

- **Channels**: no secondary device will be allowed to transmit in channels 3, 4, and 37. Channels 3 and 4 are used by legacy equipment such as VCRs and early DVD players. Channel 37 is not used for TV broadcasting; instead, it is reserved for radio astronomy and wireless medical telemetry services. This leaves the remainder of channels 2-51 for secondary use [1]. *Many of these restrictions are expected to be lifted by [40].* Further restrictions on channel usage are imposed based on location, as described below.

- **Location (cochannel)**: no secondary devices of any kind are allowed on the same channel within the protected radius (denoted r_p) of a TV tower. As the adjective "protected" implies, the TV signal should still be decodable within this region. This radius is a function of the characteristics of the TV tower and the surrounding terrain. Furthermore, each device must maintain an additional separation distance, $r_n - r_p$, from the edge of the protected region. This value ranges from 4 km to 31.2 km depending on the height of the secondary device [23, §15.712(a)(2)]. This is shown in Figure 2.2 as the "no man's land."
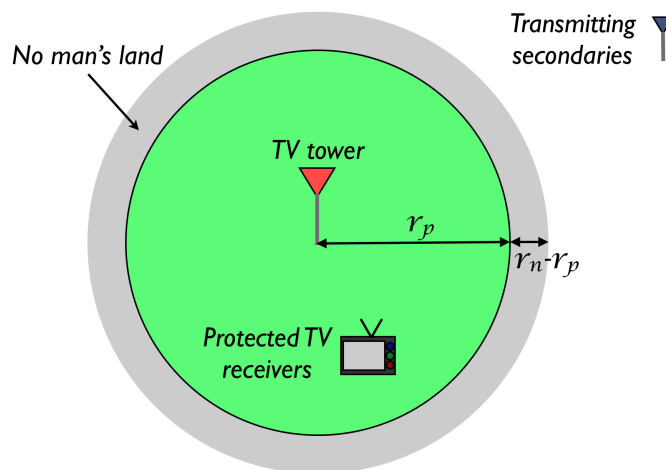
Figure 2.2: Illustration of the protection radius r_p and the no-talk radius r_n.

- **Location (adjacent channel)**: secondaries are also excluded from the circle of radius $r_{n,A}$ centered at the TV transmitter in the adjacent channel. The protected radius is the same as in the cochannel case; however the value of $r_{n,A} - r_p$ again depends on the height of the secondary transmitter and varies from 0.4 km to 2.4 km. Note that $r_n - r_p \neq r_{n,A} - r_p$, so $r_n \neq r_{n,A}$. A secondary which is outside of every tower's $r_{n,A}$ on all adjacent channels is said to meet adjacent channel separation requirements.

Although the protected region is depicted as a perfect circle in Figure 2.2, this is rarely the case. The protected region is actually defined by the received signal strength of the TV signal, which

---

[1]The FCC also bans transmission on channels 2-21 for portable devices.

Furthermore, it reserves two channels at all locations for wireless microphone usage. For example, channels 24 and 41 are reserved near Chicago and channels 15 and 16 are reserved in Berkeley.

is in turn dependent on local terrain features. More information about these calculations can be found online[2] and specific contours are available at [41].

We often do not consider the specifically excluded regions such as radio astronomy protection zones, but these are generally areas of low population and thus have little impact on our study. We also often omit PLMRS/CMRS exclusion zones, as described in Section 10.7.3. The current FCC database only lists stations in the United States, thus we also ignore TV transmitters in Canada and Mexico [3]. Further limitations of our model are discussed in Section 10.7.

2.2 Channels available to secondaries

Since TV towers tend to be located around population center and the presence of a TV station precludes secondary use in the nearby area, it may be reasonable to conclude that there are fewer channels available in higher population areas.

In order to verify this claim, we use real tower data for the United States [28] to evaluate the FCC rules [9] on a nationwide scale[4]. We enforce cochannel and adjacent channel exclusions, as defined and discussed in Section 2.1, and then count the number of available channels at each location. This is shown in Figure 2.3(a) in the form of a color-coded map. We see that up to 45 channels are available in some locations and even Los Angeles has at least four channels available in most places.

Each TV channel is 6 MHz wide. For comparison, wireless routers using the IEEE 802.11 standard and operating in the 2.4 GHz ISM band use channels which are 20 MHz wide, so a wireless router using the TV white spaces would require four TV channels for operation. However, signals attenuate faster in the ISM bands than they do in the TV bands, so 20 MHz in the TV bands is much more useful than the same amount of spectrum in the ISM bands.

Figure 2.3(b) shows the complementary cumulative distribution function[5] (CCDF) by population for the map of Figure 2.3(a). From this we can see that at least 87% of the population has at least four channels and could therefore operate a wireless router in the TV whitespaces. Furthermore, only about 1.5% of people have zero channels available (colored black on the map).

Notice that the coasts of the United States have fewer available channels than the central portion of the nation. In fact, one of the key obstacles to secondaries is that high-population areas have the least bandwidth. This phenomenon is due to the fact that TV towers (which "take" channels from secondaries) tend to be located near population centers. As we suspected, a larger population generally means fewer channels available for secondary use. We see this more clearly in Figure 2.4: at higher population densities, there are typically fewer channels available.
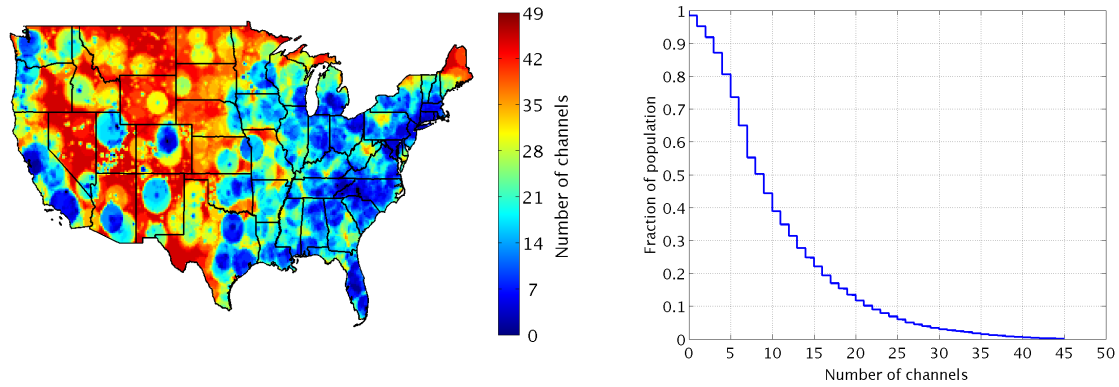
(a) Color-coded map of the continental United States

(b) CCDF by population of the number of channels available to secondaries nationwide

Figure 2.3: Number of 6 MHz channels available to secondaries in the United States
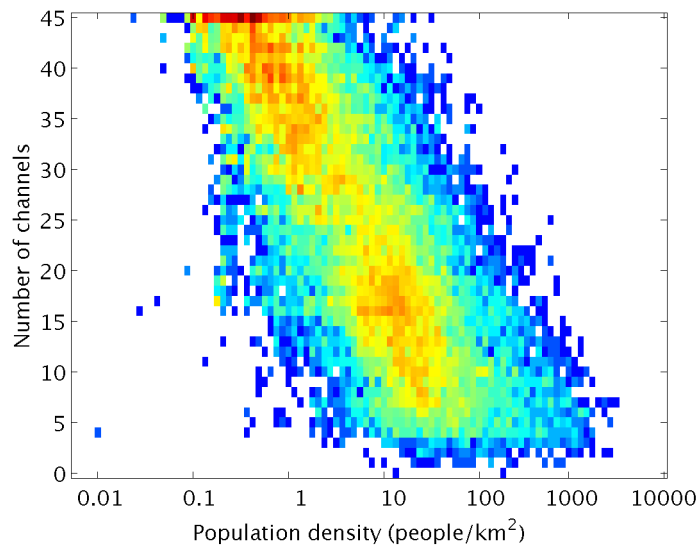
Figure 2.4: Color-coded histogram (of pixels) showing the inverse relationship between population density and the number of channels available for secondary use. Red indicates the highest frequency and blue the lowest frequency. White indicates that there were no data points in that bin.

However, the data presented are potentially misleading. Not all whitespace channels are alike because the TV signals, which do not simply end at r_p, raise the noise floor for secondaries in a spatially- and frequency-varying manner. The next section discusses this problem and quantifies its impact.

2.3 Impact of primaries on secondary utility: are the whitespaces "white"?

A simple count of the number of channels available to secondaries is not sufficient to evaluate the TV whitespaces for secondary use. Spectrum users are interested in quantities such as deliverable data rate rather than the amount of spectrum available. In exclusively-allocated channels these quantities are linearly related; however, shared bands are quite different in this regard.

Primaries affect the achievable rates of secondary users through two factors: noise from TV towers and location-based channel restrictions. We will see in this section that the effect of the noise is small relative to that of the exclusions.

To assess the available data rate, we use the information-theoretic Shannon capacity formula [42]:

$$\text{rate} = \text{bandwidth} \cdot \log_2 \left(1 + \frac{\text{received signal (desired) power}}{\text{noise (undesired) received power}} \right)$$

The the fraction of received signal (desired) power to noise (undesired) received power is commonly referred to as the signal to noise ratio (SNR). For simplicity we do not consider schemes such as dirty paper coding [43] which use information about the noise to increase the rate.

In the TV whitespaces, the bandwidth of each channel is 6 MHz and rate can be summed across channels. Thus we can express the total rate at a given location as:

$$\text{total rate} = \sum_{c \in \text{TV channels}} a_c \cdot (6 \cdot 10^6) \cdot \log_2 \left(1 + \text{SNR}_c \right)$$

where $a_c = 1$ if channel c is available for secondary use at that location and $a_c = 0$ otherwise.

---

[2]See http://www.ecfr.gov/cgi-bin/text-idx?node=se47.4.73_1625.

[3]The FCC rules state that the service areas of non-US TV towers must be protected within the country of origin up to the US border, but not beyond. Thus US residents who previously received Canadian or Mexican TV stations are not guaranteed to continue to receive these stations.

[4]Details on map generation can be found in Chapter 10.

[5]Details on CCDF calculations can be found in Chapter 10.

Unlike in many other bands, TV whitespace devices may be subject to substantial noise due to TV tower transmissions. We now examine the effect of this noise as well as protected areas on the achievable data rates for whitespace devices.

In typical wireless systems where the band is licensed, the license-holder is free from in-band interference (a.k.a. noise) and has a guaranteed limit on the amount of out-of-band interference to which the system is subjected. Thus, the majority of the interference that he faces is due to transmitters he has installed himself. With this level of control over the interference, he can carefully choose the locations of his transmitters to maximize his utility.

In unlicensed bands, devices operate at much lower powers. In this way, even though individual devices cannot control the amount of noise experienced due to other transmitters, they are reasonably assured that the total interference will be small.

The situation for the secondaries in a spectrum-sharing situation is essentially the worst of both worlds. They lack control over the other secondary transmitters in the band and they must cope with the signals from the primary transmitters which is essentially noise to them. Furthermore, they have to turn off whenever they get too close to the primary.

To understand how these effects play out, we look at a toy example in which we have a single secondary transmitter-receiver pair in the United States[6]. For this single-link example, we assume that the transmitter is operating at maximum height (30 meters) and maximum power (4 Watts EIRP[7]) on all available channels[8]. When not otherwise specified, these are the default secondary characteristics used throughout this chapter.

We also assume that all TV transmitters are transmitting omnidirectionally at their maximum power and HAAT. Furthermore, we assume that our secondary receiver is able to attenuate noise from adjacent TV channels by 50 dB. Thermal noise (*TNP*) is present at all locations on all channels.

We will specifically look at two factors: interference from the primary transmitters (which we call *pollution*) and the need to halt transmissions at locations near primary receivers (which we call *exclusions*).

In Figure 2.5 there are four color-coded maps of the continental United States which indicate the data rate available to secondaries[9]. Each map represents a hypothetical world. In Figure 2.5(a), the secondaries are operating in a clean channel (i.e. there are no other transmitters on that channel) and are allowed to operate on all channels at all locations. Adding noise (pollution) from the TV towers but still allowing them to transmit anywhere leads to Figure 2.5(b) and similarly limiting their location-channel combination but pretending they can ignore TV signals leads to Figure 2.5(c). Finally, combining these two effects (pollution and exclusions) gives us the true
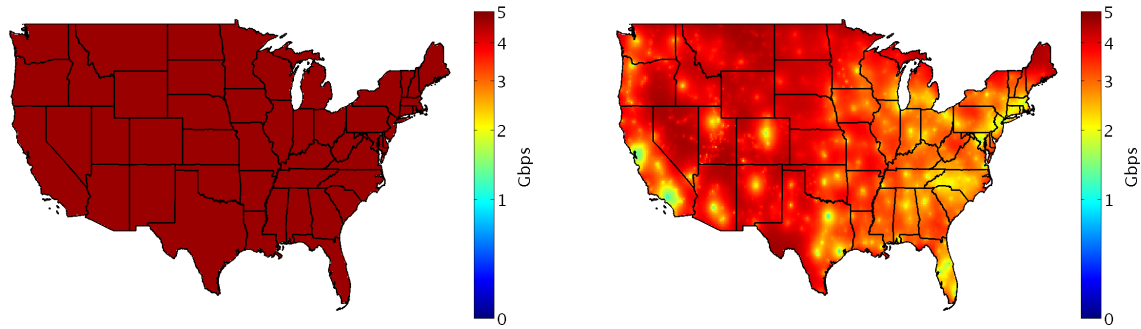
(a) No pollution, no exclusions

(b) Pollution, no exclusions
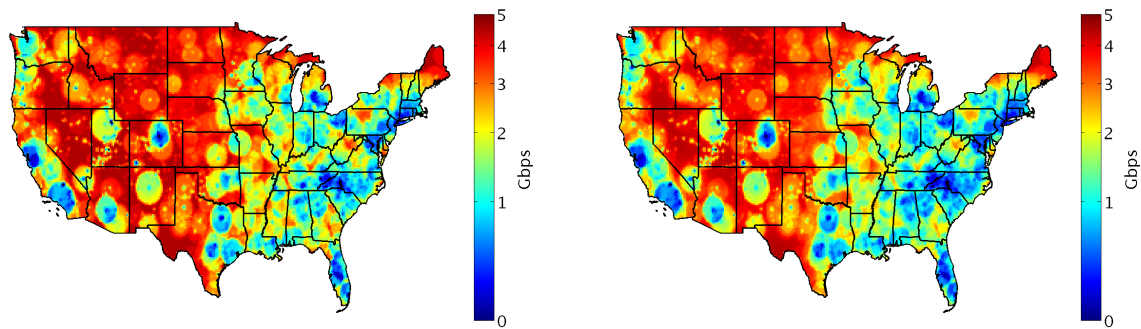
(c) Pollution, no exclusions

(d) Pollution, exclusions
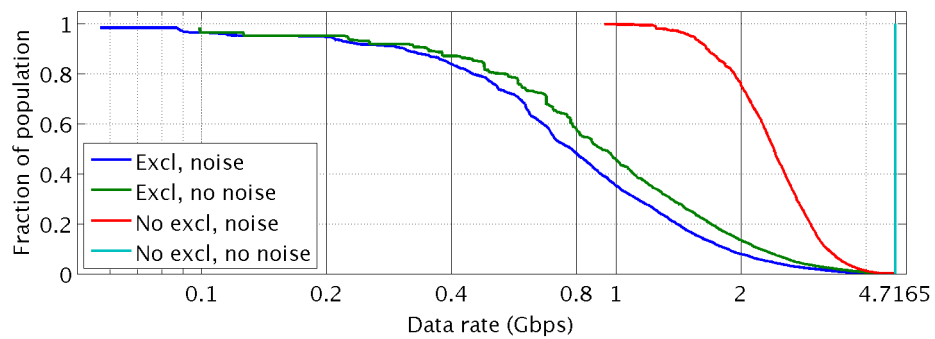
(e) CCDFs by population of maps (a)-(d)

Figure 2.5: Data rate available to a single secondary pair 1 km apart under various conditions.

state in Figure 2.5(d). It is clear from these figures that exclusions are more painful than pollution to secondaries.

In order to understand how this affects the people of the United States, we calculate a complementary cumulative distribution function (CCDF) sampled by population for each map[10]. These CCDFs, shown in Figure 2.5(e), support our earlier conclusion that exclusions are more painful than pollution. Even if secondaries could ignore TV noise, they can do no better than the green line. However, removing exclusions (but keeping the noise) would get us to the red line, whereas a clean and completely available set of bands is represented by the cyan line.

Note that the the gap between the red and cyan lines (adding in noise with no exclusions) is greater than the gap between the blue and green (adding in noise with exclusions) lines since the former allow transmission in areas which by definition are close to TV towers and are therefore noisier, resulting in a lower average SNR for the secondary.

This section has concentrated on the exogenous effects on secondaries, showing that exclusions are a much more serious problem than noise from TV towers. In the following sections, we focus on the endogenous effects that secondaries face.

2.4   Height vs. range: which matters more?

Height and transmission range are important parameters in any wireless system, including those operating in the TV whitespaces. In this section we will look at each factor separately and then show that range is dominant in most situations.

For reference, Figure 2.6 shows how a transmitted signal attenuates over distance in our propagation model for different transmitter heights[11]. Notice that height is inconsequential at some ranges.

Range

The transmission range of a communication link is the distance between the transmitter and the receiver. In Figure 2.7 we see just how dependent data rate is on the transmission range with the data rates collapsing in the 10-km case. This is due to the fact that a transmitted signal attenuates more with increasing distance, thus the SNR at the receiver is lower in the 10-km case.

---

[6]Thus interference from other secondaries is not considered; it will be discussed further in Section 2.5.

[7] [9] actually specifies a maximum transmit power of 1 Watts but allows for up to 6 dBi of directional antenna gain, thus resulting in a maximum of 4 Watts.

[8]The FCC specifically mention that the maximum total secondary power is fixed and is *not* a function of the number of channels used. For illustrative purposes, we ignore this.

[9]Details on map generation can be found in Chapter 10.

[10]Details on CCDF calculations can be found in Chapter 10.

[11]All of our models use the ITU propagation model [29].
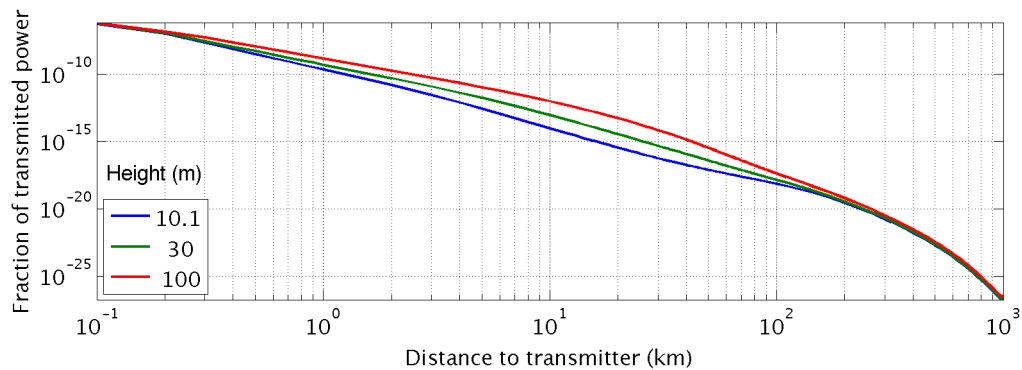
Figure 2.6: Signal strength as a function of distance from transmitter, height

(a) 1km transmission distance

(b) 10km transmission distance
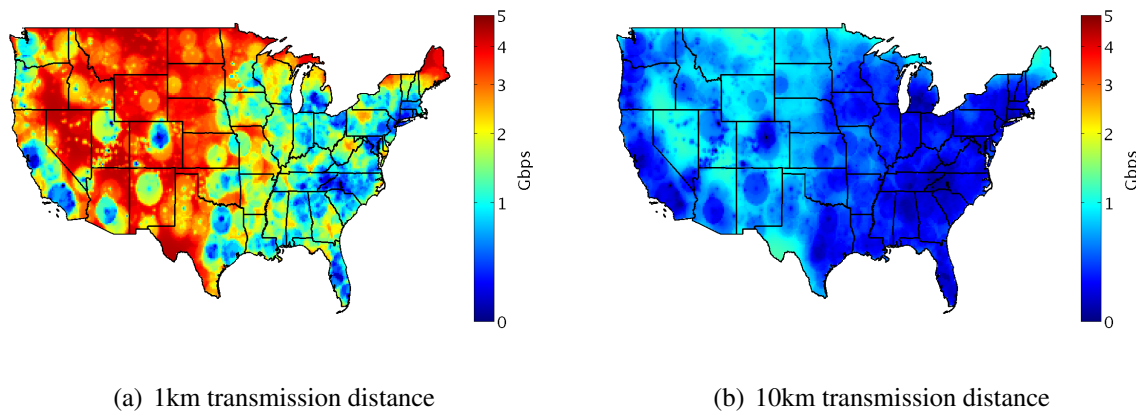
Figure 2.7: Data rate available to a single secondary pair with differing ranges.

Height

The height of a transmitter affects the propagation characteristics of its transmitted signal. Increasing the transmitter height increases the signal strength at any given location. However, we can see from Figures 2.6 and that this effect is minimal compared to that of range.

Comparison

To find the relative magnitude of these effects in the TV whitespaces, we created nationwide data for different (height, range) pairs[12]. For each pair, we found the rate of the average person and used that as a final data point, as shown in Figure 2.9. Again, the height boost only occurs at

---

[12]This comparison isn't entirely fair since the FCC likely would not have chosen the same power limit if they allowed secondaries that were 100 meters tall.

(a) 30m HAAT   (b) 100m HAAT
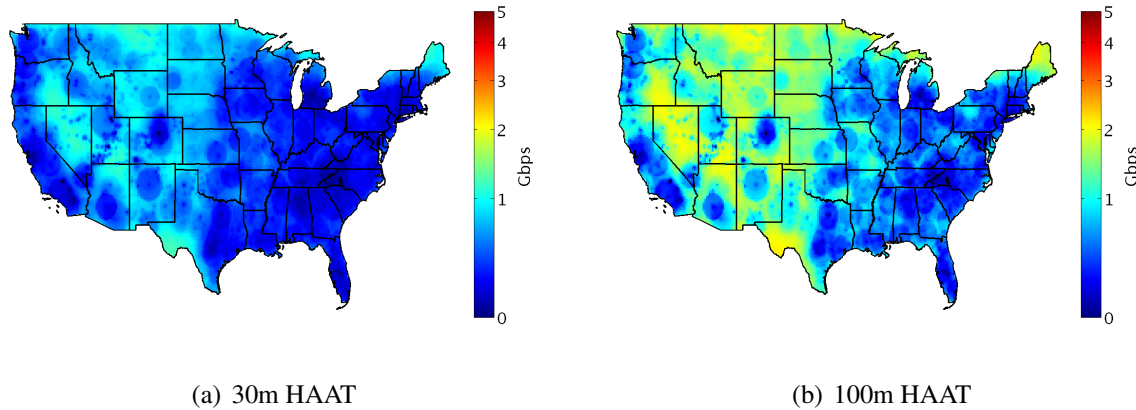
Figure 2.8: Data rate available to a single secondary pair with differing heights.

certain ranges since the signal strengths only differ at certain ranges. However, these ranges – approximately 5 km to 150 km – are precisely those that are likely to be used for long-range applications. In these cases, it is important to have and to exercise flexibility in transmitter height in order to maximize utility. Choosing the wrong height for a 50-km-range application may mean an order-of-magnitude decrease in utility.
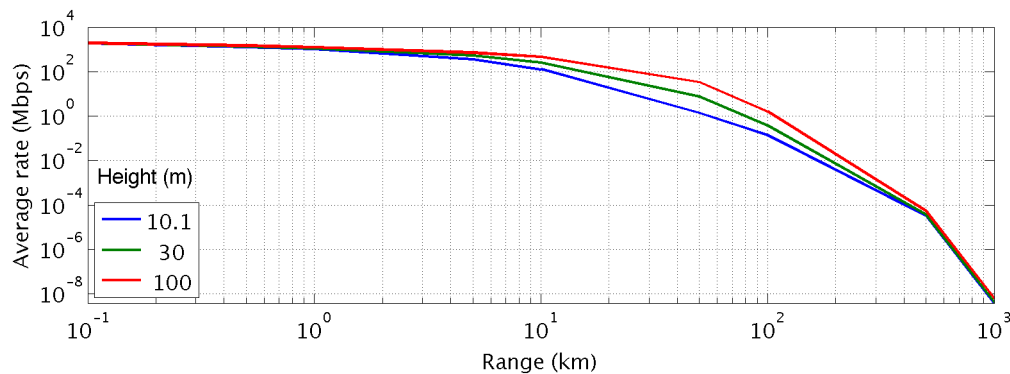
Figure 2.9: Nationwide average achievable secondary rate as a function of range and transmitter height

2.5 Self-interference: an unfortunate reality

The simple single-link model used in the previous sections promises incredibly high data rates because it ignores the effects of self-interference among secondary users. To a receiver, any unintended signal it receives is considered noise (also known as interference), even those from

other secondary systems. This section implements a secondary sharing protocol to enhance the model and shows that the single-link rates are absurdly optimistic.

The Shannon capacity formula can be written more explicitly as

$$\text{rate} = \text{bandwidth} \cdot \log_2 \left(1 + \frac{\text{received signal power from transmitter of interest}}{\sum \text{received TV signal power} + \sum \text{received power from other secondaries}} \right)$$

Thus each secondary suffers a performance loss due to nearby secondaries that are concurrently transmitting. Some systems mitigate the effects of self-interference by using a medium access control (MAC) scheme in which neighbors take turns transmitting [44]. We will now consider such a system.

Specific model details are given in Section 10.5. In essence, we ban all but one secondary transmission within a certain area called the *MAC footprint*. The *capacity per area* is then the transmitting link's rate divided by the MAC footprint. This is shown in Figure 2.10 where we again see a large difference in the achievable data rates based on the range.

(a) 1km transmission distance　　　　　　　　　(b) 10km transmission distance
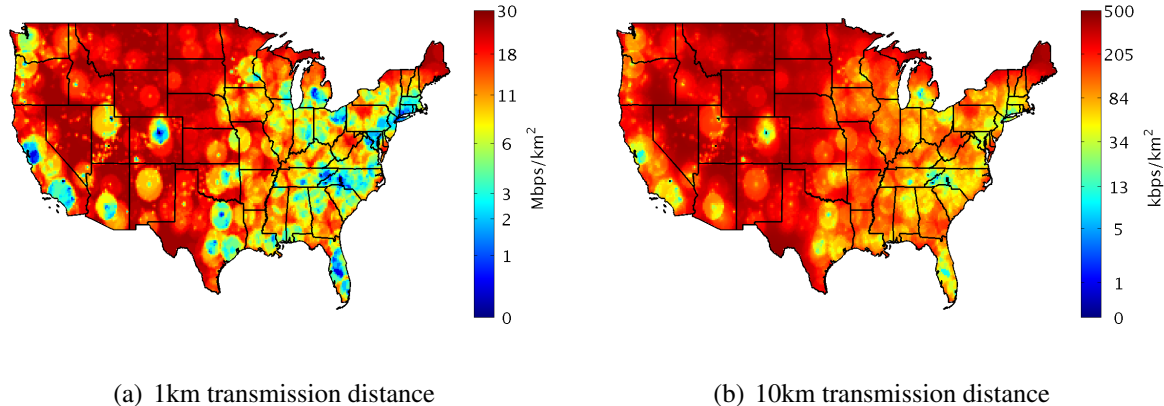
Figure 2.10: Capacity per area using MAC scheme.

Further, we can consider the effective *capacity per person* by dividing the capacity per area by the local population density. The population density[13] is shown for reference in Figure 2.5 and the capacity per person is shown in Figure 2.12.

In Figure 2.13, we see the associated CCDFs for both the MAC model and the single-link model. The single-link capacity is clearly too optimistic as its rates greatly exceed those of the MAC model.

---

[13]Details on population data and calculations can be found in Section 10.2.
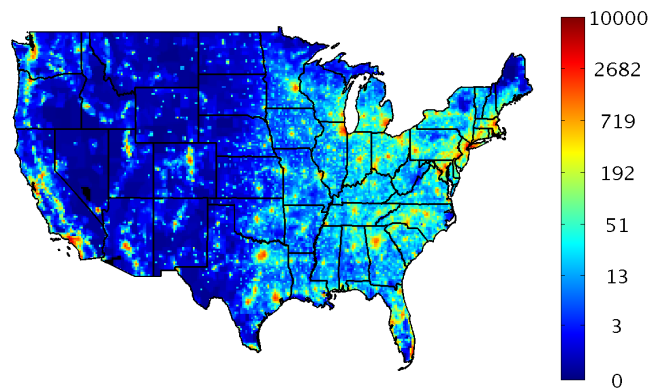
Figure 2.11: Population density (people per km^2) [10, 11]
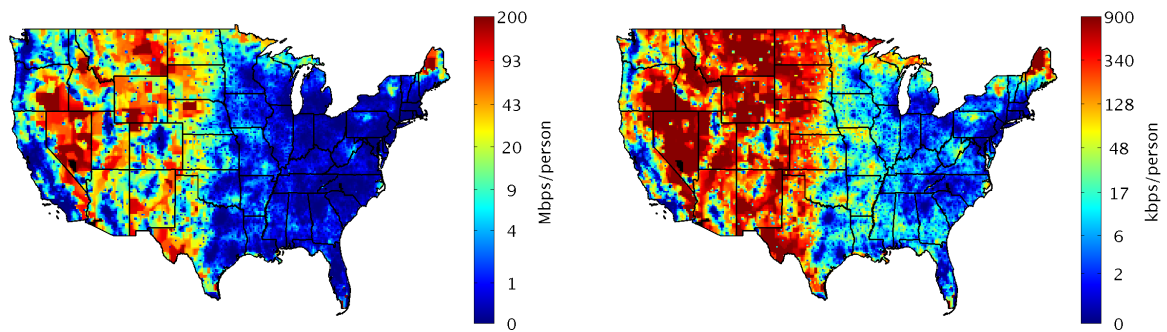
(a) 1km transmission distance

(b) 10km transmission distance

Figure 2.12: Capacity per person using MAC scheme.
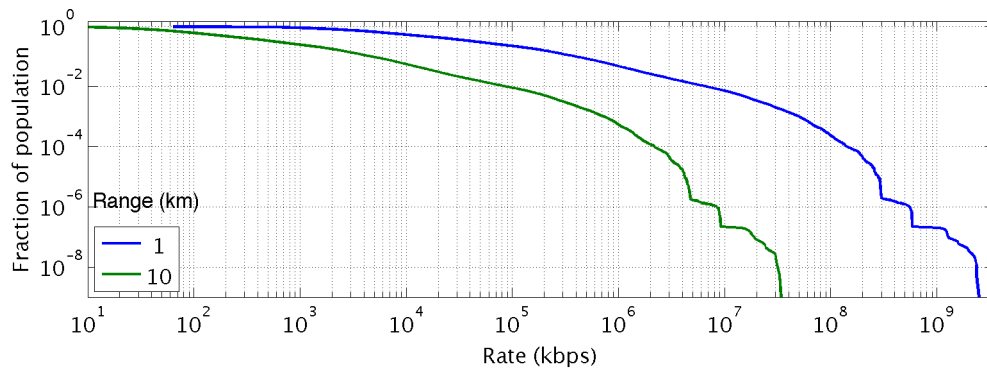
Figure 2.13: Distribution of rate per person in the MAC model with different transmitter ranges

These maps and CCDFs also highlight the spread of achievable rates. In particular, we see extremely high rates in low-population areas and much lower rates in high-population areas. Recall from Section 2.2 that urban regions in general have fewer channels available for secondary use and the channels that are available are often noisier than in rural regions. Furthermore, the few resources available must now be split among many users whereas rural areas have many resources for few users. From these observations it seems that the TV whitespaces might be best suited for rural use. The following section addresses one of the potential flaws in this reasoning.

2.6   Range tied to population: a qualitative shift

While some applications are immune to changes in population density, many are not. Consider a walkie-talkie example, an application which one may think of as fixed range. Suppose person A wishes to communicate with person B, where person B is chosen at random from the $p = 2000$ people nearest to person A (we tend to talk to our neighbors). In a city, person B is likely to live closer to person A than if they lived in the countryside and this is simply due to the relative population densities. Thus range is often tied to population density rather than being fixed at 1 km or 10 km. Specifically, we assume that one user in the pair is at the edge of the neighborhood containing p people and the other is at the center, as shown in Figure 2.14.
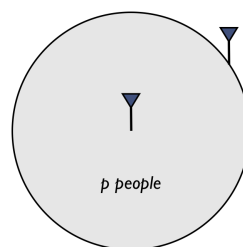
Figure 2.14: Single-link model where range is tied to population

Under this new model, low-population areas have much longer ranges than high-population areas. Figure 2.15 shows a qualitative shift: although rural areas generally have more available spectrum, their rates are typically lower than those in urban areas because of the longer ranges. In previous models where the range was constant, the situation was reversed: rural regions performed better than urban regions due to their extra spectrum.
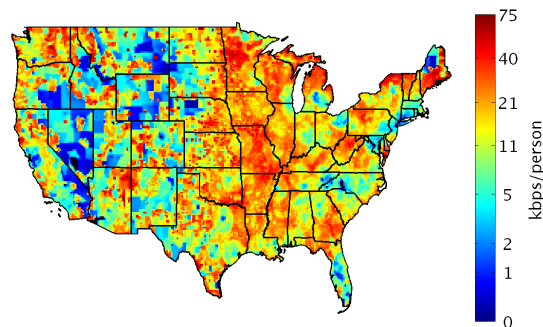
Figure 2.15: Capacity per person using MAC scheme with range tied to population density ($p = 2000$)

Figure 2.16 shows how the size of the neighborhood, p people, changes the mean and median data rate in the single-link and MAC models. Intuitively the results make sense: the closer your communication partner, the higher your rate will be.

We discuss another important wireless model, the cellular model, in the following section. In this model, a service provider builds towers which will service non-overlapping regions. Each is sized to contain p people under the assumption that p customers are required to recoup the cost of providing the service.

2.7 Cellular models

In addition to point-to-point links such as the walkie-talkie case, whitespace users may be interested in building a cellular architecture, for example as a wireless Internet service provider. We will show that the overall effect on data rates is similar to that of the MAC model.

While the MAC model captures the variety of ranges seen throughout the nation, it misses the variations in range on the smaller scale. In particular, a tower will not be communicating at a fixed range to its users, rather it will have a cell full of users with which it needs to communicate. This is
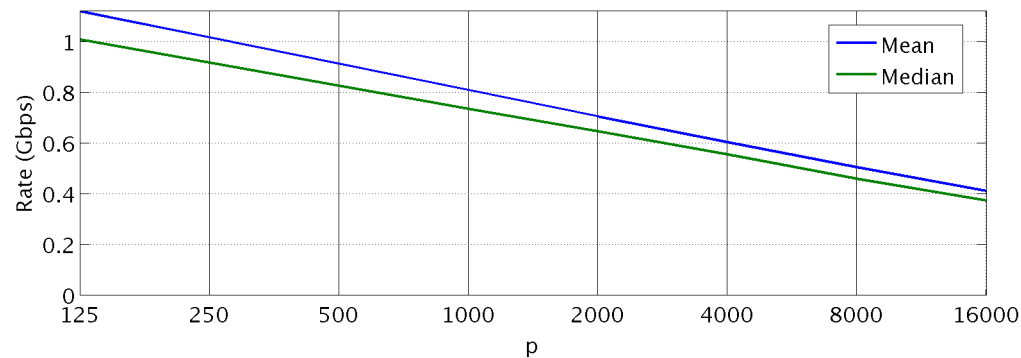
(a) Single-link model
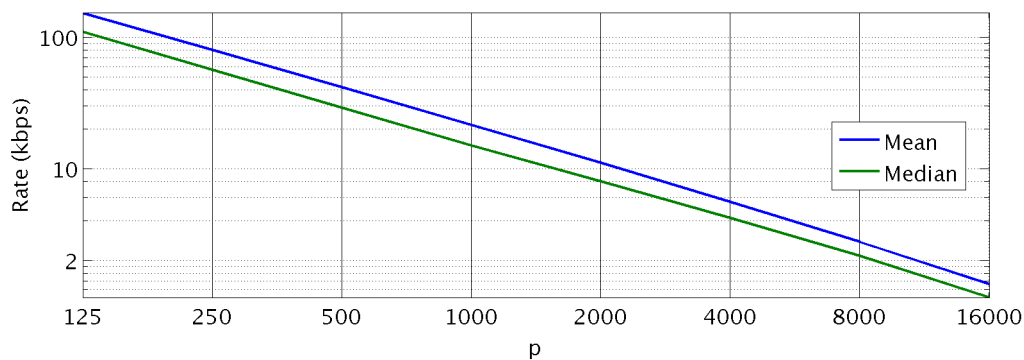
(b) MAC model

Figure 2.16: Effect of tying range to population in the single-link and MAC models.

illustrated in Figure 2.17(a) where the tower is at the center of the cell and the users are uniformly scattered within the cell[14].

We consider a downlink-only model in which each tower transmits to each of the users in its cell in turn and at a fixed power. Each receiver will experience interference from neighboring cell towers[15] as shown in Figure 2.17(b).

The cellular model can be viewed as similar to the MAC model but with a fixed footprint. However, there is one important difference: secondaries in the MAC model can choose to accept any SNR[16], but the worst-case secondary at the edge of the cell almost always receives an SNR of approximately one-half regardless of the cell size and power used. This is because his predominant sources of noise are the two cell towers nearest him whose received signal strengths each match

---

[14]We cap the cell size to $\pi \cdot 100^2$ km^2.

[15]All neighboring cells are assumed to be the same size which is based on the assumption that population is locally constant. As discussed in Section 10.7, this is not correct in general.

[16]In cases where the pollution is too high, this may actually maximize the capacity per area.
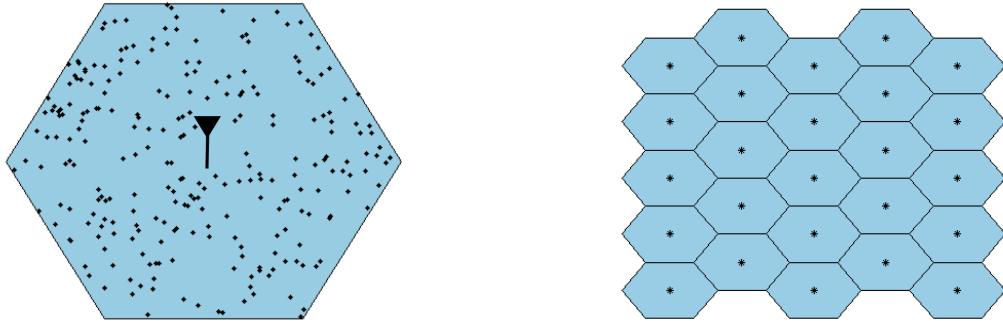
(a) User locations in one cell. The base station is located in the center of the cell.

(b) Cell arrangement and tower locations.

Figure 2.17: Hexagon model illustrations.

the strength of his desired signal. If P_r is the received power of each and N is the noise from the primary transmitters, we can write his SNR as

$$\text{SNR} = \frac{P_r}{N + 2P_r} \approx \frac{1}{2} \quad \text{if } N \ll P_r$$

Following the spirit of the previous section, each cell is sized to fit p people[17], thus high-population areas will have smaller cells which means both a shorter transmission distance but also greater self-interference. Conversely, low-population areas will have larger cells which means a longer transmission distance and less self-interference. We see this nationwide in Figure 2.18. As in the MAC model, we see rates in rural areas collapsing compared to rates in urban areas.

The effect of changing the number of people per tower, as shown in Figure 2.19, is also similar to that in the MAC model.

---

[17]The idea behind this is that it takes p people to fund a tower's construction and maintenance. Back-of-the-envelope calculations suggest that $p = 2000$ is a reasonable number if we assume the following:

- Four people per household

- Households are willing to pay $15/month for the service

- One-hundred percent market penetration

- $50,000 per year is required to build and operate the tower
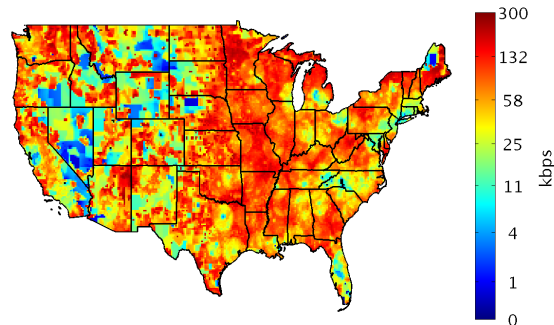
Figure 2.18: Hexagon model, $p = 2000$
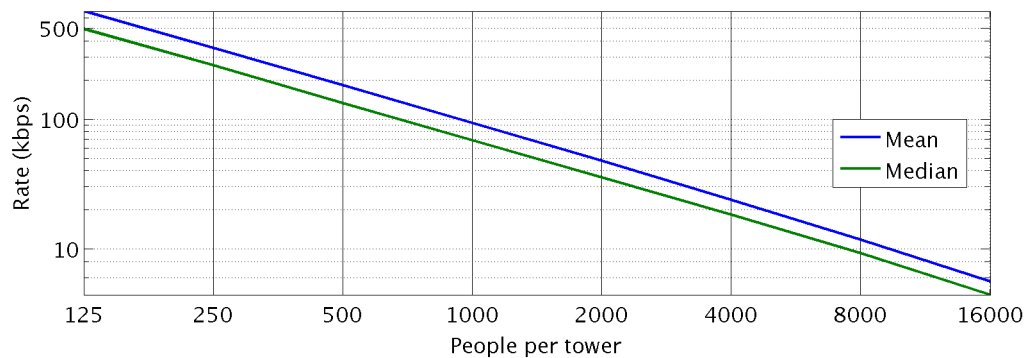
Figure 2.19: Effect of number of people per cell in the hexagonal-cell model

These figures give some insight into the secondary economy. In particular, they can be used to assess the viability of a particular business model and deployment strategy given the expected market penetration. One common option for reducing tower costs is to "piggy-back" on existing towers by paying their owner for the right to use their facilities rather than building a separate facility. This reduced cost would allow for smaller cells to be built, thus increasing the utility for the users of the system. In future work, we would like to assess the opportunity available in such an arrangement.

2.8 Conclusions

This chapter has taken a step toward quantifying the TV whitespaces opportunity and presenting some of its challenges. We began by showing that high-population areas have fewer channels than low-population areas which means that there will be less spectrum precisely where it is needed.

Despite this effect, at least 87% of the US population has enough spectrum to operate a wireless router using the IEEE 802.11 standard which requires 20 MHz. However, secondary systems are subject to substantial noise from the TV towers ("pollution"), making this straightforward channel-count misleading. We suggest using achievable data rate rather than a spectrum count in order to more accurately quantify the whitespace opportunity. Using this metric and a simple single-link model, we then showed that although this pollution is troublesome, exclusions make the biggest difference in achievable secondary rates.

The secondaries' own system parameters also play a large part in determining their achievable rates. For example, we showed that an erroneous choice of tower height may result in a 50% decrease in rate on average. However, extending the transmission range can have a much greater effect, decreasing rates by up to two orders of magnitude with only a tenfold increase in range.

In addition to pollution, whitespace devices must also cope with interference from other secondaries. We first implemented a MAC sharing protocol which allowed secondaries to increase their individual rates via cooperation and time-sharing. The results from this model show that the single-link rates are overly optimistic.

The results from the MAC model also suggested that whitespace devices might be better suited to rural regions which have more spectrum and fewer people with whom to share it. However, we argued that, rather than being fixed, range is often tied to population density–even for a point-to-point link. Under this model, we saw a qualitative shift: urban areas now outperform rural areas despite their comparatively low amount of spectrum.

Finally, we developed another sharing method, the cellular model. Under this model, each secondary transmitter serves the nearest p users. Random user placement captures the variation in range on a smaller scale. Since cells are sized to hold p users, the variation in cell size captures the variation in range on a large scale.

Chapter 3

Power scaling: increasing utility

Some or all of the work in this chapter appeared in the author's Master's thesis, [38].

In this chapter we show that due to interference aggregation the current FCC regulations may not be sufficient to provide the protection to primary receivers that they guarantee. Interference aggregation is the term which describes the cumulative effect of unwanted signals, in this case those from secondary transmitters, at the primary receiver. Since these regulations were designed to be safe in the presence of a single secondary transmitter[1], they fall short when multiple secondaries are operating. We will show the magnitude of this problem and suggest a remedy in the form of an enforced maximum power density, described in detail below.

We further show how this simple solution to the problem — made possible by the existence of databases — can actually be used to help solve another problem faced by the FCC: that of favoring one use case over another. We describe this tension below and offer a potential solution.

---

[1]From the FCC's 2008 ruling in which they develop the separation distance requirements:

"In developing the table of separation distances, we believe it is desirable to minimize complexity for compliance. In this regard, we have balanced this goal of simplicity with the need to provide assurance that TV services will be adequately protected. Given that the power of fixed TVBDs will be limited to 4 watts EIRP, the most important variable in determining the separation distance between a particular TVBD and a TV station's protected contour is the height of the device's antenna above ground. For example, using the FCC curves in Section 73.699 of the rules and the D/U protection ratios specified above, we find that a transmit antenna at a height of 30 meters transmitting with 4 watts EIRP could cause co-channel interference to a TV receiver with an antenna 10 meters above ground at a distance of 14.4 kilometers and adjacent channel interference at 0.74 kilometers. For transmitting antennas at lesser heights, the FCC curves do not provide usable data, so the Okumura propagation model is applied. Using that same transmit antenna at less than 10 meters above ground, interference could be caused by a TVBD to a TV receiver at a distance of 8.0 kilometers to a co-channel TV station and 0.1 kilometers to an adjacent channel TV receiver. A similar calculation applied to a TVBD antenna at 3 meters above ground level calculates that interference can be avoided if separation distances of 6.0 kilometers and 0.1 kilometers are maintained for co and adjacent channel TV stations, respectively." [1, ¶181]

This is reaffirmed in the FCC's 2010 ruling:

"We affirm our decisions regarding the protection contours for TV stations. First, we decline to change the method that must be used to calculate TV station protected contours." [9, ¶21]

3.1 Good intentions and failures: interference aggregation

The regulations chosen by the FCC reflect a desire to use light-handed regulation for secondaries so as not to unnecessarily constrain their design. However, while the current rules do provide a remarkable amount of flexibility, it is also imperative that they maintain the quality of service that they guarantee to primaries. Due to the nature of the rules, the current limits do not provide the necessary level of protection. Interference from secondaries aggregates at primary receivers, potentially causing an outage. This is especially a concern in urban areas where the secondary device density is likely to be much higher.

To see the effect that interference aggregation has on a single TV tower's protected region, we create a toy world in which only one TV tower exists. This tower has a 500 meter height above average terrain (HAAT) and is transmitting on channel 21 with an ERP of 100 kW. We then place secondaries around the TV tower while respecting the FCC's rules. We assume that there is one secondary transmitter per 500 people and a population density of 379 people/km^2 (the median population density — by population — in the United States). Even though all secondary transmitters are obeying the FCC's rules, we see in Figure 3.1 that this actually shrinks the protected region's radius by over 10 km from 108.5 km to 97.4 km. This means that even those TV receivers which are well within the FCC's *protected region* may face too much interference to be able to receive TV.

| (a) | (b) |
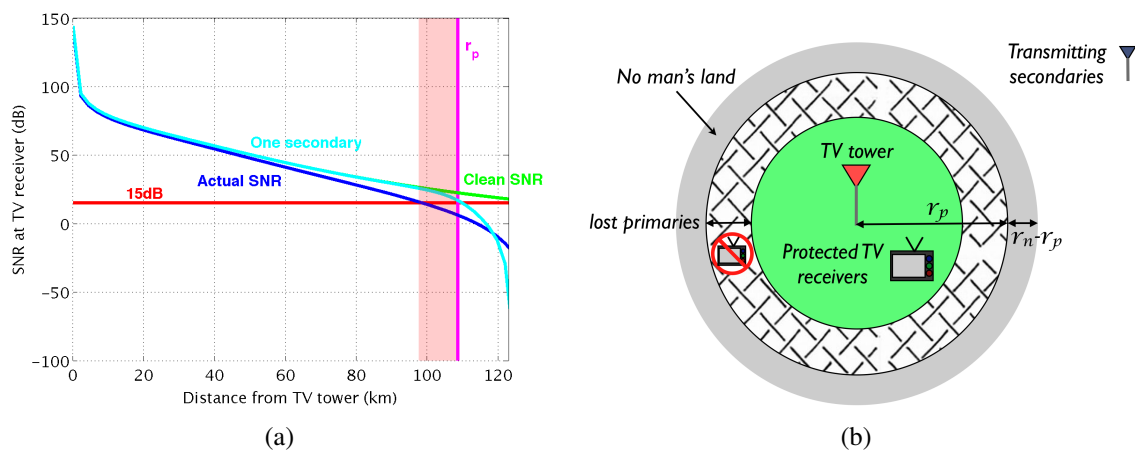
Figure 3.1: Lost area for one tower

To verify that this toy is not simply a pathological example, we perform a similar calculation for actual TV towers [28] using real population information [10, 11]. First, we find the magnitude of the maximum-strength TV signal for each pixel[2]. We then find the secondary interference

---

[2]Note that we ignore the possibility of both constructive and deconstructive interference from other TV signals. That is, we do not add TV signal strengths together, nor do we count the non-dominant TV signals as noise. The truth lies in one of these scenarios but the answer is different for each set of TV towers and we have not yet developed a way to determine which case we are in.

level at each pixel given that the secondaries are obeying the FCC's rules and that there is one 1-Watt transmitter for every 40 people (think of wireless routers)[3]. These two pieces of information combine to form the signal-to-noise ratio (SNR) for TV receivers across the nation. We can then compare the areas with an SNR greater than 15 dB[4] to those areas inside of r_p. The difference is shown in Figure 3.2.

Note that we have shown only the locations that were guaranteed reception under the FCC's rules but have potentially lost reception; that is, we're omitting those locations that lost channels simply due to the FCC's ruling (which will be discussed further in Section 5.1).

It is also important to understand that our nationwide model underestimates the amount of interference experienced at the TV receiver because of the scale: the large distance between pixels means that effect of local interference is severely understated. Since transmission powers attenuate very quickly with distance, local interference can actually account for a large portion of the total interference. Thus our estimates are much too conservative and need to be improved in future work.
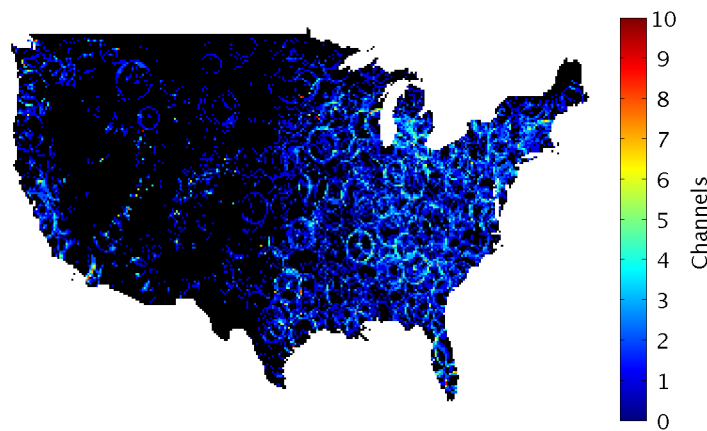
Figure 3.2: Number of channels at risk due to bad rules, $p = 40$

---

[3]We assume that all secondaries are transmitting all of the time. If you don't believe that that many devices would be active at the same time, consider the potential for hacking. Think of something like a wireless router for market penetration.

[4]We assume that 15 dB is the decodability threshold for TV signals [45]. Note that r_p is not explicitly designed by the FCC to have an SNR of 15 dB in the presence of secondaries.
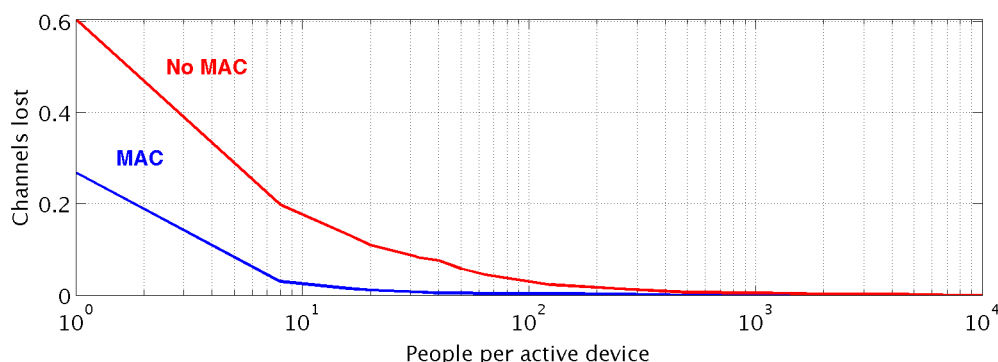
Figure 3.3: Conservative estimate of the average (by population) number of *protected* channels lost as we vary the popularity of secondary devices

We have discussed in previous sections that secondary devices might find it desirable to do some sort of self-regulated power density (e.g. using a MAC exclusion model or a minimum cell size). For example, two devices which are very near to one another would both prefer to take turns transmitting rather than have to deal with the other's noise. At some distance this is no longer true and the two devices will choose to transmit at the same time rather than take turns. This distance effectively sets the local power density in the MAC model.

However, this self-regulated power density is not sufficient to protect primaries, as shown in Figure 3.3 by the "MAC" line[5]. This stems from the fact that the self-regulated power density will not necessarily be the same as the "safe" power density.

The simplest solution[6] to this problem is to increase the separation distance $(r_n - r_p)$, but as Figure 3.4 shows, the population density in the United States varies too greatly for a one-size-fits-all rule to work well. In order to work, it would have to plan for the worst-case scenario. Since the difference between the 10th and 90th percentiles is two orders of magnitude, this would require serious compromise for much of the country.

It is impossible to anticipate the prevalence of whitespace devices, so we cannot design rules based on an expected device density. Instead, we must regulate a *power density*. We will illustrate this concept with a simple example. Consider two devices, A and B, which are co-located. Suppose device A is transmitting at the maximum safe power, P. When device B begins transmitting, their

---

[5]In this case, we assume that devices must be separated by at least 200 meters in order to simultaneously transmit; that is, there is one active device per 0.0314 km^2. This is a reasonable distance for applications such as a short-range wireless service.

[6]It became apparent during the writing of this thesis that there is in fact another simple solution: choosing the separation distance $r_n - r_p$ for each tower individually (and perhaps dynamically) based on local characteristics rather than using the same separation distance for all towers. This would require minimal changes to the functionality of the databases and thus may be a more desirable solution to the problem of interference aggregation. The impact to secondaries of such a rule is a topic of future work.
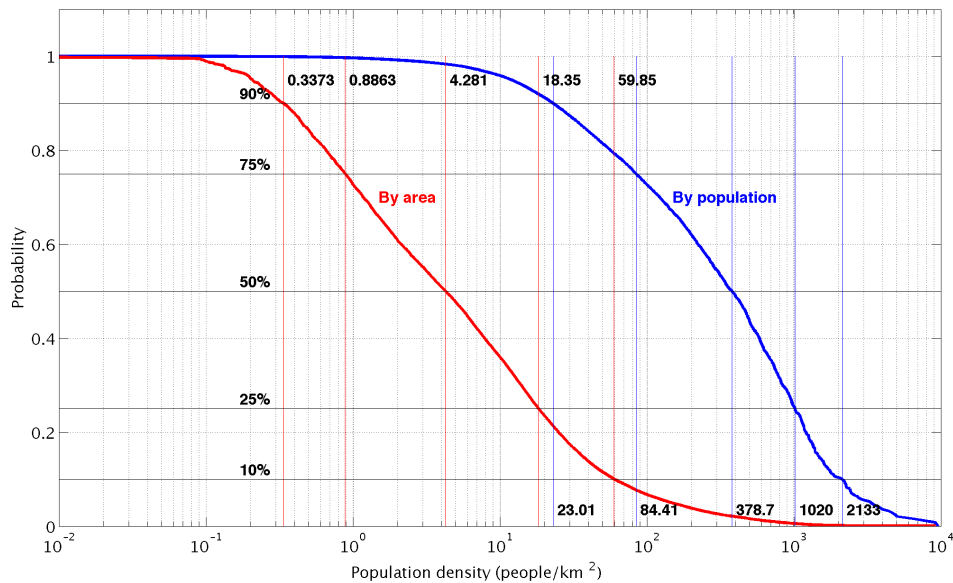
Figure 3.4: Population density of the United States, sampled by population and by area

combined power can be no greater than P in order to maintain the power density and thus primary safety. As new secondaries turn on, nearby secondaries must adjust their power so that ultimately the rules are obeyed.

De Vany et al. [15] expressed very similar ideas with regards to how to control interference from secondaries. In their paper, they suggest that rights be defined in terms of an area in which the secondary is allowed to transmit. Furthermore, it shall be his responsibility to guarantee that his aggregate transmissions are below a certain threshold outside of his area. Most importantly, he suggested that anyone proposing to divide their transmission area (perhaps in order to sell part of it) would need to observe the original rules. That is, the worst-case scenario to neighboring areas remains the same regardless of whether or not the area is partitioned. In the extreme case where these property rights are partitioned into infinitesimally small pieces, we arrive at the notion of power density.

There are two ways that the power density can be controlled:

1. A database gives a local power limit to a set of devices which control their powers collaboratively. Although harder to achieve, this decentralized solution will use spectrum more efficiently.

2. A database gives a power limit to individual devices. This has the advantage of minimizing device complexity and interoperability. However, it puts a large burden on the databases and relies heavily on propagation models [46].

Notice that both schemes require the various database providers to communicate with one another in near-real-time. However, this is not as imposing as it may seem: intercommunication between databases is already a requirement[7] due to the need to exchange time-varying registrations such as those for wireless microphones[8].

Both power-control methods are subject to the problem of an untruthful secondary which can either request more power than necessary or underreport its usage for malicious purposes. We do not consider these enforcement issues here. In our model, a region is given a power density and we assume that it perfectly obeys this mandate.

We took a rule of this nature and applied it to the United States in Figure 3.6(a). We found the greatest safe uniform power density for each channel assuming that we are following the FCC's regulations regarding exclusions and that TV receivers at r_p can tolerate interference from secondaries that is the same as thermal noise (that is, we assume that a receiver at the FCC's r_p has an SNR of 18 dB in a clean channel and that 15 dB is required for TV reception [45]). In our model, each secondary is given power according to his footprint (in a traditional cellular system, this may be the coverage area of the cell). In particular,

$$\text{power} = \text{area of footprint} \times \text{power density}$$

First, we show in Figure 3.5 that using this power density is safe. Notice that we do lose a few channels when we compare the areas that receive TV to the FCC's protected regions. However, our models assumed that it was safe to erode 3dB despite the fact that the separation distance $(r_n - r_p)$ is not sufficient to guarantee this in some cases. Those stations for which this separation was inadequate will experience some outages in this model. Future work will incorporate this varying level of acceptable interference.

Figure 3.6(b) shows for comparison the rates for secondary devices under the FCC's current rules. While urban users receive roughly the same quality of service, rural users tend to do better because their larger cells allow them to use a power proportional to their area, thus helping to make up for their abysmally long ranges.

For a constant power density, increasing the area of a cell does not scale the allowed transmit power fast enough to offset the growing distance to the receiver. Consider an example of a cell which has radius r with a transmitter at the center communicating to a receiver at the edge of the cell, distance r away. Using a standard inverse-power-law pathloss model, the power necessary to maintain a constant rate to this receiver in a clean channel will increase with r^α where $\alpha > 2$. However, the transmit power is scaling only with respect to the cell area under our power density rule and thus

---

[7]*"The Commission further required that, if multiple database administrators are authorized, the database administrators are to cooperate to develop a standardized process for sharing data on a daily basis, or more often, as appropriate, to ensure consistency in the records of protected facilities."* [9, ¶95]

[8]Events may register their venue as a temporary excluded region if they deem that the existing channels are not sufficient for their purposes.
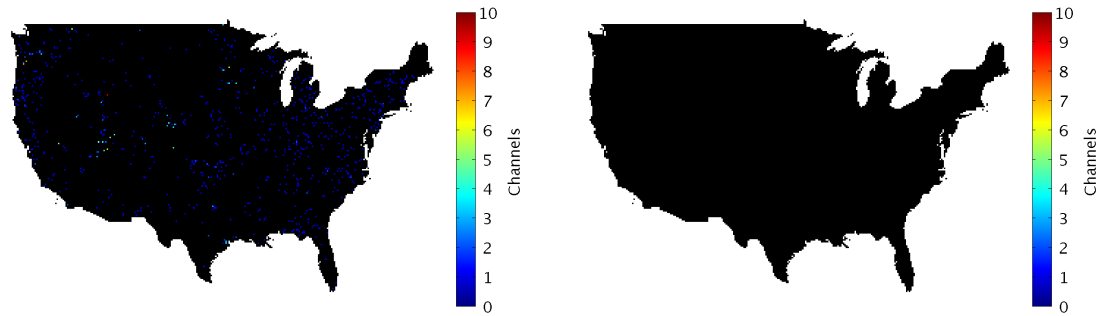
(a) FCC's protected region

(b) Regions which have an 18-dB SNR without the presence of secondaries

Figure 3.5: Number of TV channels potentially lost when using a uniform power density and two different protected regions.
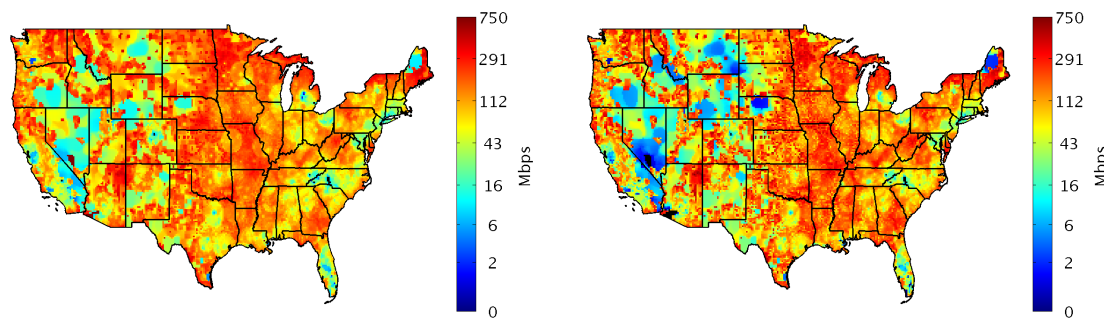
(a) Overly-cautious but simple power rule

(b) FCC's current rules

Figure 3.6: Rates available to secondaries under different power rules ($p = 2000$)

is proportional to r^2. Therefore scaling the power based on area alone is not sufficient to maintain a set link quality in large cells. This suggests that we need to find another way of increasing the power available to rural users. We will address this problem in the next section by *safely* increasing the power density in rural areas.

3.2 Analysis: candidate rules

While it may seem straightforward to allow users far from TV towers to transmit with higher powers than those near TV towers, finding a *safe* and *fair* power scaling rule requires some thought. In this section we first show why the obvious solutions do not work and then offer a potential solution.

3.2.1 Why is power scaling even possible?

First, it is important to note that databases are the key piece of technology which enable us to realistically discuss power scaling schemes. Until now, unlicensed devices were certified for compliance with a uniform power limit–the only feasible option–and then deployed without any need for run-time adaptations. The ever-changing whitespaces (e.g. time-varying primary transmissions, event-specific wireless microphone exclusions) make such adaptability all but necessity.

Some devices will attain this adaptability via sensing, but we believe that many will make use of databases because they increase the *usable* whitespace opportunity [47,48]. If we accept that these devices are already contacting the database daily[9] in order to receive a list of available channels, it is not outrageous to assume that they could also be assigned a varying maximum power limit.

In fact, Karimi of Ofcom has already suggested that this be part of the database capabilities [49]. Further, the Electronics Communications Commission of Europe has already suggested that rudimentary power scaling — but not power *density* scaling — may be advisable for maximum spectral efficiency and has provided a table of potential EIRP limits which vary as a function of the separation distance [46, §9.1].

3.2.2 Toy example: why doesn't straightforward power scaling work?

A nave approach to power scaling would be to allow each secondary, regardless of its distance to r_p, to cause the same amount of noise at the protected TV receiver. We can quickly reject this scheme by realizing that the eroded fade margin[10] would need to be split between a potentially very large number of secondaries, leaving each with only a small amount. Furthermore, this intuitively scales the power too quickly (it will scale as the inverse of the pathloss), thus giving regions far

---

[9]*"The Commission required fixed and Mode II TV bands devices to re-check the database, at a minimum, on a daily basis to provide for timely protection of wireless microphones and other new or modified licensed facilities."* [9, ¶95]

[10]The eroded fade margin is defined in Section 5.2. Briefly, it defines the total allowed noise that may be caused at the TV receiver from all secondaries.

from the TV tower much more power then they need[11] while taking it away from nearby regions. Clearly we need a fairer way of dividing the margin. We will do this by looking at what each region "dreams" of having.

We refuse to allow secondaries to be wholly unrealistic in their "dreams" by pretending that they live in a world all by themselves. Instead, we take the viewpoint of a secondary who understands that r_n represents the boundary for secondary transmissions but that the value for r_n may be negotiated. Indeed, each person's dream is that they are living at the edge of the no-man's-land (r_n). In that way, they need not sacrifice power in order to allow anyone "in front of" them and they are allowed to use the maximum safe power. However, we cannot fulfill this dream for every user at the same time. We use the following toy model to explain why this is true.

Consider the toy one-dimensional[12] world illustrated in Figure 3.7. There is a primary transmitter, shown in red, and a primary receiver (i.e. television set) shown in green. Since this receiver is at the edge of r_p, optimality of the rules implies that he should be on the verge of losing reception when the secondary system is fully loaded. For this example, we will assume that this means that the aggregate secondary interference cannot be more than T, the amount of noise in a clean channel[13]. Thus the secondaries may collectively cause 3 dB of interference.
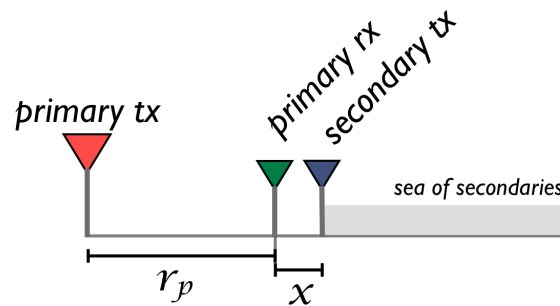
Figure 3.7: Toy one-dimensional world with one primary transmitter, one primary receiver at r_p, and a multitude of secondaries, the foremost of which is at $r_p + x$.

We use a standard theoretical model for the attenuation of a transmitted signal. In particular, if the transmitted power is P, then the signal strength distance r away is $P \cdot r^{-\alpha}$ where $\alpha > 1$ (in a two-dimensional world we would require $\alpha > 2$).

---

[11]Due to the saturated nature of the log function which describes their data rate, increasing the power has diminishing returns. Furthermore, devices which are in the interference-limited regime find that increasing their power is only marginally helpful: increasing their power necessarily increases that of their neighbors which has the net effect of decreasing the relative noise power from the TV.

[12]The forthcoming results can be generalized to two dimensions but for clarity we have used one dimension.

[13]Even a clean channel has "noise," called thermal noise.

Consider a secondary transmitter which is at distance x from r_p, shown in blue in Figure 3.7. His dream is that he is the closest secondary transmitter to r_p, thus the power density is maximized for him. We will call his dream power density $P_{\text{dream}}(x)$. The aggregate interference must be less than or equal to T, and thus we have the following condition:

$$
\begin{aligned}
T &= P_{\text{dream}}(x) \int_x^\infty r^{-\alpha} dr \\
&= P_{\text{dream}}(x) \frac{x^{-\alpha+1}}{\alpha - 1} \\
\implies \quad P_{\text{dream}}(x) &= T(\alpha - 1) x^{\alpha - 1}
\end{aligned}
$$
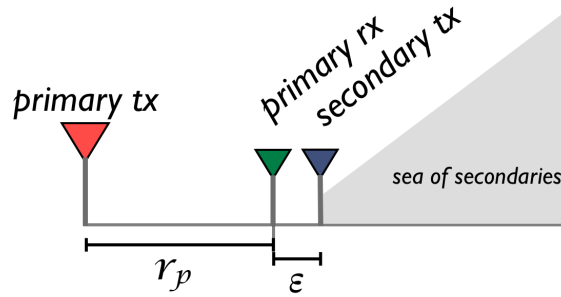
Figure 3.8: Dream power scaling for locations outside of $r_n = r_p + \epsilon$ (power density is shown on logarithmic scale).

While choosing this uniform power density will be safe under the secondary's assumption, it does not resolve the tension between the rural and urban users[14] since we must choose a particular x' to favor. To see why we can not give all their locations their dream power, suppose we chose $x' = \epsilon$ and let each secondary transmitter at location x_i use power density $P_{\text{dream}}(x_i)$ as shown in Figure 3.8. In this case, the interference at the primary receiver is unbounded:

$$
\begin{aligned}
\int_\epsilon^\infty P_{\text{dream}}(r) \cdot r^{-\alpha} \cdot dr &= \int_\epsilon^\infty \left[T(\alpha - 1) r^{\alpha - 1} \right] \cdot r^{-\alpha} \cdot dr \\
&= T \cdot (\alpha - 1) \cdot \int_\epsilon^\infty \frac{1}{r} \cdot dr \qquad (3.1) \\
&= \infty
\end{aligned}
$$

The calculations above show that we cannot simultaneously give everyone everything that they want. Similar results were shown by Hoven [50] but he did not suggest a solution. We will now suggest a candidate rule to overcome this challenge in a way that is fair to urban and rural users alike.

---

[14]For our purposes, we assume that TV towers are located at population centers. That is, urban users are those nearest to the TV tower while rural users are those furthest away.

3.2.3 Toy example: a candidate rule for fair power scaling

The previous section demonstrated why we cannot give all secondaries their "dream" power density simultaneously. We suggest the following fair compromise: if a secondary would have received rate R using his "dream" power density, we allow him to use the power necessary to achieve rate $\gamma \cdot R$ where $\gamma < 1$ is a tunable parameter which is used to ensure primary protection. Because of the logarithmic nature of the rate function, this results in a power density scaling which grows at a rate slower than $x^{\alpha-1}$, meaning that the integrand of Equation 3.1 is no longer $1/r$ and thus the interference at the primary receiver is bounded. This section will show these calculations in detail.

If a secondary would have received rate $R_{\text{dream}}(x)$ while using power $P_{\text{dream}}(x)$ in a clean channel, then under the compromise he will receive rate $R_{\text{new}}(x) = \gamma \cdot R_{\text{dream}}(x)$ by using power $P_{\text{new}}(x)$ where $0 \leq \gamma < 1$. The tunable parameter γ can be reduced until the SNR requirement for the primary receiver is met.

As we saw in previous sections, we can use the Shannon capacity formula to find the theoretical rate based on the bandwidth (we use a unit bandwidth in our calculations), the signal power, and the noise power [42]:

$$\text{rate} = \text{bandwidth} \cdot \log_2 \left(1 + \frac{\text{signal power}}{\text{noise power}} \right)$$

Our pathloss model dictates that signal power $= P \cdot d^{-\alpha}$ when P is the transmit power and d is the transmission distance. Notice that in order to calculate a rate, *we must assume a transmission distance, d.* We will come back to this problem later. Furthermore, in order to calculate the transmit power we need to multiply the power density by the footprint or "area" of the link, A. In our one-dimensional world we can think of area as length.

We assume incorrectly but for simplicity[15] that we are operating in a clean channel, thus noise power $= T$.

We will now show that adding the parameter γ does indeed cause the integral to converge. To begin, we calculate the old rate and corresponding new rate[16]:

$$
\begin{aligned}
R_{\text{dream}}(x, d, A) &= \log_2 \left(1 + \frac{A \cdot P_{\text{dream}}(x) d^{-\alpha}}{T} \right) \\
&\approx \log_2 \left(\frac{A \cdot P_{\text{dream}}(x) d^{-\alpha}}{T} \right)
\end{aligned}
$$

---

[15]We would otherwise need to take into account not only the noise due to the primary but also noise from all other transmitting secondaries.

[16]We assume that we are not in the linear region of the $\log_2(\cdot)$ function.

$$
\begin{aligned}
R_{\text{new}}(x, d, A, \gamma) &\approx \gamma \cdot \log_2\left(\frac{A \cdot P_{\text{dream}}(x)d^{-\alpha}}{T}\right) \\
&= \log_2\left(\frac{A^\gamma \cdot P_{\text{dream}}^\gamma(x)d^{-\alpha \cdot \gamma}}{T^\gamma}\right) \\
&= \log_2\left(\frac{\left[A^\gamma\left(T^\gamma(\alpha-1)^\gamma x^{\gamma(\alpha-1)}\right)d^{\alpha(1-\gamma)}\right]d^{-\alpha}}{T^\gamma}\right) \\
\implies \quad P_{\text{new}}&(x, d, \gamma) = A^\gamma \cdot \left(T(\alpha-1)^\gamma d^{\alpha(1-\gamma)}\right) \cdot x^{\gamma(1-\alpha)}
\end{aligned}
$$

Notice that the power now scales with $x^{\gamma(1-\alpha)}$. This means that the integrand is now $r^{-(\gamma+(1-\gamma)\alpha)}$:

$$
\int_\epsilon^\infty P_{\text{new}}(r, d, \gamma)r^{-\alpha}dr = A^\gamma \cdot T(\alpha-1)^\gamma d^{\alpha(1-\gamma)} \int_\epsilon^\infty r^{-(\gamma+(1-\gamma)\alpha)}dr
$$

Thus we require $\gamma + (1-\gamma)\alpha > 1$ for convergence. We easily verify this by noting that $\alpha > 1$, $\gamma > 0$, and

$$
\gamma + (1-\gamma)\alpha > 1 \Leftrightarrow (1-\gamma)\alpha > 1 - \gamma
$$

Further, we can easily see that the aggregate interference condition is met if and only if

$$
A^\gamma \cdot \frac{(\alpha-1)^{-(1-\gamma)}}{1-\gamma} \cdot d^{\alpha(1-\gamma)} \cdot \epsilon^{-(\alpha-1)(1-\gamma)} \leq 1 \tag{3.2}
$$

Using this equation we can find the optimal γ if we can determine A, d and ϵ. However, this still assumes that there are no secondaries closer than ϵ to r_p, which is analogous to the FCC's $r_n - r_p$. We see how γ depends on ϵ in Figure 3.9. We can see already that the optimal choice of ϵ (a.k.a. $r_n - r_p$) is very dependent on the transmission distance, d.

Of course, secondaries closer than ϵ to r_p will not be able to transmit at all on this channel and thus have zero *whitespace* rate. This may seem like a large sacrifice when viewed as a percentage loss, but on an absolute scale it is rather small if we choose ϵ carefully. These excluded secondaries would have a low rate anyway since γ would have had to be almost zero to include them and therefore we have not taken much away from them. Furthermore, devices always have the option of using the ISM bands instead of or in addition to the TV whitespaces. Finally, almost all people have access to at least two channels under the FCC's current regulations, implying they may be able to use another whitespace channel as well.
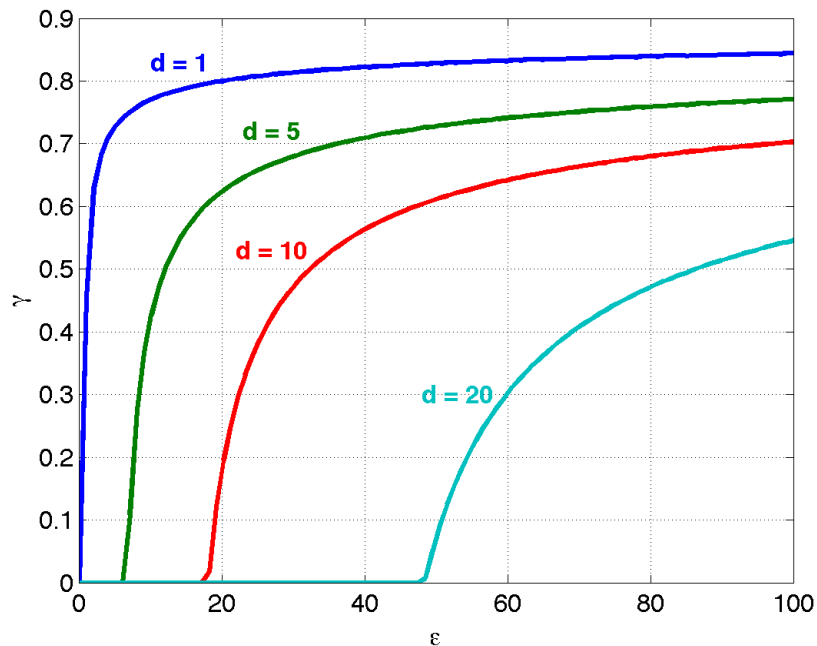
Figure 3.9: Optimal γ as a function of ϵ from Equation 3.2.

In our model, we choose ϵ dynamically (for each tower[17]) using a threshold β as described in Equation 3.3: secondaries whose dream rate does not meet this threshold are not allowed to transmit and secondaries who exceed the threshold are allowed to transmit.

$$\epsilon \quad = \quad \arg\max_{x} x \tag{3.3}$$
$$\text{such that } P_{\text{dream}}(x) \leq \beta$$

Thus we have obtained a (β, γ) approximately-optimal rule: the achieved rate for each secondary differs by no more than a factor of γ or the amount β from the dream rate. We now look at the effect of this rule on a nationwide scale.

3.3 Nationwide evaluation of candidate rule

In this section we will see the results of applying the candidate rule from the previous section to the United States. We will first describe the two models (cellular and hotspot) that we used to test our rule, then we will go on to contrast the two models in terms of dream powers and then show

---

[17]This is done independently for each TV tower and the results are later synthesized into a nationwide map. Further details can be found in Section 10.6.

their realistic rates achieved. Finally, we'll show that TV reception is preserved under both models regardless of the device density.

The power rule developed in Section 3.2.3 is dependent on three variables: transmission range, the separation distance $\epsilon = r_n - r_p$ (set via parameter β as described in Section 3.2.3), and the rate-scaling factor γ. We choose the transmission range via the model as described in the subsection immediately below. below. We set the utility threshold $\beta = 0.5$ bits/second/Hertz since even this dreamed-of rate is not very useful. Finally, we maximize γ for each tower subject to the primary's interference constraint.

The interested reader is highly encouraged to review additional details regarding the generation of the following data which can be found in Section 10.6 .

3.3.1 Two models: cellular and hotspot

Throughout this section we will consider two usage cases: the cellular case and the "hotspot" case. We define them here and henceforth refer to them by name only.

The cellular case has previously been described in Section 2.7. It essentially consists of a hexagonal cell with a transmitting (we consider downlink only) tower at the center and receivers (i.e. users) scattered uniformly throughout the cell. This is the typical cellular model. This is shown in Figure 3.10(b) where gray represents the footprint of the cell and green represents the support for the distribution of the users. For simplicity, we use a single user's range to calculate the rules but evaluate the rules based on all of the points in the cell. In particular, we consider a point approximately half-way out in the cell.

(a) Cellular model   (b) Hotspot model
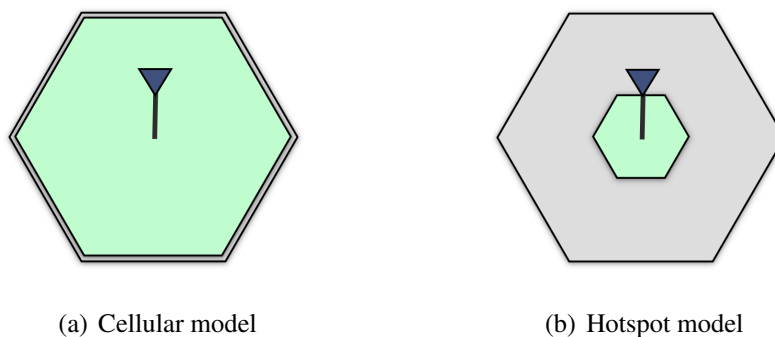
Figure 3.10: Cellular vs. hotspot models: gray represents the footprint of the cell and green represents the area in which users may be. The transmitting secondary tower is located at the center of the cell. Cellular users are scattered throughout the entire cell and hotspot users are clustered near the center of the cell.

The hotspot model differs from the cellular model only in that users are clustered near the center of the cell, as shown in Figure 3.10(b). Regardless of cell size, hotspot users are no further than 100

meters from their base station. One can envision a coffee shop scenario: there is one coffee shop on each block (hence the footprint size is one block) but each shop's wireless Internet signal does not extend beyond its walls. Again, we use a single user's range to calculate the rules but evaluate using all users. This range is 100 meters.

Under both models, each receiver will have a different potential rate which depends on his location in the cell, so in the spirit of fairness we assume that time is shared unequally among users so that each user gets the same overall rate regardless of his position in the cell. This is actually the harmonic mean of the potential rates. We call this the *fair rate* and it is our primary metric for evaluating cellular data rates.

As before, we set the cell footprint to be inversely proportional to the local population. We use the value $p = 2000$ throughout when not otherwise mentioned.

Again, many details have been omitted for the sake of brevity and readability. Please see Section 10.6 for full details.

3.3.2 Powers and rates

In this section we will first show the "dream" powers and rates for each model (hotspot vs. cellular) and then we will show the results using the power scaling rule suggested in Section 3.2.3.

The average (across whitespace channels) dream power[18] is shown in Figure 3.11. Note that the dream power is the same for both the cellular and the hotspot models because it does not make any assumptions about the usage case. As expected, we see the dream power increasing as we go further and further from TV towers. Also, we see a high degree of variation in the dream powers. This further makes the case that a uniform power density would make a lot of places (and people) very unhappy.

We have applied a form of adjacent channel exclusions: we assume that adjacent-channel receivers can attenuate by 50dB[19] and then use the lowest of the cochannel and (potentially) two adjacent channel safe powers.

The rates resulting from using the dream power in each of the models are shown in 3.12. Our model incorporates both pollution and self-interference from nearby secondaries[20]. Note that this utilizes a new method for calculating secondary interference which is detailed in Section 10.6.

As with the uniform power density, rural areas do quite well now because their large cells allow them to use a higher power which is almost enough to make up for their large transmission distances. The difference is that with the power scaling version of the power density rule, they now

---

[18]Actually, we've taken the average not including zero-power channels.

[19]This was inspired difference in separation distances for cochannel versus adjacent-channel protections, presumably because TV receivers can reject adjacent-channel interference by approximately 50 dB.

[20]Here, we no longer assume that nearby cells use the same power. Instead, we use the actual power densities.
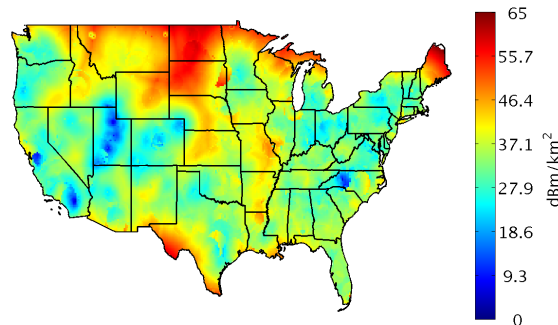
Figure 3.11: Dream power (same for cellular and hotspot models).

do even better than before because of the inverse relationship between population and whitespace channels seen in Section 2.2: fewer people means bigger cells but also fewer TV towers and hence a larger power.

We also notice that the hotspot model does extremely well. This is because the transmission range does not grow while the footprint does, thus it gets the best of both worlds.
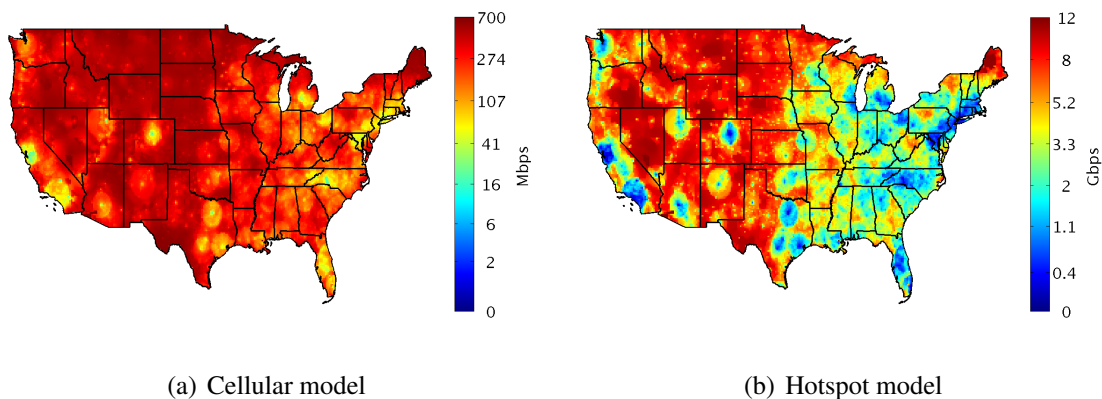
(a) Cellular model

(b) Hotspot model

Figure 3.12: Dream rates

These dream rates are unrealistic because the power density scales too quickly as we saw in Section 3.2.2. After we apply our candidate rule to the cellular model, we see that the resulting rates are about 20-30% lower than the dream rates, as shown in Figure 3.14. This means that most people are getting at least 70% of their dream rates. This is further corroborated by the CCDFs of Figure 3.16 where we see that at least 87% of people get at least 70% of their dream rate. The median

person receives about 87% of his dream rate. And — most importantly — rural regions are doing about as well as urban regions.

The hotspot model naturally fares much better than the cellular model due to its constant short range. It loses only about 10% of its data rate from our candidate rule. Under this model, 96% of people get at least 70% of their dream rate. The median person receives about 92% of his dream rate. Rural regions do better than urban regions because in addition to generally having a greater power density their larger cell size allows them to use more total power.

(a) Cellular model   (b) Hotspot model

Figure 3.13: Actual rates

(a) Cellular model   (b) Hotspot model
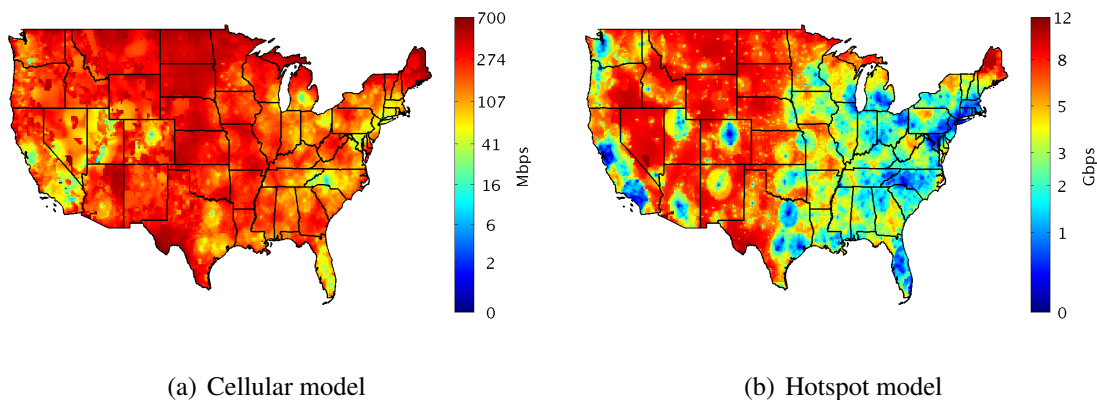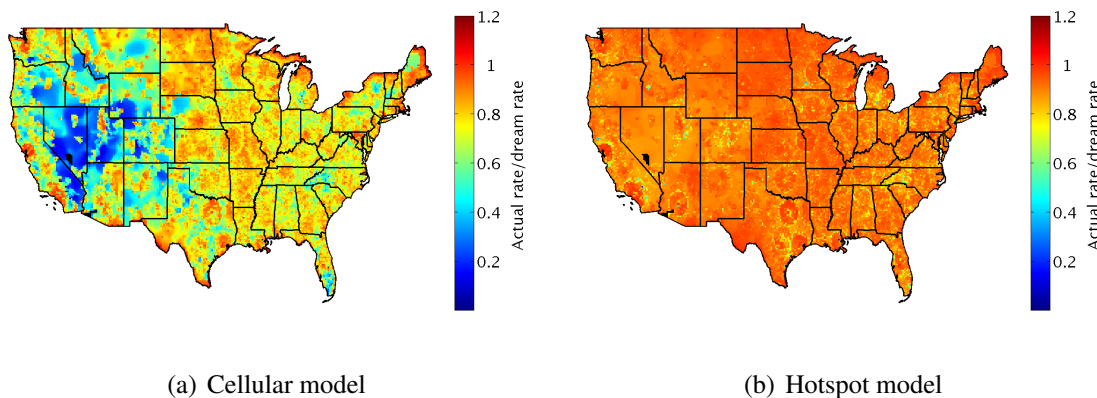
Figure 3.14: Ratio of actual rate to dream rate

As discussed in Section 2.19, we may have different "device popularity" levels, represented by a change in the value of p. We briefly show the effect of changing p on the average rate per person, shown in Figure 3.17. The effect seen here is two-fold: increasing the number of people in a cell
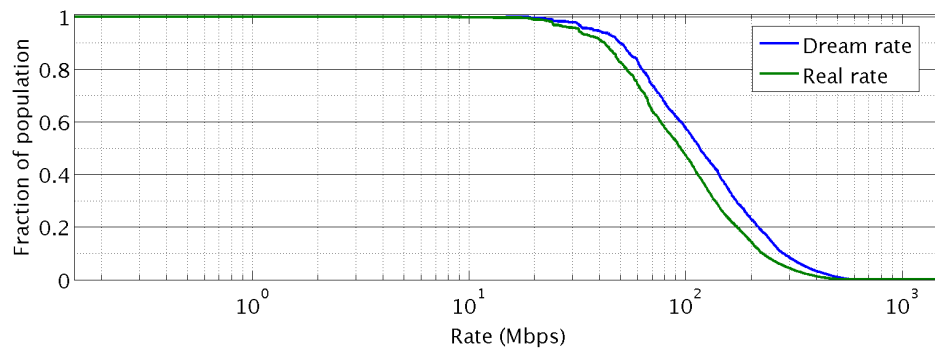
(a) Cellular model
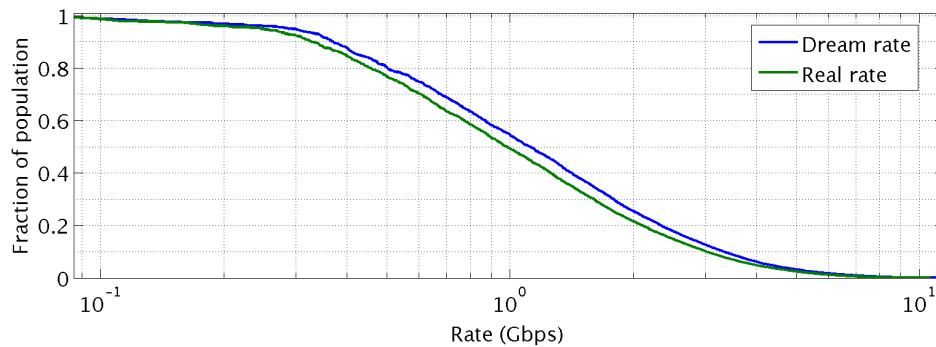
(b) Hotspot model

Figure 3.15: CCDFs by population of dream rates and actual rates

not only increases its area (thus making a set rate harder to achieve) but also that same dismal rate is shared with even more people.

A more interesting result shows up in Figure 3.18 where we plot the total rate per cell as a function of p. The hotspot model experiences an increase in utility because it has an increasing footprint and it does not have to cope with an increasing range. However, the cellular operator sees a downturn in his rates once his range grows too large. Notice however that his dream power is scaling fast enough for his dream rates to be generally increasing.

Finally, we do one last reality check to make sure that our power density rule is safe. This is done in the following section.

3.3.3   Reception for TVs preserved

In this section we simply provide proof that our scaling power density model continues to provide reception to TV receivers. We showed earlier that a uniform power density was safe and Figure 3.19 shows that indeed the scaled power density is also generally safe.
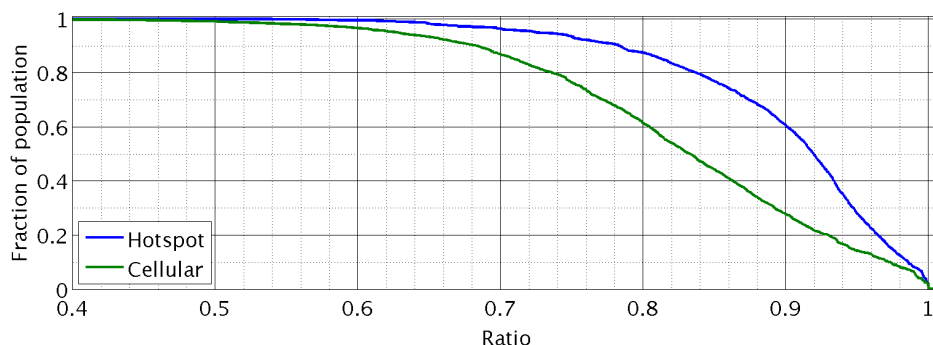
Figure 3.16: CCDF by population of ratio of actual rate to dream rate

We note for future work the problem of investigating why the few at-risk regions exist.

3.4   Criticism of the candidate rule: tension between short-range and long-range applications

Although our candidate rule helps to mitigate the tension between the two usage cases (cellular and hotspot), this section explains why we have not completely overcome this problem. A complete solution is left for future work.

As we will show in Chapter 5, there are many ways in which to choose the protected region (r_p) and the size of the no-man's-land ($r_n - r_p$) for each TV tower. Each choice yields a power density limit, just as the FCC rules did. However, it is intuitively clear that users in urban areas want exactly the opposite of what rural users want. In particular, urban users who generally have a shorter transmission range and access to fewer channels than rural users (as shown in Section 2.2) will be willing to sacrifice some power in exchange for a greater coverage area. On the other hand, rural users are not nearly as hindered by exclusions and if acting purely for self-interest would lobby for a greater no-man's-land in exchange for greater transmission powers.

Unfortunately, our power-density-scaling rule requires us to assume a transmission range d and a footprint A when calculating the "dream" rate which in turn influenced the final power density level. When the mismatch between the expected deployed transmission ranges and footprints is high–in particular, high enough to cause a transition from the saturated region of the rate function (log) to the linear region or vice versa–the deployed devices will have a high amount of regret. Even having excess power is regrettable: receiving more power than necessary in one region necessarily means depriving another region.
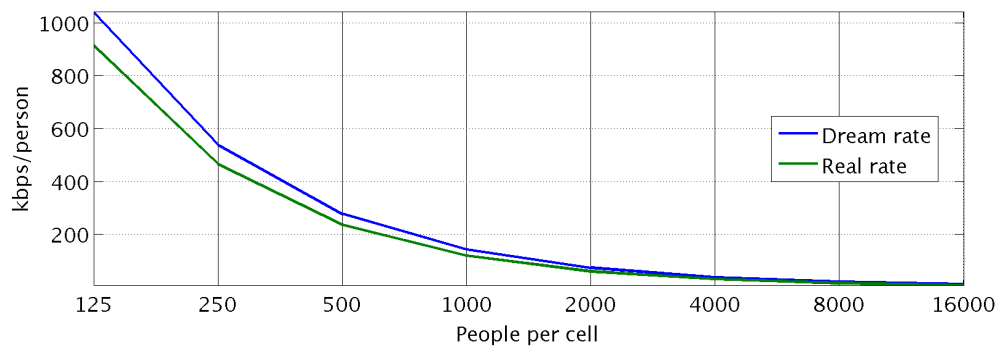
(a) Cellular model
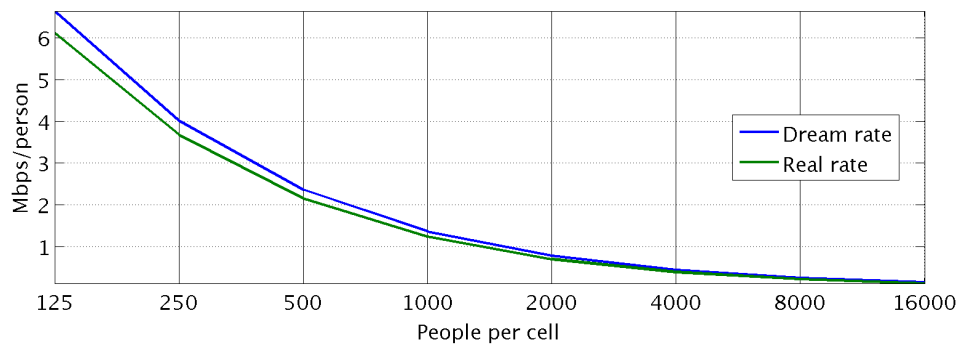
(b) Hotspot model

Figure 3.17: Average rate per *person* as a function of p, the number of people in the cell (the area of the cell is a function of p and the local population density)

It is not obvious how to resolve this problem in a fair way. We suspect that artificially introducing heterogeneity into the power density levels seen across channels at a given location may help: in this way, a device may be able to pick a channel whose power density has been optimized for his usage case.

3.5 Conclusions

At the beginning of this section we showed the potential consequences of the current FCC regulations. In particular, we showed that TV channels may be lost *in addition to those already lost by allowing the whitespaces* (see Chapter 5). This is an inevitable consequence of interference aggregation and it must be dealt with.

We propose a simple solution of using a power density in order to keep the aggregate interference at acceptable levels. We then show the rate available to secondaries under such a power density rule and compare it with the rate available under existing rules. Amazingly, this rule which is safer for primaries is also better for secondaries!
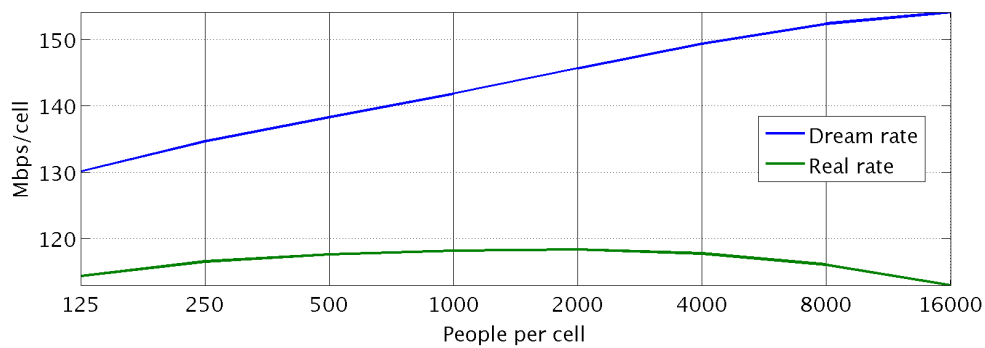
(a) Cellular model
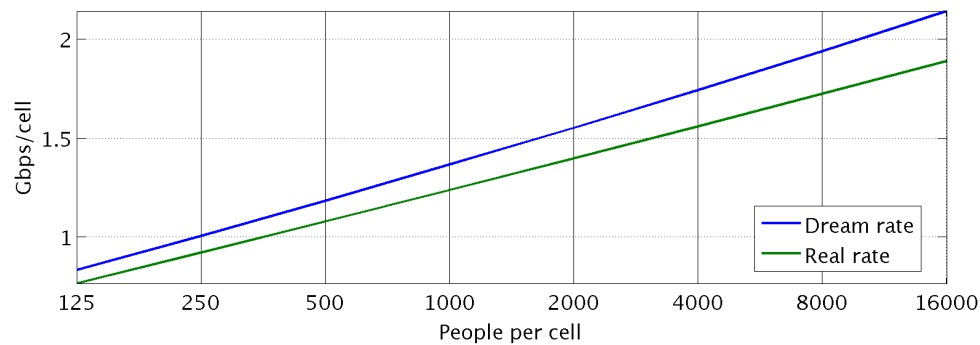
(b) Hotspot model

Figure 3.18: Average rate per *cell* as a function of p, the number of people in the cell (the area of the cell is a function of p and the local population density)

However, rural users are still suffering from ridiculously long ranges due to larger cells[21]. We seek to design a safe power scaling rule which mitigates this problem while maintaining quality of service for urban users and settle on a candidate approximately-optimal power-density rule. Under this candidate rule we see that the achieved rates have roughly equalized across the country.

Finally, we acknowledged the shortcomings of our candidate rule suggested an as-yet untested solution. In particular, we were unable to find a rule which was agnostic of the secondary usage scenario. In future work we would like to find a rule which fairly navigates the tradeoff between the many different deployment scenarios.

---

[21]This is a consequence of our "economic feasibility" model which was developed in Sections 2.6 and 2.7.
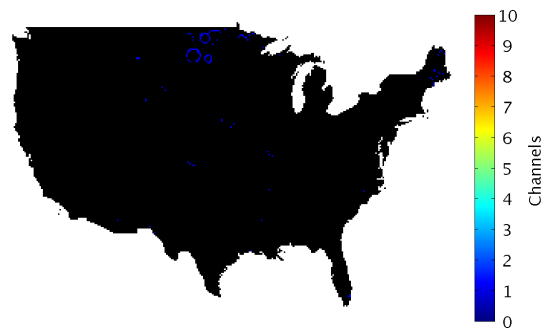
Figure 3.19: The FCC's protected regions are still protected under the power-density-scaling rule.

Chapter 4

Context-aware regulations

Some or all of the work in this chapter appeared in [51]. That paper won the best student paper award for the policy track at IEEE DySpAN 2012.

The TV whitespaces represent an incredible opportunity for innovation. Along with this opportunity come significant challenges: the whitespaces are very heterogeneous in terms of channel availability as well as quality. This poses a challenge to anyone wishing to achieve uniform or even similar quality of service nationwide.

In this chapter, we consider using heterogeneity in the emissions limits to counteract the inherent heterogeneity of the whitespaces, ultimately resulting in a more homogeneous quality of service for secondary users. However, heterogeneity cannot be added to regulations in a haphazard manner. Rather, it must be carefully crafted as a result of looking at the bigger picture and fully exploiting the capabilities of databases.

Databases should not be seen as a necessary evil but rather as an exciting opportunity for improvement and innovation. In particular, rather than being viewed as a simple repository of data, the "database" can be viewed as a cloud-based entity that reports the result of a default kind of Coasian-bargaining that could be expected to occur among frequency-agile radios. We conclude by showing a few small examples of positive heterogeneity as it applies to real-world data for the TV whitespaces in parts of the United States.

This works serves primarily as a proof of concept for dynamic and context-aware regulations. Section 4.5 details some of the natural follow-on work. One of the key observations is that the paradigm of databases as a central coordinator affords us new flexibility. This flexibility can be used to test new algorithms on specific regions, devices, or at specific times; it can also be used to iterate toward a full-scale solution as our understanding increases.

4.1 Introduction

As discussed in Chapter 1, the FCC approved regulations in 2008 which allow wireless transmissions by unlicensed devices within the spectrum reserved for over-the-air TV broadcasts[1] [1]. The regulations define a *protected region* surrounding each TV tower (a.k.a. primary transmitter, in reference to its priority) inside of which TV reception is theoretically guaranteed. Unlicensed devices (a.k.a. secondary transmitters) may transmit once they are safely outside the protected region but they are subjected to a maximum power constraint intended to keep aggregate interference at the primary's receivers (i.e. television sets) at acceptable levels. Since the rules are location-dependent, a secondary transmitter must[2] contact a database to determine which channels are available for use at its location. The databases are operated by neutral third parties and issue coordinating commands to the secondary transmitters.

The 2008 FCC rulemaking was groundbreaking. It brought to the table a new method of spectrum utilization which provokes many interesting technological and regulatory questions. It will revolutionize the use of spectrum as surely as did the original spectrum assignment that first took place in the late 1920s. Not only do the whitespaces open up large portions of viable spectrum but they also hint at novel and interesting ways to think about spectrum regulations in this new era.

The dynamic nature of the whitespaces made databases all but necessary. We already know from previous studies that the use of databases drastically increases the whitespace opportunity [47]. This is a revolutionary way of thinking of wireless spectrum because we can now *depend* on having ubiquitous Internet connectivity through which to contact a database to determine critical parameters for establishing a physical-layer connection. This dependence simply was not conceivable two decades ago.

4.1.1 The power of databases

Discussions of whitespaces typically revolve around the technical challenges of building frequency-agile and cognitive radios and ensuring that they operate safely. However, the TV whitespaces have introduced a far more exciting element into the wireless regulatory ecosystem: databases. The whitespaces represent our chance to fully understand the power of *dynamic* spectrum access: rules which are aware of the bigger picture and evolve over time without the need to re-certify devices. Databases also give us a tool by which we can adapt regulations after deployment if they are found to be unsafe or too conservative [52]. With the rise of cloud computing, software-as-a-service

---

[1]These rules were subsequently updated in 2010 [9] and again in 2012 [23].

[2]The regulations include provisions for sensing-only devices; however, we do not consider such devices in this work.

providers can now rapidly test deploy new features; databases provide the analogous functionality in the whitespaces and spectrum management generally[3].

Furthermore, databases can be used to shift trust away from the devices and to the databases themselves. No longer does the regulator have to trust that all the cognitive radios will correctly interoperate to carry out a distributed inference regarding a newly updated piece of policy language that they have downloaded [53]. This is accomplished by computing the implications of the current policies in the database itself and issuing only basic commands to the devices (e.g. allowed frequencies and emissions limits). This simultaneously makes certification both simpler and more robust, thus increasing assurances that whitespace devices will operate safely[4].

The President's Council of Advisors on Science and Technology (PCAST) recently recognized that the database approach easily and naturally scales. Implicit in their proposal for use of federal spectrum is the idea that databases (and the related administrative software) can be trusted more than the devices themselves and that databases should do the heavy lifting with regards to policy implementation:

> The heart of the proposed SAS is a database that holds information about what spectrum is occupied for a a given location and time; the parameters of the signal, such as power and bandwidth; constraints for specific locations, such as no transmission in blasting zones or along international boarders; and the price for accessing the spectrum. The Radio Access Coordination and Management and Optimization function provides frequency assignments and authorizations. *It may work to optimize overall spectrum efficiency over a given region*, but above all will insure that legacy Federal retain priority access to spectrum. [54, §2.4]

The authors of [52] have discussed some of the potential that databases hold (e.g. incorporating rules which allow secondary transmitters to increase their power as they increase their distance from the primary receivers) but few have truly explored the implications of this technology. Databases mean that regulators are no longer tied to rules made at auction-time: we can now have truly *dynamic* spectrum access. Furthermore, rules can change with location, time, and other variables. This functionality has been partially explored by the FCC in that they require secondary transmitters to reveal their location to the database in order to obtain a list of channels that are available at that location, but there is no reason or need to stop there. We advocated

---

[3]Under the current paradigm, devices are certified to comply with a particular set of rules, e.g. a maximum emission limit. Traditionally these rules have been determined at certification time and were implemented on the device. However, databases give us the opportunity to certify only that a device complies with a set of simple instructions and allow the database to compute the relevant instructions given the current rules. This would mean that policy updates could happen on a much shorter time scale as they would not require any coordination with the devices, only the databases.

[4]This also makes foreign operation trivial: devices need not be aware of the underlying policy engine and only the basic interface must be standardized.

strongly in Chapter 3 that it is imperative to also include some sort of information regarding safe transmit power levels which vary with location: any one-size-fits-all rule would be either far too conservative or far too dangerous.

There are many ways to choose safe power levels; the real trick is to find *good* power levels which foster creativity and innovation in new devices. It is clear that these power-level rules should be application-agnostic and that they should support nationwide-applicable business models. With only these constraints in mind, we can already see a multitude of problems: unlike typical allocations, the whitespaces are incredibly heterogeneous in nature. This is a large obstacle in terms of nationwide coverage because it means that we need to overcome the drastic heterogeneity in the whitespaces in a way that results in a relatively homogeneous quality of service nationwide. Of course, spatial and temporal heterogeneity is a generic feature of the new world of spectrum sharing.

4.1.2 Heterogeneity in the whitespaces

The heterogeneity in the whitespaces is two-fold: protected regions near primaries cause blackout regions for secondary transitters *and* signals from the primary transmitters practically need to be treated as noise by the secondaries. This heterogeneity is not an unknown phenomenon: all authors of whitespace papers are aware of this reality and a few have attempted to quantify it using real-world data (see Chapter 2 and [31, 32, 47, 55]) for the United States and European countries. We see from Figure 4.1 that the distribution of the number of whitespace channels in the United States has a long tail, indicating large amounts of heterogeneity. This poses a challenge for anyone wishing to offer services that can scale to nationwide coverage.
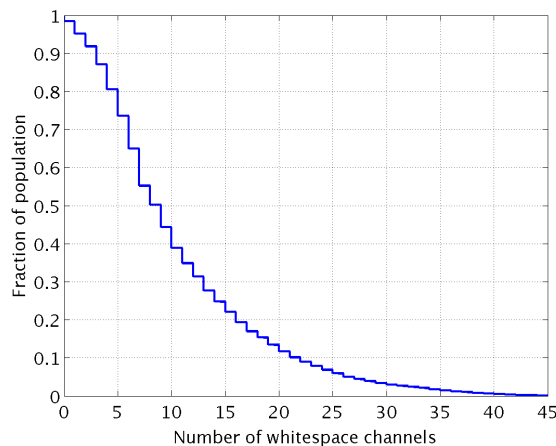
Figure 4.1: Complementary cumulative distribution function (CCDF) by population of the number of channels available to secondary devices in the United States (adjacent channel exclusions included).

Some authors attempt to counteract this heterogeneity via intelligent channel allocation algorithms for use within secondary systems [56–58]. For example, [57] considers link allocations in the presence of channel heterogeneity (both in quantity and in reward). These authors recognize that it is important to avoid self-interference among secondaries and therefore often advocate against frequency-reuse-1 schemes. Self-interference is often modeled through an "interference radius": transmitters that are inside the interference radius cannot operate on the same frequency without effectively jamming one another's transmission.

One can imagine using similar rudimentary interference models to understand the interaction among primaries and secondaries in the whitespaces. For example, one could use the existing frameworks but alter them slightly by declaring a second class of "secondaries" (i.e. primaries) which are automatically allocated any spectrum they request. Indeed, this seems to be the approach of the FCC[5].

However, secondary-to-primary aggregate interference has been shown to be a significant effect (see Chapter 3 and [59]). Intuitively, aggregate interference matters in the primary-and-secondary situation — but not as much in the secondary-and-secondary situation — for the following two reasons:

1. The distances between secondary transmitters are generally small which means that the aggregate interference felt by a transmitter is dominated by his nearest neighbor. On the other hand, primary receivers are somewhat far from the secondary transmitters and therefore the aggregate interference to the primary is not dominated by any single secondary transmitter.

2. Unlicensed devices are designed to have resilience to fluctuating noise levels. The primary network was designed specifically to have to deal with only its own interference and therefore we cannot assume that it is tolerant of high noise levels.

4.1.3 Existing heterogeneity in the rules

The 2008 version of the FCC regulations [1] included provisions for licensed wireless microphones which would allow operators to register the wireless microphones as temporary primaries. This would allow operators of such licensed devices to reserve spectrum for events utilizing wireless microphones. However, Carlson Wireless, Motorola, and WISPA (Wireless Internet Service Providers Association) argued that there should be a portion of the whitespaces reserved for unlicensed wireless microphones at all times and at all locations [9, ¶26]. As a result, the FCC amended its regulations in 2010 with additional provisions for wireless microphones [9]. However,

---

[5]From the FCC's 2008 ruling in which they develop the separation distance requirements: "In developing the table of separation distances, we believe it is desirable to minimize complexity for compliance. In this regard, we have balanced this goal of simplicity with the need to provide assurance that TV services will be adequately protected. ... We find that a transmit antenna at a height of 30 meters transmitting with 4 watts EIRP could cause co-channel interference to a TV receiver with an antenna 10 meters above ground at a distance of 14.4 kilometers and adjacent channel interference at 0.74 kilometers." [1, ¶181]

adding nationwide spectrum for the wireless microphones is not as simple as defining a few channels as off-limits to secondaries regardless of primary presence. Any reasonably-sized set of channels would have regions of unavailability due to the presence of TV transmitters (which clearly have priority) and thus would not suffice. Instead, it was necessary to introduce rules which vary based on location:

> All TVBDs [TV bands devices, a.k.a. secondary devices] are permitted to operate [in] available channels ... subject to the interference protection requirements in §§15.711 and 15.712, **except that operation of TVBDs is prohibited on the first channel above and the first channel below TV channel 37 (608-614 MHz) that are available**, i.e., not occupied by an authorized service. If a channel is not available both above and below channel 37, operation is prohibited on the first two channels nearest to channel 37. These channels will be identified and protected in the TV bands database(s). [9, §15.707a] *(emphasis added)*

Rules	Spatial awareness	Frequency awareness
Current FCC regulations	None: rules do not depend on activity in other locations ✗	Minimal* ✗
Power scaling rules**	Devices obey *aggregate* interference constraints which induces location dependency ✓	None: channels are treated independently ✗
SRASC method***	1) Devices obey aggregate interference constraints 2) Secondary quality-of-service guarantee for *all* locations ✓	Locations may opt to leave fallow available channels to improve the systemwide utility ✓

* The specific channels chosen for wireless microphone exclusions depend on which other channels. are available at that location. Adjacent-channel exclusions also create short-range inter-channel dependencies.
** These power scaling rules were developed in our DySPAN 2011 paper.
*** These rules will be developed later in this paper.

Figure 4.2: Comparison of context-awareness in proposed rules. "Power scaling rules" refer the rules presented in Chapter 3 and [12].

This rule essentially reserves the first two non-primary-inhabited channels above and below channel 37 for wireless microphones. Because of the variability of whitespaces due to the variety of primary locations and channels, the specific channel chosen will vary depending on location. For example, in Berkeley these two channels are 15 and 16 whereas near Chicago they are 24 and 41. From this example, we see that the FCC has recognized the need for heterogeneity in order to provide homogeneous quality of service. It is necessary to use databases to implement these rules but the same databases can also be used to implement other heterogeneous regulations.

4.1.4 Context-aware rules

The FCC's provisions for wireless microphones are an excellent example of what we will call a *frequency-aware rule*: that is, a rule whose outcome for a given channel depends on the characteristics of other channels available at a particular location. The remaining rules for secondary devices in the TV whitespaces are minimally frequency-aware: while cochannel exclusions are independent, adjacent-channel exclusions[6] do induce some dependence. This means that these rules are frequency-aware in only a limited way.

Rules can also be *spatially aware*. We gave examples of such rules[7] in Chapter 3 where we considered methods for scaling the maximum power limit based on the power limits *in other locations*. The power limits were coupled across locations because we enforced an aggregate interference constraint. Indeed any rule that has such a constraint will be spatially aware: there is a finite "budget" (the amount of interference a primary receiver will tolerate) and each transmitter uses a nonzero portion of the budget which affects the decisions made for devices at other locations. In contrast, the FCC regulations are *not* spatially aware because the transmit power of one device does not affect the permissible transmit power of another device; however, these regulations are location-aware. In this chapter we differentiate between location-aware and spatially-aware. The former refers to the cognizance of a device's *own* location whereas the latter refers to the awareness of the potential locations of *all* devices.

4.1.5 Frequency agility

In our work, we assume that devices do not have a preference for a particular channel or set of channels: devices using the spectrum for data communications will view channels as substitutable. Furthermore, the amount of variability in the whitespaces means that any device which has a narrow operating range will not be well-suited to take advantage of whitespaces regardless of the rules. A service without a wide range of acceptable frequencies will suffer from preventable blackouts. In this way frequency agility is essentially a requirement for any marketable device operating in the whitespaces. Any service which cannot in principle be frequency agile (e.g. radio astronomy) should be allocated its own band.

Furthermore, we assume that devices do not have a preference for a contiguous block of channels. Technology has advanced to the point where frequency agility is not uncommon. In most cases, devices will not transmit on all available channels. For example, consider the case in which there are many secondary devices at the same location. Each device will receive a fraction of the total rate available at that location and each device will be able to achieve its rate using only a subset of the available channels.

---

[6]Cochannel exclusions refer to the ban on transmission using channel c when a device is located inside the protected region of any primary operating on channel c. Adjacent-channel exclusions are similar but consider primaries operating in adjacent frequency bands.

[7]Technically these rules were also somewhat frequency-aware because we enforced adjacent-channel exclusions as well.

4.1.6 Overview of chapter

In this chapter, we argue that good rules will need to be both frequency- and spatially-aware. To make our point, we show that either of these qualities alone is insufficient and that in combination they are quite potent. We will begin with two motivating examples which show the various consequences of regulations and contrast these outcomes with that of a hypothetical secondary market. We then formulate and discuss several optimization problems which represent the potential goals of regulations. Then we test our hypotheses using a simple one-dimensional world but real TV tower data [28]. Finally, we conclude by showing that in our model it is important to be both frequency- and spatially-aware and we quantify the gains yielded by these two qualities.

4.2 Motivating examples

In this section we will discuss two examples which motivate our argument for context-aware rules. The model for each example is shown in Figures 4.3 and 4.4, respectively. There are two channels, red and blue, and primary transmitters on each of these channels. The protected region[8] is represented by colored circles around the primary transmitters and the no-talk region[9] (where applicable) is shown with black semi-circles. Secondary transmitters are shown in discrete locations as green dots whose size indicates their relative power. We assume that a secondary transmitter will use all available channels (i.e. channels with nonzero secondary power). For simplicity, we do not consider the effects of self-interference among secondaries. Here we do not consider adjacent-channel or microphone-related exclusions.

Each of the subfigures shows the power allocation chosen under different regulations. The first follows the current FCC regulations. The second follows a power scaling rule such as that in Chapter 3. The third is a rudimentary context-aware rule — termed the "SRASC method" — which will be further defined in Section 4.3.3.3. The properties of the three rules are given in Figure 4.2 and an analysis is presented below.

4.2.1 Existing FCC-style rules

Consider the FCC-style rules shown in Figures 4.3(a) and 4.4(a). We note that the separation distance is fixed (specifically, it is 14.4 km) and that secondaries outside of the no-talk region operate with a fixed transmit power. As we argued in Chapter 3, there are two main problems with these rules:

1. The fundamental idea of a per-device power limit is based on an expected device density. If there are too many secondaries operating simultaneously, the aggregate interference may

---

[8]The protected region is the space where secondaries are not allowed to transmit under any rules. Primary receivers inside this protected region are theoretically guaranteed reception.

[9]This is the region is the space where secondaries are not allowed to transmit. The no-talk region necessarily includes the protected region.

(a) FCC-style rules

(b) Power scaling rules

(c) Context-aware rules
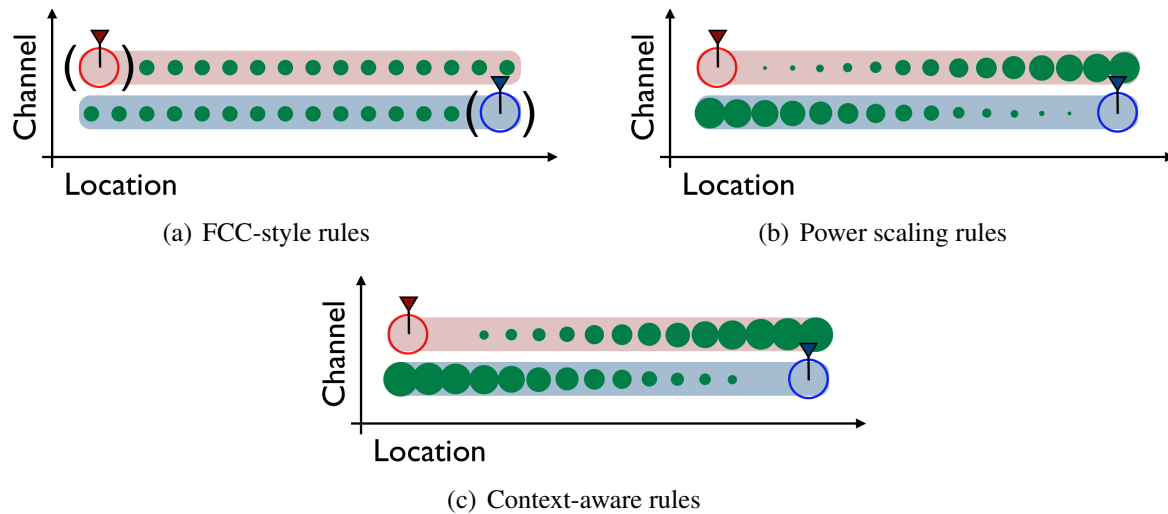
Figure 4.3: Model for example 1. There are two channels (red and blue) with one primary tower each. The protected regions are shown as circles around the primary transmitters and the no-talk regions are shown using black semi-circles. Secondaries are shown in discrete locations and their sizes indicate their relative transmit powers. We assume that a secondary transmitter will use all available channels (i.e. channels with nonzero secondary power). In this example we see a large difference between the FCC-style rules and the power scaling rules but minimal difference between the power scaling rules and the context-aware rules. Note that the FCC-style rules do not guarantee that the primary's aggregate interference constraint is obeyed.

cause some primary receivers in the protected region to experience an outage. This is a big threat to the future of whitespaces: services which feel threatened or made vulnerable by the regulations may advocate strongly for more conservative regulations in order to allay their fears.

2. The choice of separation distance is a political one: since urban areas tend to be near protected regions, a large separation distance unfairly discounts urban areas. On the other hand, increasing the separation distance means that a higher maximum transmit power can be chosen; this would greatly benefit areas which are able to transmit. This tradeoff makes it difficult to give the same quality of service to everyone.

4.2.2 Power scaling rules

Based on the arguments above, we developed candidate power scaling rules that attempted to "blur" the no-talk region and create a graceful degradation of power — and hence data rate — as one neared the protected region (see Chapter 3), as shown in Figures 4.3(b) and 4.4(b). These rules had the advantage of being safe for primaries while being flexible for secondaries. Notice that there is no explicit no-talk region in these rules.

(a) FCC-style rules

(b) Power scaling rules

(c) Context-aware rules
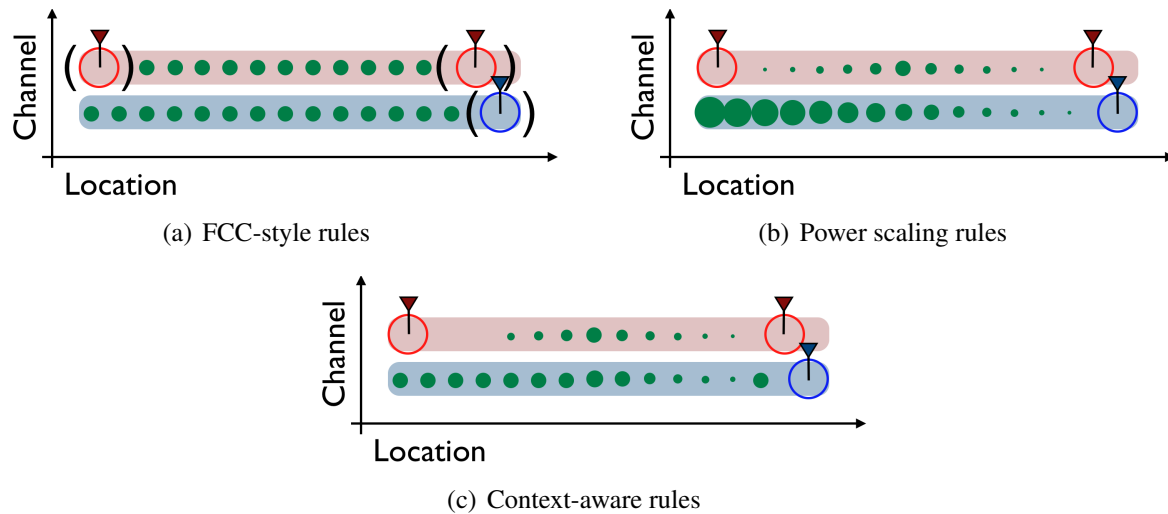
Figure 4.4: Model for example 2. There are two channels (red and blue) with one primary tower on the blue channel and two primary towers on the red channel. The protected regions are shown as circles around the primary transmitters and the no-talk regions are shown using black semi-circles. Secondaries are shown in discrete locations and their sizes indicate their relative transmit powers. We assume that a secondary transmitter will use all available channels (i.e. channels with nonzero secondary power). In the context-aware rules we allocate power to a point near the protected region on the blue channel because it is unable to use the red channel. Note that the FCC-style rules do not guarantee that the primary's aggregate interference constraint is obeyed so their powers are not comparable to the power scaling and context-aware rules.

Because the rules enforced the aggregate interference constraint for the primary, they were *spatially aware*: powers were coordinated across locations to ensure that the primary's interference constraint was obeyed.

The important thing to notice here is that secondaries which are very near to the protected region would cause much more interference to the primary. For example, see Figure 4.5 which shows the relative impact of a secondary's transmission on the aggregate interference based on his distance to the protected region. We see that the impact at a distance of 1 km is almost four orders of magnitude higher than the impact at 10 km: this means that his power would have to be 10,000 times lower to cause the same interference to the protected primary receiver. Because of this fact, we can think of secondary power as being more "expensive" near the protected region.

The power scaling rules presented in Chapter 3 are far from the only example of variable power limits. Indeed, the Electronic Communications Committee (ECC) in Europe noted that "location specific output power seems to be better from a spectrum usage view" [46, §9.1]. The UK's Ofcom has also included provisions for scaling the transmit power [60].

By examining Figure 4.3(b), we notice that locations which are near the protected region on the red channel are far from the protected region on the blue channel and vice versa. This suggests that

taking into account the location-specific alternatives may yield considerable gains over assuming it to be an all-or-nothing game on each channel.
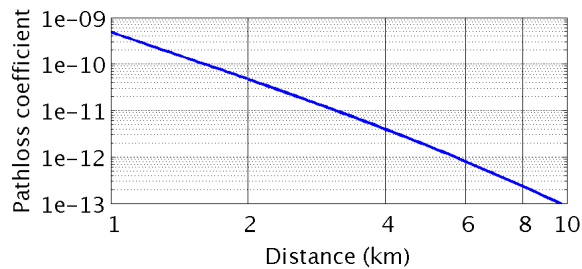
Figure 4.5: Pathloss coefficients as a function of distance on TV channel 21. The pathloss coefficient is also the interference weight of a secondary transmitter as a function of distance to the protected region. Notice that the weights decrease by almost four orders of magnitude over a distance of 9 km, making power much "cheaper" at 10 km than it is at 1 km.

4.2.3 How would a secondary market behave?

We are now at a point where we understand the competing interests of the secondary transmitters at each location. We assume for simplicity that each location is equally interested in increasing its achievable rate.

In a simple aggregate-interference model we can think of primary receivers inside the protected region[10] as having an "interference budget" which reflects the amount of interference from secondary transmitters that they are able to tolerate. Due to the difference in pathloss shown in Figure 4.5, we know that secondary power used near the protected region is "expensive" (i.e. it uses more of the budget per unit of secondary power) while secondary power used at far-away locations is "cheap."

Using this knowledge, we can describe the relative prices in the example of Figure 4.3:

- Power used at locations toward the left is expensive in the red channel but cheap in the blue channel.

- Power used at locations toward the right is expensive in the blue channel but cheap in the red channel.

- Power costs are roughly the same near the center.

---

[10]Since signal attenuation is monotonically decreasing with distance, we can take the worst-case viewpoint of a primary receiver which is at the edge of his respective protected region.

Taking the power scaling rules as a starting point, the Coasian-bargaining behavior of the participants is easy to predict: they will "sell" some or all of their expensive power on one channel and use the "profits" to "buy" more cheap power on the other channel. For example, a secondary located near the left edge of the model will relinquish his right to transmit in the red channel in exchange for increasing his power in the blue channel. The secondary does this because the increase in rate in the blue channel exceeds his loss in the red channel.

In Figure 4.4 we see a similar example which is asymmetric. The point on the right-hand side of the model now has the blue channel as his only option and thus "buys" power from other locations on the blue channel in order to achieve a comparable rate. In both cases, we see that frequency-agile or wide-band secondaries are making decisions based on the characteristics of the channels available to them; in other words, they are demonstrating frequency awareness.

Unfortunately, an actual money- and transaction-based secondary market such as this may turn out to be very complex due to the sheer number of participants. The advantage of the whitespace-style regulations has always been that they provide a reasonable default way to access spectrum without engaging in complex transactions, for example in the provisions for wireless microphones. Given that frequency awareness can deliver value, it is worth seeing if that can also be done in a good default way. The power of databases is that we can simulate trading without needing actual trades to occur.

4.2.4   Context-aware rules

The power scaling rules shown earlier in Figures 4.3(b) and 4.4(b) are safe but they are ultimately too conservative. In making the rules, we assumed that shutting someone out on one channel was the worst thing you could do because it would give them no rate at all. However, many locations have alternative channels on which they are not so close to the relevant protected region. The Coasian-bargaining solution has these locations selling their right to transmit in the "expensive" channels and increasing their rate in the alternative channels. For those locations which don't have alternatives — such as the secondaries on the right-hand side of Figure 4.4(c) — we need *frequency-aware* regulations which allow them to transmit near the protected region if that is their only option.

In this particular example, we see in Figures 4.3(b) and 4.4(b) that most locations have both the red and blue channels available. We can therefore judiciously restrict some locations to only one channel if it benefits the greater good[11]. The choice of who to "kick" from which channels then depends on the definition of "greater good." We consider three options below.

---

[11]Note that is always better to use power P in each of two identical channels than power $2P$ in one channel. This is because rate (Shannon capacity) is linear in bandwidth but logarithmic in power. However, the choice given here is more complex so sometimes it is better to use fewer channels in one location so that other locations will be able to use more channels and/or power.

4.2.4.1 Maximize total power used

Given the primary's interference constraint, we could consider trying to maximize the total amount of power available for secondary use. The solution to this problem turns out to be equivalent to maximizing the amount of power that can be used at the point furthest from the protected region since that is where power is "cheapest." Intuitively we know that this is not a good allocation of resources because it favors only one location, i.e. it is not fair. There is no fairness incentive because of the linearity of the objective function. In reality, we know that the utility of power is not linear but rather logarithmic due to the nature of the information-theoretic capacity bound. In order to incentivize fairness we consider maximizing the average utility in the next section.

4.2.4.2 Maximize average rate

It is tempting to think that maximizing the average rate to all secondaries will maximize the individual utility of each secondary. If all secondaries are subjected to the same amount of noise (with or without self-interference), have the same amount of spectrum available, and there is only a maximum system-wide power constraint — i.e. in homogeneous spectrum — it *is* optimal in terms of fairness to give each secondary an equal amount of power because of the concave nature of the rate function[12].

However, the properties of the whitespace environment mean that we are not operating in such a simple world. First of all, the utility of one unit of power depends heavily on the location since the noise level is now location-dependent. Locations closer to the primary transmitter will experience higher noise levels. Secondly, we no longer have a sum power constraint but instead a *weighted* sum power constraint. Our only constraint is that the primary receiver (without loss of generality we consider one at the nearest edge of the protected region) observes aggregate noise from the secondaries which is no greater than some fixed value N_s. As we saw in Figure 4.5, the relative impact of a secondary's transmit power varies greatly depending on his distance to the primary receiver. Thus we can think of power as "cheaper" for secondaries which are far from the protected region and more "expensive" for those near the protected region.

These properties combined mean that it is both "cheaper" and more beneficial in terms of the average rate to allocate power to secondaries which are far from the protected region. Note that if there were multiple secondaries at the same location, we would still split the power equally among them (as in the previous example). However, this type of objective function does not favor fairness among locations.

---

[12]For example, suppose the maximum power limit is $10P$ and there are 10 secondaries in the system, each with a noise floor N. We could obtain rate $\log(1 + P/N)$ on each of 10 secondaries or rate $\log(1 + 10P/N)$ for one secondary and 0 for all the rest. We know that $10 \cdot \log(1 + P/N) > \log(1 + 10 \cdot P/N)$ and therefore the first scheme is optimal.

4.2.4.3 Maximize minimum rate

We saw in the previous example that it may be important to explicitly incorporate fairness into our objective function. As a result, we will look at the most fair objective function: the max-min rate. This objective function seeks to provide all locations with a maximal quality-of-service guarantee. The optimal solution will have the following properties:

- Locations which are inside of the protected region on all channels will not be able to transmit on any whitespace channel. This is unfortunate but unavoidable because the protected region is a hard constraint. We will not consider these locations to be within the feasible set of the minimization function because any rules which would allow them to use the whitespaces would necessarily violate the primary's interference constraints.

- In general, each location should try to use the channels with the minimum interference weight, thus reducing its impact on other locations. This lower interference weight also allows it to use a higher power if necessary.

- Locations which are near the protected region on all channels are still allowed to transmit. These locations are "expensive" but it is unfair to deny them service.

We see an example of this allocation in Figure 4.3(c). The locations near the protected region on the red channel have graciously moved to the blue channel where they can increase their power to make up for the rate they lost by vacating the red channel. The same thing happens on the other side with locations near the protected region on the red channel. In Figure 4.4(c), we see the main difference between this method and the other rules: the point on the far right is allowed to transmit despite being near the protected region.

This system has two advantages:

- The no-talk region is not fixed which allows leniency for devices with few or no alternatives.

- By removing as many people from the neighborhood of the protected region as possible, we have increased the total amount of power we can use in our system and therefore the rates.

Like the power density rules, this approach is spatially aware because it enforces an aggregate interference constraint and thus induces a dependency between locations. However, this solution is also inherently *frequency aware*. That is, in order to make these decisions, we really need to know what other options exist for a given location. We will see examples later which demonstrate why we cannot achieve this performance without frequency awareness.

The answer is relatively obvious for the toy examples discussed above but how does it work in the real world? In the next section, we will use an example one-dimensional world within the United States and check our hypothesis there.

4.3   One-dimensional test in the United States

In this section, we apply our hypothesis to the US using actual tower data [28] and the ITU propagation model [29].

4.3.1   Assumptions

For simplicity, we restrict ourselves to a one-dimensional line[13] connecting Berkeley and New York City, pictured in Figure 4.6. Future work may extend our results to two dimensions. The main difficulty lies in the computational complexity and we do not expect the qualitative results to change. We also make several other simplifying assumptions:

- We refer to the theoretical achievable capacity (a.k.a. the Shannon capacity) as the "rate." If the power level is P, the noise level is N, the bandwidth is B, and the pathloss coefficient is γ we can write the resulting rate as

$$R = B \cdot \log_2 \left(1 + \frac{\gamma \cdot P}{N} \right)$$

- The distance between the secondary transmitter and the secondary receiver is constant (which implies that γ is constant). This implies that noise, power, and bandwidth are the only factors in the rate calculation.

- In this particular example, there are no primary towers on channel 5 whose protected regions intersect the line. This implies that there are no interference constraints on channel 5. As a result, there is no power limit and therefore no rate limit. We exclude secondaries from using channel 5 in our calculations in order to make our results meaningful and interesting.

4.3.2   Heterogeneity

We have previously seen statistics showing the extreme heterogeneity found in the whitespaces. Here, we wish to show the reader the exact amount of heterogeneity present along our one-dimensional line. The line and nearby TV towers are shown in Figure 4.6; this illustration informs Figures 4.7 and 4.8.

In Figure 4.7, we see the number of channels available under the FCC rules for secondary use along the same line. The difference between the two lines in this graph underscores the effect of adjacent-channel exclusions[14] In Berkeley, there are only a few channels available due to the preponderance of TV stations on TV towers in the area (e.g. Sutro Tower and the San Bruno Tower). As we pass

---

[13]We only worry about protecting the TV towers whose protected regions include some portion of the line. This excludes towers with protected regions adjacent to but not overlapping with the line so this is only a toy.

[14]For all results in this chapter we enforce adjacent-channel exclusions.
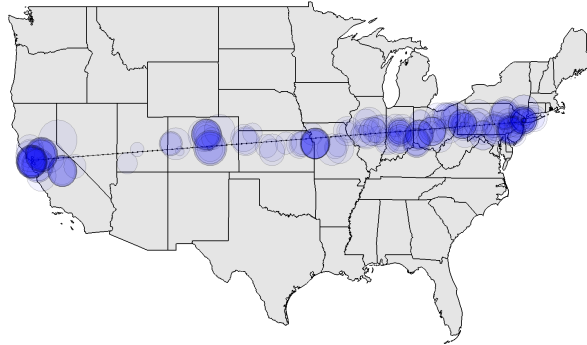
Figure 4.6: Trip across the United States from Berkeley to New York City. Blue circles indicate the protected regions of nearby TV towers.

through other less populous western states, we see the number of available channels increases. In the eastern United States, the population density and consequently the TV tower density increases, leading to a decrease in the number of available channels. Finally, upon reaching the east coast we once again see a drastic decrease in the number of available channels since New York City has many local TV stations. We see the locations of these towers also reflected in the average noise floor, shown in Figure 4.8.
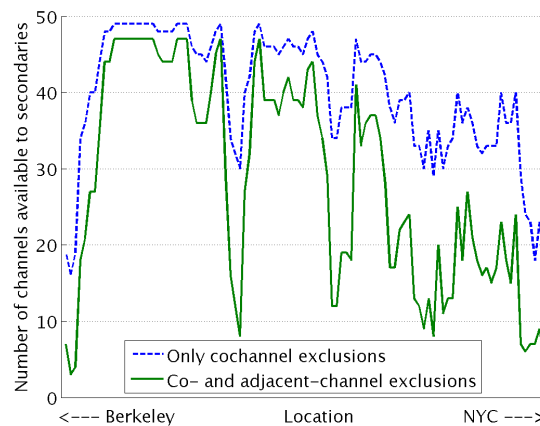
Figure 4.7: Number of channels potentially available to secondaries in the TV whitespaces along the line shown in Figure 4.6 .
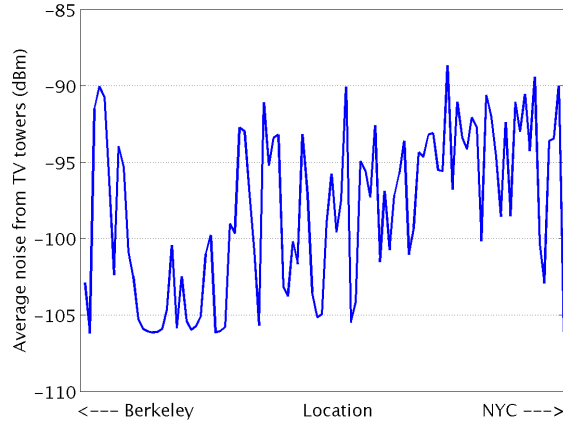
Figure 4.8: Average noise level for secondaries operating in the TV whitespaces along the line shown in Figure 4.6.

4.3.3 Methods for whitespace power allocation

In this section, we look at several algorithms for allocating power to each secondary (location, channel) pair. We present the details of each algorithm and compare the results. A comparison of the algorithms can also be found in Figure 4.9. Note that at optimality, all locations achieve the same rate R.

Nomenclature and variables: we will use the following nomenclature and variable names:

- l is an index indicating the discrete location of a secondary. We will refer to "location" and "secondary transmitter" interchangeably.

- c is an index indicating the channel (i.e. frequency).

- N_c is the total number of channels in the model.

- $N_c(l)$ denotes the number of channels available at location l. $1 \leq N_c(l) \leq N_c$ for all l.

- $R(l, c) \geq 0$ is the rate achieved by the secondary at location l on channel c.

- R denotes the total rate achieved at every location. That is, $\sum_{c=1}^{N_c} R(l, c) = R(l) = R$ for all l.

Our goal is to maximize (over power allocations) the minimum rate (over all locations) subject to the aggregate interference constraints of the primary receivers in the protected regions. We can

write this as

$$\max_{powers} \; \min_{l} \; \sum_{c=1}^{N_c} R(l,c) \qquad (4.1)$$

subject to: aggregate interference constraints

We will now examine several algorithms which attack this optimization problem with different sets of knowledge.

Method	Spatial awareness	Frequency awareness
SRASC Split rate among a subset of channels	1) Devices obey aggregate interference constraints 2) QoS guarantee for *all* locations ✓	Locations may opt to leave fallow available channels to improve the systemwide utility ✓
SREAC Split rate equally among channels	1) Devices obey aggregate interference constraints 2) QoS guarantee for *all* locations ✓	Minimal: only considers the number of available channels **?**
MECR Maximize each channel's rate	1) Devices obey aggregate interference constraints 2) QoS guarantee for *all* locations ✓	None: each channel is maximized independently ✗
FPE Fixed power everywhere	Does not obey aggregate interference constraints ✗	None: power does not change with the number of available channels ✗
FPMCQ Fixed power, minimum channel quality	Does not obey aggregate interference constraints ✗	Transmits only on channels which are of sufficient quality ✓

Figure 4.9: Comparison of the properties of the methods presented in Section 4.3.3.

4.3.3.1 Maximize each channel's rate ("MECR")

One natural approach to this problem is to maximize the achievable rate on each channel independently. This solution ignores the dependence between channels and is therefore frequency-unaware. However, we will enforce the aggregate interference constraint so that it is spatially aware.

In this method, the power is allocated such that the rate $R(l,c)$ is R_c if location l can transmit on channel c and 0 otherwise. Note that each location will have a different set of available channels which causes spatial variation in the total rates achieved. The max-min rate problem can be written as

$$\min_{l} \; \sum_{c=1}^{N_c} R(l,c)$$

s.t. R_c is the maximum achievable rate on channel c

Note that the maximization already occurred in the choice of R_c. The method presented in the next section adds a minimal amount of frequency awareness.

4.3.3.2 Split rate equally among channels ("SREAC")

This method is motivated by the concept of a service at location l which desires a fixed rate R and blindly attempts to achieve this rate by splitting it equally among its $N_c(l)$ channels, thus achieving rate $R/N_c(l)$ on each. The max-min rate R is chosen such that the aggregate interference constraint is obeyed.

This method is spatially aware since it obeys the aggregate interference constraint but it is only minimally frequency-aware since it does not consider the relative quality of channels. The SREAC method can be viewed as the starting condition for the SRASC method discussed in the next section.

4.3.3.3 Split rate among a subset of channels ("SRASC")

This method is a heuristic approach rather than an exact solution to the max-min problem[15] of (4.1). It uses the SREAC method as a starting point but greedily "bans" locations from specific channels (i.e. imposes additional constraints). This has the effect of creating a flexible separation margin: locations which have a "cheaper" channel available will vacate "expensive" channels (thus making space for others) while those who have only the "expensive" channels are still allowed to use them[16]. Note that this is the method shown in Figures 4.3(c) and 4.4(c). This method is both frequency- and spatially-aware.

4.3.3.4 Fixed power everywhere ("FPE")

These rules are inspired by the FCC's regulations which allow transmitters located outside of the no-talk region to operate at 4 Watts. Likewise, we allow each location to transmit at 4 Watts on each channel. In this model, we found that a separation margin of less than 26 km did not protect receivers inside the primary's protected region. Unlike the previous methods we will not enforce the aggregate interference constraint. For these reasons the FPE method is both frequency- and spatially-unaware.

---

[15]Although the problem stated in Equation 4.1 is one of concave maximization over a convex set, conventional solvers such as CVX and the Matlab Optimization Toolbox do not seem to have the necessary precision. The constraints define the interior of a polytope and $R(l, c)$ is clearly concave due to the logarithm. Summation and pointwise minimization preserve concavity, thus the problem is concave.

[16]There are situations in which this is not an entirely selfless behavior. For example, consider the example given in Figure 4.3(c): locations near the protected region on one channel used the other channel exclusively. This left more of the "budget" for far-away locations, meaning that they received a boost to their powers. Since the example is symmetric, vacating one channel meant receiving a much higher power in the remaining channel.

4.3.3.5 Fixed power, minimum channel quality ("FPMCQ")

These rules were inspired by the idea of a device using the FCC rules but which has a minimum channel quality constraint. This reflects the cost-benefit analysis that will be done by transmitters when considering whether or not to transmit on an additional channel.

Since these rules do not obey the aggregate interference constraint, they are considered spatially-unaware. However, since they take into account the relative quality of the available channels they are frequency aware.

Regardless of the channel quality threshold these rules will never perform better than the FPE rules since the powers used will be less than or equal to those from the FPE rules. We have included these rules for completeness but we will make no further mention of them since they are obviously strictly inferior to the FPE rules for all metrics considered here[17].

4.3.4 Evaluation

In this section, we discuss the relative performance of the methods described in the previous section. We begin by looking at the rate achievable at all locations[18]:

Method	Max-min rate R	Safe for primaries?
SRASC	588 Mbps	Yes
SREAC	0.13 Mbps	Yes
MECR	97 Mbps	Yes
FPE	254 Mbps	*No*
FPE, 26 km	197 Mbps	Yes

The rates in the table above are much higher than one might expect to see in the TV whitespaces. However, we wish to remind the reader that these results hold for a *one-dimensional* world and thus reflect a drastic decrease in the number of constraints. This allows some locations to use a much larger power than would be safe in the two-dimensional case[19]. We expect that investigating a two-dimensional example will confirm the same basic picture revealed here.

We see that the SRASC method is the clear winner, achieving over twice as much as the FPE method. This is especially impressive because the power allocation used in the FPE method is *not* "safe" for primary receivers (i.e. it does not obey the primary's aggregate interference constraint) while the power allocation from the SRASC method *is* safe. This improvement is due to the judicious use of available channels employed in the SRASC method. The MECR and SREAC methods perform poorly for the same reason. In an attempt to give these other methods a fighting chance, we consider two variations of the problem.

---

[17]The only potential advantage these rules have is that they will require a smaller separation margin to be safe; however, since they are not being tested for safety this is irrelevant.

[18]In our example one-dimensional world, all points are outside of the protected region and the adjacent-channel protected region on at least one channel. This can be verified using Figure 4.7.

[19]However, we saw little change in the number after artificially capping the secondary transmit power at 100 W.

4.3.4.1 In terms of separation distance

Here we consider increasing the separation distance to determine if it improves performance. The separation distance is the enforced distance between the protected region and the nearest transmitting secondary[20]. The existence of a separation distance blindly bans secondary transmitters from channels without considering the alternatives available to them. As the separation distance increases, secondaries will need to increase the power used on remaining channels in order to compensate for a decreasing number of available channels. We can see the resulting max-min rate in Figure 4.10 as a function of the separation distance. We discuss the general trend of each method in turn.

SRASC: This method already considered "banning" secondaries from technically-available channels. Although it generally increased the distance between secondaries and the protected region, this was not a hard constraint. Introducing this as a hard constraint decreases the max-min rate. To see why, consider a secondary with two available channels which is near (but not inside) the protected region on each channel. As the separation distance increases, he loses one of these two channels. In order to maintain the same rate R, he would need to increase his power on his remaining channel. However, he cannot always increase it enough to maintain rate R due to the interference constraint. This accounts for the downward jumps as the separation distance increases. At a separation distance of about 140 km, at least one location no longer has any channels available and so the minimum rate drops to zero for all methods.

SREAC: Unlike the other methods, we see that the trend for this method is not monotone decreasing nor is it monotone increasing. We explain these two behaviors as follows:

- *Increases:* As the separation distance increases, secondaries are banned from using the channels which are near the protected region (therefore "expensive;" see Figure 4.5). By using cheaper channels, the net power usage increases and consequently the rates increase.

- *Decreases:* Occasionally the separation distance increases so much that a location which really depended on the use of a particular channel is now "banned." This location is now unable to achieve his former rate without affecting the power allocation for other locations.

The interplay between these two phenomena causes the function to follow an interesting path. Ultimately, though, this method performs worse than the SRASC method because it makes a hard decision about channel availability without considering the bigger picture.

MECR: The max-min rate for this method is strictly increasing with the separation distance. The rates are very low for small separation distances due to allocation of power to locations which are "expensive." As we increase the separation margin, we start using cheaper locations and the

---

[20]For simplicity, we assume that the separation distance is the same for cochannel and adjacent-channel exclusions. The FCC rules use different separation distances.

expensive (location, channel) pairs now receive their rate using cheaper channels. This method still performs worse than the SRASC method at every point because it is forced to serve all locations outside of the no-talk region on each channel: anything less runs the risk of under-serving the minimizing location. In contrast the SRASC method can intelligently leave fallow certain (location, channel) pairs if this yields an overall gain.

FPE: This method simply did not have the power to be able to compete with the SRASC method. Even though it is unsafe for separation distances smaller than 26 km, it still uses a lower power than the SRASC method when far from the protected regions (i.e. it doesn't employ power scaling).

Conclusion: Ultimately the goal is to maximize the minimum rate. We have seen by examining Figure 4.10 that this maximum is achieved by using the SRASC method with no minimum enforced separation distance (i.e. maximum flexibility). Recall that the SRASC method is a heuristic approach to solving the max-min problem and furthermore the algorithm itself is not fully optimized. However, the solutions for all other methods are exact so the advantage of the context awareness is actually understated.
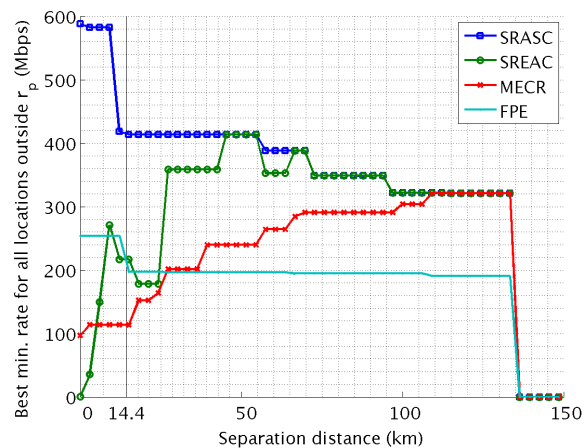
Figure 4.10: Comparison of max-min rates achieved with four power-allocation methods while varying the separation distance. We see that the SRASC method is strictly superior to the other three methods.

4.3.4.2 In terms of coverage

We saw in Figure 4.7 that some points — especially those near the coasts — suffer from a severe lack of channels. Nonetheless, we have been requiring our algorithms to accommodate these locations. To check that there aren't a few "trouble locations" which are giving an unfair advantage to the SRASC method, we consider making the conscious decision to deny service to some locations.

Practically, we did this by removing locations from the feasible set in the minimization problem. The algorithm is as follows for each method (individually):

1. For each location as l, consider removing l from the feasible set in the minimization. Evaluate the max-min optimization problem to get the potential rate as a function of the removed point, l.

2. Find the location l' which, when removed, gives the greatest potential rate. Note that l' need not be the same point for each method.

3. Remove the point l' from the feasible set and return to the first step.

By construction, the max-min rate will monotonically increase as we deny coverage to more and more locations. Indeed we see this behavior in Figure 4.11. For ease of interpretation, we do not consider varying the separation distance in this exercise and instead assume that a secondary is eligible to transmit if it is outside of the protected region[21].

From these results we see that not only is the SRASC method outperforming all other methods but in fact it does about three times better than the SREAC method and 2.5 times better than the MECR method *even when we remove the most bothersome points*. This demonstrates the important gains made possible by full context awareness.
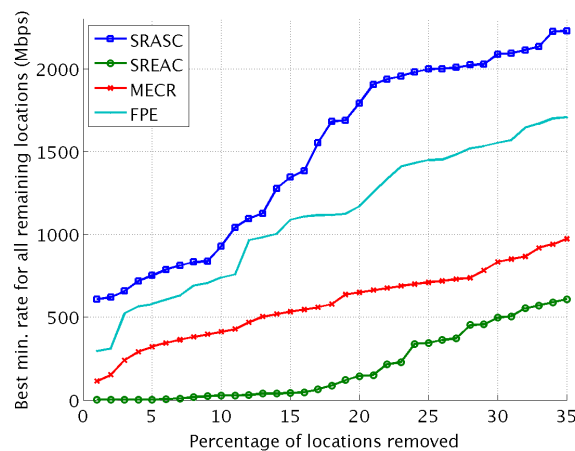
Figure 4.11: Comparison of the max-min rates achieved with four power-allocation methods while varying the desired amount of coverage. Notice that the SRASC method is strictly superior to the other three methods.

---

[21]As always, adjacent-channel exclusions are enforced.

4.3.4.3 Discussion

We have seen that under all examined conditions, the SRASC method outperforms the other four methods. From Figure 4.9 we see that it is the only method which has true spatial and frequency awareness. This suggests that these qualities are necessary in regulations which seek to provide a fair quality of service to all constituents.

4.4 Conclusions

The creation of a spectrum regulatory agency in 1927 and the subsequent rearrangement of the spectrum brought about a brand new way of utilizing the spectrum. It reduced interference between competing services and improved performance. This method worked well for almost a century until spectrum became artificially scarce, at which point the FCC recognized the need to utilize the whitespaces. The dynamic nature of the whitespaces coupled with the need to keep the primaries safe really makes databases essential to the use of the whitespaces. However, databases should not be seen as a necessary burden but rather as an exciting opportunity for exploration and improvement. In particular, we explored the potential gains of including a variable secondary power limit which is disseminated via databases.

We have argued that the extreme heterogeneity in the whitespaces means that context-aware rules can provide huge gains over simpler rules. We first showed these improvements using rudimentary examples including a thought-experiment about the behavior of a secondary power market. The lesson was that there was a powerful incentive to trade which raised the question: why not let the database do some of the obvious trades itself? In fact, the only situation in which the database cannot compute the optimal Coasian-bargaining solution is when the secondary devices have some sort of specialized or local knowledge not available to the database ahead of time. When considering the advantages of secondary markets we should therefore compare them to the optimal default database-reliant rules rather than the current one-size-fits-all rules.

We then considered a more complex example which included real-world data on TV transmitters in the United States. The gain here from using a context-aware method was substantial and remained so even after modifying the example to remove the apparent disadvantage of the other methods.

All of our results suggest that context-aware regulations perform better for all metrics, so what sort of situation *would not* benefit from the use of context-aware rules? If we look at context-aware rules as the steady-state result of Coasian-bargaining, the answer is clear: context-aware rules provide no gain when resources *cannot* be traded. For example, radioastronomy applications rely on the use of a specific frequency and cannot use any other frequency no matter how high the incentive. However, whitespaces are not meant to be used for such selective applications but rather as a breeding ground for innovative devices and applications. Well-chosen context-aware rules will foster this innovation, ensuring that the full potential of the whitespaces is harnessed.

4.5 Future work

We have identified several directions for future work:

- We expect our results to generalize from a one-dimensional line to a two-dimensional map. The main difficulty is in the computational complexity that this problem presents with numerical precision issues.

- We suspect that our results hold for general risk-averse (i.e. coverage-inclined) objective functions but we have not yet worked out these results.

- Heterogeneity in the whitespaces is not confined to the variability of the noise floor and the number of available channels. We fully expect that a wide variety of devices will utilize the whitespaces and it is extremely important to ensure that regulations are application-agnostic [58]. Anything less than this undermines the notion of the whitespaces as a place for discovery and growth. The authors of [56] began an exploration of this space when they considered assigning a range to each channel.

- Given physical limitations of current cognitive radio technology, we may need to add constraints that address the maximum frequency agility of a device. For example, devices may not be able to transmit simultaneously at 60 MHz and 600 MHz.

- It may also be useful to impose a "minimum utility" constraint (e.g. minimum spectral efficiency) on a channel: if a channel offers too little, it will go unused and the resources would have been better allocated elsewhere.

Finally, as important as it is to choose good regulations, we do not necessarily need the full solution at the time of deployment: databases afford us new regulatory flexibility that can be used to our advantage.

Part III

Spectrum utilization

Chapter 5

Repurposing vs. whitespaces, part I

Some or all of the work in this chapter appeared in the author's Master's thesis, [38].

In this chapter we discuss two different families of rules, refarming TV channels and reducing protections for TV towers. We will compare these two families to the FCC's choice along the tradeoff curve between whitespaces and incumbent protection.

5.1 Repurposing the channels: the traditional approach

Here, we consider the fundamental decision to even *have* TV whitespaces. Rather than create the TV whitespaces, the FCC could have simply reassigned TV stations to a smaller set of channels and opened up the remaining bands to unlicensed use. Alternatively, they could have refarmed some of the channels in this manner while opening up the remainder as TV whitespaces. The tradeoff for these two scenarios is shown in Figure 5.1 as the "no sharing" and "sharing" curves, respectively. The black crosshair marks the tradeoff point that the FCC chose.

In these calculations, we progressively remove channels[1] from TV use and give them over to exclusive "secondary" use. Each dot marks the "removal" of one channel, so the second dot indicates that one channel was refarmed, the third dot indicates two refarmed channels, etc. In the no-sharing model, channels still used for television are completely off-limits for secondaries. In the sharing model, secondaries may additionally transmit on TV-inhabited channels using the current FCC rules for TV whitespaces[2]. Notice that the sharing model will do strictly better than

---

[1]Channels are heuristically ranked in order of greatest gain in a knapsack-like problem in which the value is the potential secondary-cum-primary rate and the weight is the number of TV viewers that will be lost.

[2]For simplicity we ignore adjacent-channel exclusions for the sharing model (the green line in Figure 5.1). Although this may seem a minor issue, we will see in the next section that adjacent-channel exclusions actually account for a large portion of lost secondary opportunity. We can see this already in the horizontal distance between the the FCC's tradeoff point (shown as a black crosshair) and the first green point. This green point represents the case where all channels are used as whitespace but with only cochannel exclusions; the FCC's tradeoff point is identical except that it also includes adjacent-channel exclusions.

(a) 1km transmission distance, MAC model

(b) 10km transmission distance, MAC model

(c) $p = 2000$, MAC model

(d) $p = 2000$, hexagon model
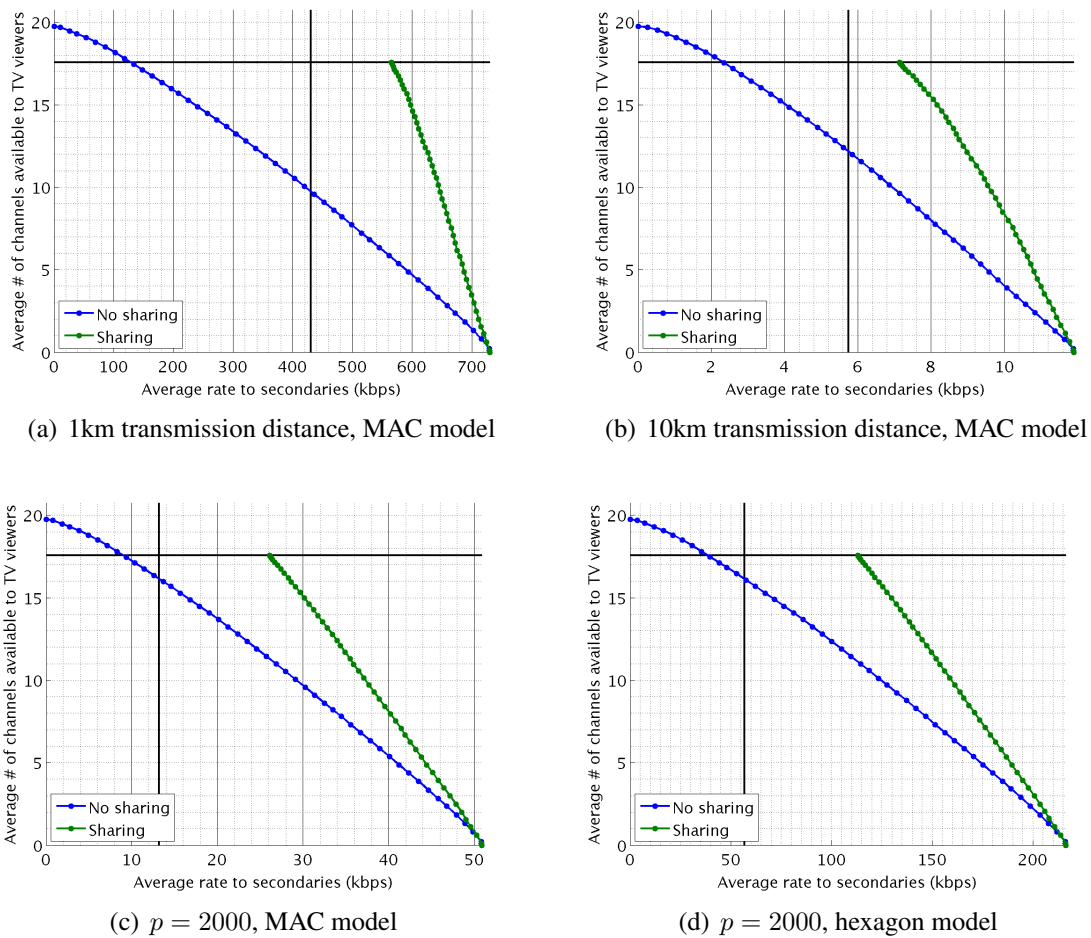
Figure 5.1: Tradeoff between number of TV channels and secondary data rate

the no-sharing model since the set of allowed (location, frequency) pairs under the sharing model is a strict superset of those in the no-sharing model.

There are two interesting features in this graph. First, note that the FCC is accepting a potential loss of over two channels per person on average just by allowing the existence of secondaries. This difference arises from the fact that a receiver near the limit of reception faces more noise with secondaries than without secondaries, thus fewer locations can receive TV. This issue was addressed in Chapter 3.

Second, to achieve the level of service that the FCC has already given to secondaries via refarming alone, they would have to relinquish an average of 6 *additional* channels per person in the $p = 2000$ cases, 11 *additonal* channels in the range = 1 km case, and 7 *additional* channels in the range = 10 km case. Thus whitespaces are more efficient than straightforward refarming for this level of secondary service.

Note that in these scenarios we have not accounted for where those TV stations will go should they be evicted from their current bands. This is considered in the next chapter, where we "repack" these incumbents into spectrum that will remain allocated for TV.

This section has shown that whitespaces appear to be better than typical channel reassignment given the FCC's desire to preserve most TV coverage areas. However the next section will show that whitespaces themselves can be designed in a variety of ways.

5.2   Making more whitespace: flexibility in the new approach

The selection of the protected radius (r_p) and the separation distance from the protected area ($r_n - r_p$) also represent a *choice* between secondaries and primaries. In this section we look at other options for these values and the tradeoff they present in terms of number of TV channels available and achievable data rate.

Here we consider a method of varying the value for r_p which is illustrated in Figure 5.2. This method was also developed in [47] but is recapitulated here for completeness. We correspondingly adjust the value of $r_n - r_p$ based on the transmit power of the secondary so that the desired fade margin is maintained.

Suppose that there exists only one TV tower, i.e. no other primaries nor secondaries. Due to signal attenuation and constant noise, the SNR for a receiver decreases as his distance to the TV tower increases. With digital television[3], the reception threshold is about 15 dB [45]; that is, when a receiver's SNR is below 15 dB, he can no longer watch TV. Thus there is already a naturally-defined maximum coverage area for each TV tower. We call this distance from the TV tower the 0 dB protected radius.

Clearly we cannot preserve this coverage area while allowing secondary operation. Any amount of additional noise, no matter how small, necessarily decreases the SNR and thus the coverage area. Therefore in order to add other transmitters on this channel it is necessary to sacrifice some of the coverage area, but how much?

We define an *eroded fade margin* to be the total amount, in decibels, by which the secondaries may decrease the primary receiver's SNR as compared to a clean channel SNR. For example, a margin of 3 dB allows secondaries to cause the same amount of noise as that found in a clean channel. At the same time, it decreases the coverage area to those areas which would have had an SNR of at least 15 + 3 = 18 dB in a clean channel. Conceptually, increasing the fade margin decreases the protected area. We see this illustrated in Figure 5.2.

Note that the FCC's choice of protected radius does not in general directly correspond to any particular fade margin.
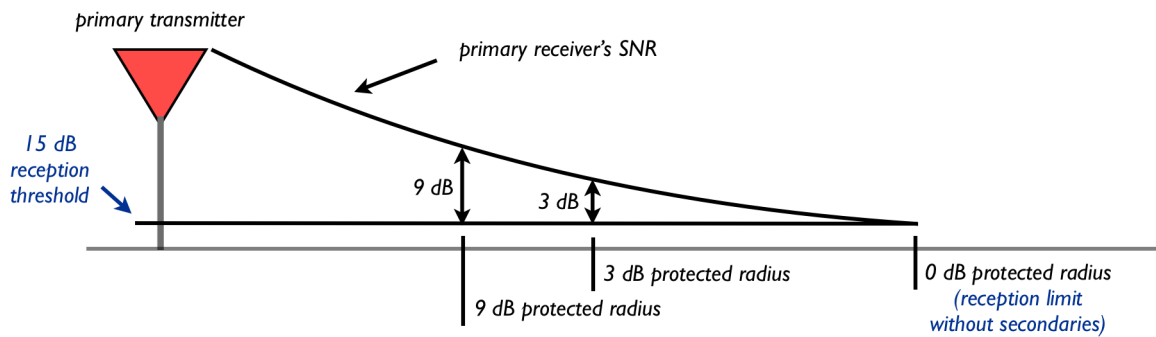
Figure 5.2: Fade margin illustration

In Figure 5.3 we see the effect on the secondaries of varying the fade margin. This helps us to visualize the impact of pollution and exclusions. Each point represents the median capacity (as sampled by population) of a single-user system with the corresponding exclusions.

Pollution alone appears to cause a great loss in capacity. However, recall that as in Figure 2.5(e) this is overstated since a complete lack of exclusions allows secondaries to use the extremely polluted areas they would otherwise be barred from.

Notice though that the noise has more of an effect on the 10-km range cases; this is due to the fact that the secondary's SNR is much lower due to signal attenuation over a larger distance. This also explains the relatively small effect that noise has in the $p = 2000$ case: polluted areas are typically near population centers and this high population causes the range to shrink, thus the signal is often greater when the pollution is higher.

Among the two types of exclusions, cochannel and adjacent-channel, we see that the latter often has greater impact. Intuitively this makes sense because each tower has (typically) two adjacent channels, thus there are effectively twice as many towers to exclude on adjacent channels as on the same channel. Naturally as the fade margin increases, the overall effect of exclusions diminishes[4].

---

[3]For simplicity, we assume that all current television stations are digital. In reality, many stations are still analog.

[4]The graphs in Figure 5.3 are misleading: as discussed in Section 2.3, shrinking the protected area indeed increases secondary utility but not uniformly. Areas near the TV tower are less valuable than those far away due to pollution, thus increasing the eroded fade margin has diminishing returns.
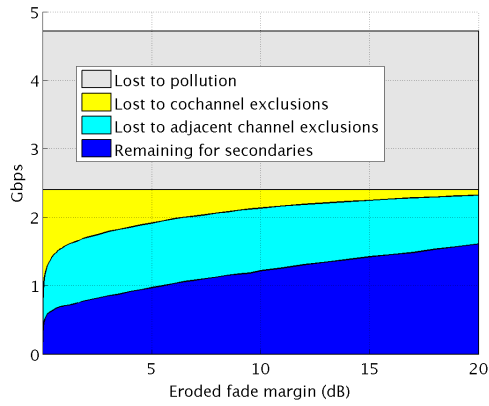
(a) 1 km range, single-link model
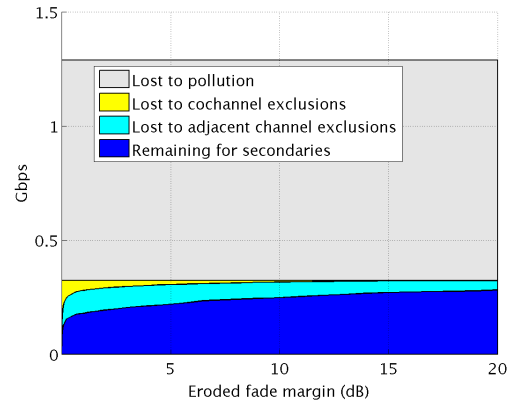
(b) 10 km range, single-link model
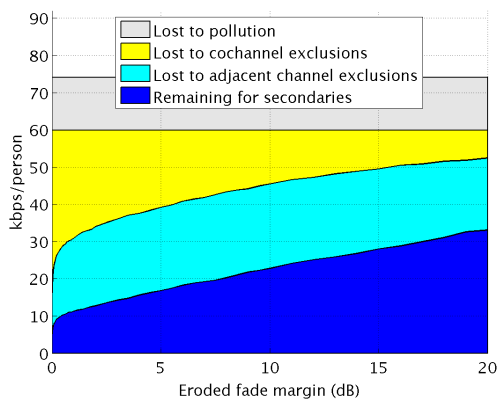
(c) 1 km range, MAC model
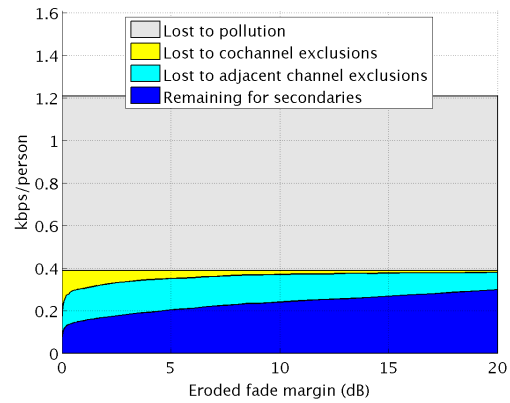
(d) 10 km range, MAC model
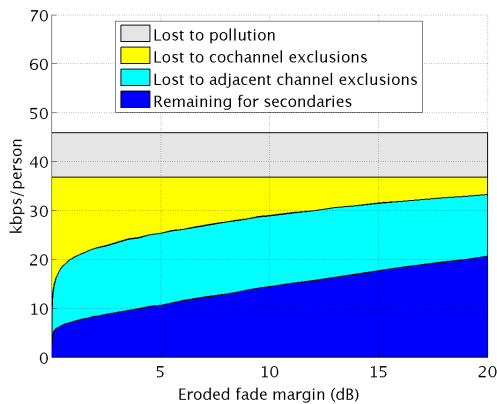
(e) $p = 2000$, MAC model
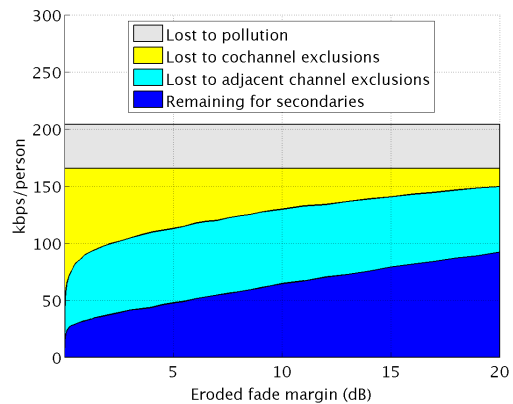
(f) $p = 2000$, hexagon model

Figure 5.3: Impact of TV noise and exclusions

5.3 Conclusion: whitespaces best for minimal primary impact

We have examined two potential alternatives to the FCC's chosen regulations. The first alternative follows the traditional spectrum-reallocation method by refarming TV channels for unlicensed use, shown as the blue curves in Figure 5.4. The second, shown in red, creates whitespaces but varies the size of the protected region[5].

We clearly see in all four secondary-deployment scenarios that whitespaces represent the best tradeoff if few TV channels can be sacrificed. Once we are willing to lose about three channels, simple refarming generally gives better results than whitespaces.

Another option is to use a combination of refarming and whitespaces, shown by the green line in Figure 5.4. Refarmed channels are unlicensed but maintain a power limit and the remaining TV channels are used as whitespaces under the current FCC regulations[6]. This method always outperforms the simple channel-refarming scenario but of course they converge to the same point when all of the channels are given over to unlicensed use.

We will see in the next chapter that a further blend can be achieved by "trading" whitespaces for incumbent service. This is accomplished by altering the operating channels of incumbents so that they transmit in the would-be-whitespaces in the remaining TV spectrum rather than simply evicting them.

It is important to notice that even with relatively small excluded regions (i.e. the ends of the red and teal curves), whitespaces cannot achieve the same data rates as refarming. This is due to the inherent pollution from TV towers that secondaries must face. If a higher rate is desired, refarming is the only option.

It is important to point out that the FCC's tradeoff point (the black crosshair) does not lie directly on the red line which represents varying the eroded fade margin using both cochannel and adjacent-channel exclusions. This is likely due to the fact that the FCC assumes the use of directional antennas for TV receivers but we do not.

These conclusions are important to take under advisement when setting the rules: they tell us when the whitespaces are the right choice and when they are not.

---

[5]This is identical to the eroded-fade-margin (EFM) method of the previous section. For each margin value we have calculated the TV reception maps in order to find the average number of channels remaining for TV viewing. This allows us to plot the number of viewable channels against the secondary utility in the EFM scenario.

[6]For simplicity we do not consider adjacent-channel exclusions in this calculation. In the future we would like to create a more sophisticated model which incorporates these effects.
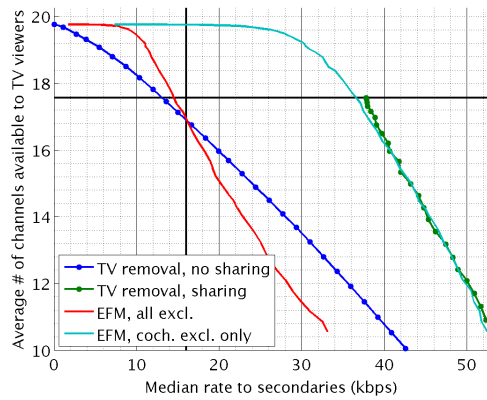
(a) 1 km range, MAC model
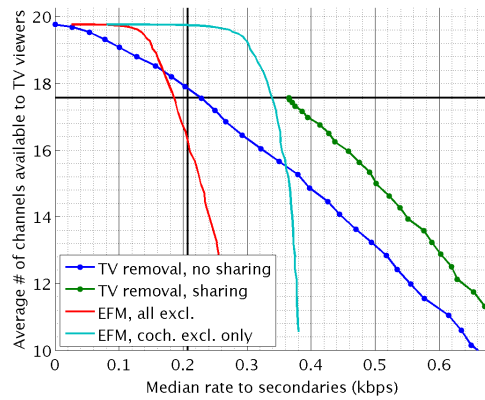
(b) 10 km range, MAC model
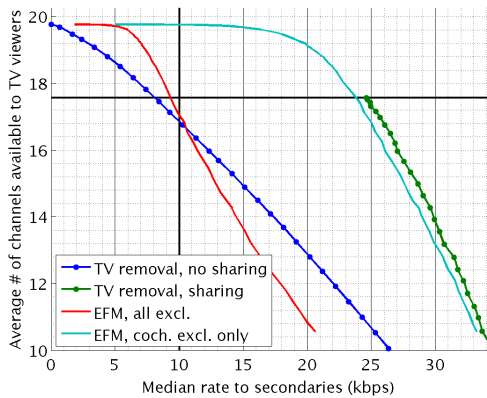
(c) $p = 2000$, MAC model
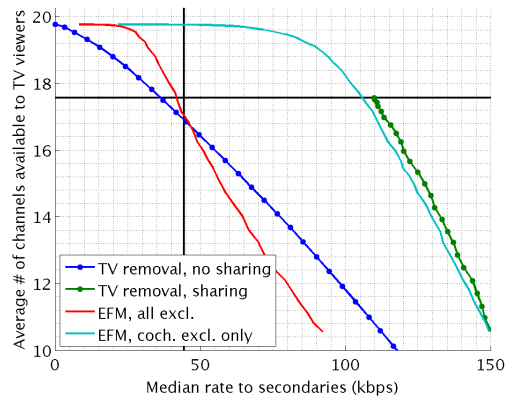
(d) $p = 2000$, hexagon model

Figure 5.4: Comparison of TV removal and eroded fade margin (EFM) schemes. The black crosshair represents the FCC's chosen tradeoff point.

Chapter 6

Repurposing vs. whitespaces, part II

Some or all of the work in this chapter appeared in [61].

Spectrum has traditionally been allocated for single uses and by now most of the "prime" spectrum has well-entrenched incumbent users. When a new service needs spectrum, there are two qualitatively distinct ways of making bandwidth available for it. A swath of incumbent users can be removed from a band, with the cleared band being reallocated for the new service. Alternatively, the new users can be allowed to utilize the interstitial spectrum holes (i.e. whitespaces) between incumbent users, with the requirement to protect the incumbents' QoS. But these can also be used in combination by partially clearing a band and opening up the rest for whitespace-style sharing, as seen in the previous chapter. In this chapter, we consider the case where regulators are additionally able to "repack" incumbents, e.g. alter their operating channels, which can reduce the need to evict them. An open question has been how whitespaces and partial spectrum clearing interact with each other and the ability to repack incumbents. Do efficient repacks completely eliminate whitespaces?

The USA FCC's upcoming incentive auction in the TV bands is the first large-scale attempt to repack a major band of spectrum in order to clear spectrum for LTE. This auction is meant to navigate the tradeoff between incumbent TV services and LTE networks. In preparation, the FCC has made a large and complex data set of repacking constraints available for the first time. We have repurposed this data and built our own repacking engine in order to study a more general version of the tradeoff between whitespaces and cleared spectrum.

We conclude that (1) repacking enables clearing of significantly more spectrum than just removing incumbents; (2) the total amount of spectrum available for new uses is relatively insensitive to how incumbents are removed; (3) efficient repackings basically trade whitespace spectrum for cleared spectrum; (4) even the most efficient repackings leave plenty of whitespace — an amount that can be comparable with the amount of cleared spectrum.

This work leverages WEST in a major way, as described in Section B.1.3. In particular, it uses WEST to compute tens of thousands of estimates of whitespace and TV availability under various

hypothetical repacking scenarios. In the face of uncertainty regarding repacking outcomes, we computed many plausible outcomes for each scenario and aggregated them using means, medians, and standard deviations. Although these are standard statistical metrics, they had not been applied to whitespaces before this work.

6.1   Introduction

TV spectrum has recently become a very popular topic due to its proximity to mobile spectrum as well as the TV whitespaces, which give access to spectrum necessary for economic development. There are many interesting aspects to the field of cognitive radio and whitespaces, such as co-existence techniques, network planning, system architecture, and security and robustness, whose unique challenges have been studied to varying degrees. However, few studies address a very simple question: when is it better to completely reallocate a band vs. to share it?
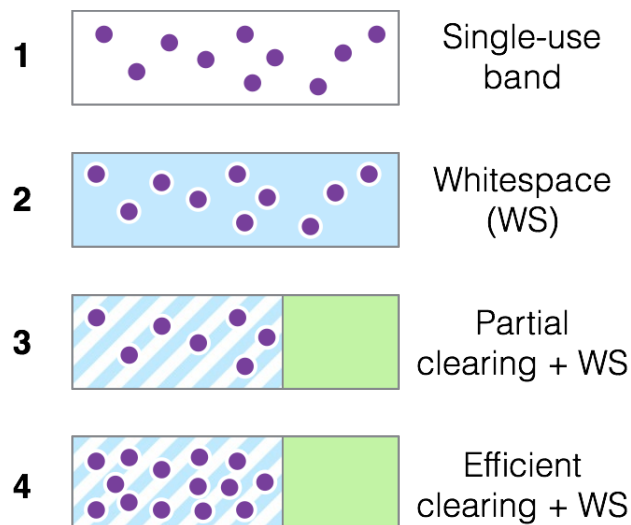
Figure 6.1: An illustration of the various options for spectrum repurposing. Incumbents are shown as purple dots while whitespaces are blue and cleared spectrum is green. White represents unused spectrum (in the case of whitespaces, this is a buffer which is necessary to maintain the incumbent's quality of service). The white and blue hashed pattern represents spectrum that could but need not support whitespace rules.

In fact, there are several different options for making "new" spectrum, as shown in Figure 6.1:

1. *Completely reallocate the band as a single-use band*. Until recently this was the standard way of reallocating spectrum. Complete clearing is especially useful for applications which cannot or will not share spectrum.

2. *Declare the entire band potential whitespace while preserving the quality-of-service of the incumbents via sharing rules.* This is becoming the de facto way of "generating" new spectrum, especially after the publication of the PCAST report[1] [54]. Whitespace regulations naturally have to navigate a tradeoff between quality-of-service for the incumbent vs. the secondary users. This has been explored in Chapter 5 as well as [62, 63].

3. *Partial clearing of the band.* Pristine spectrum is carved out while a portion of the incumbents remain. The uncleared spectrum may be designated as either single-use spectrum or as whitespace with the incumbents as the primary users. Partial clearing is preferable when it is not possible or desirable to remove all of the incumbents.

4. *Efficient partial clearing of the band.* The spirit and use cases are very similar to scenario 3 except that this option maximizes the number of incumbents that remain after a partial clearing. Rather than remove the incumbents which were in the now-cleared spectrum, these incumbents are efficiently packed into the remaining (uncleared) spectrum whenever possible. This approach essentially sacrifices would-be whitespace in order to "house" an incumbent[2]. Thus more incumbents remain in service at the expense of secondary users.

In the previous chapter, we considered the first three scenarios which are relatively easy to analyze. However, the FCC's upcoming incentive auctions provide an excellent chance to study the fourth option in the context of the United States. Briefly, the incentive auctions give television broadcasters a chance to bid one of two choices: (1) relinquish their spectrum usage rights or (2) be "repacked" (i.e. moved to another channel) within the TV bands[3]. As overseer of the entire auction, the FCC will subsequently decide which stations will be "repacked" vs. removed, in the process creating a situation akin to scenario 4. The cleared spectrum will be auctioned off in a manner which encourages prospective LTE-system builders (e.g. AT&T, Verizon, Sprint, and T-Mobile) to buy it. In the interest of focus, we have simplified the incredibly complex incentive auction. We encourage interested readers to read Section B.4 which provides a more detailed overview of the auction.

"Repacking" incumbents is not a trivial task (as proof: the FCC is currently being sued by the National Association of Broadcasters over its repacking methodology [64]). However, the FCC has made their repacking process and constraints public [65] which gives us the opportunity to conduct hypothetical spectrum repacks of our own, independent of the auction's actual course. Thus in this chapter we use the incentive auction data to explore the fourth spectrum-scrounging scenario.

---

[1]This report, submitted as a recommendation to the President of the United States by the President's Council of Advisors on Science and Technology in 2012, emphasized the need to find at least 1,000 MHz of spectrum as soon as possible and highlighted spectrum sharing as the best way to accomplish this goal.

[2]Note that this is not an option in smaller bands where the incumbent has a fixed bandwidth equal to the width of the band. However, there are many channelized bands which would be good candidates for this kind of solution.

[3]Broadcasters are actually given several choices [13, ¶365] but in this chapter we simplify to the most important choice for brevity and clarity.

We first look at how many incumbents would need to be removed in order to meet a variety of spectrum clearing targets and show that the ability to efficiently repack drastically reduces the need to remove incumbent users. We show that repacking also concentrates the removal of incumbents to the geographic areas where it is strictly necessary, reducing unnecessary loss of incumbent services. Although efficient repackings aim to pack the spectrum as tightly as possible, we find that interstitial spectrum holes remain even in the places which already had few whitespaces.

Finally, we explore the true tradeoff between incumbent services (e.g. TV) and spectrum for new services. We find that for the same sacrifice of incumbent service, the total spectrum available to an opportunistic device that could use both TV whitespaces and cleared spectrum is surprisingly insensitive to the clearing method. The effect of repacking efficiency is in modulating the tradeoff between whitespaces and cleared spectrum, rather than between incumbent and new services. In particular, more efficient repackings reduce whitespaces but create an equivalent amount of cleared spectrum.

6.1.1 Prior work

The idea of incentive auctions can be traced back to the proposal in 2002 by Kwerel and Williams of the FCC, which pressed for a rapid transition to a market-based allocation of spectrum and specifically called for a large-scale two-sided auction for repurposing spectrum from incumbents who are willing to relinquish their rights [66]. The National Broadband Plan of 2010 [20] proposed the use of such a two-sided "incentive auction" for repurposing spectrum from broadcast television services to mobile broadband services (LTE). In February 2012, Congress authorized the FCC to conduct these auctions through what came to be known as the Spectrum Act [67]. Soon after, the FCC announced the preliminary plan for this auction in their October 2012 Notice of Proposed Rulemaking (NPRM) [68]. A report and order released by the FCC in June 2014 laid out the structure of the auction process [13] and a follow-on Public Notice released in December 2014 [69] provided additional details. Since its inception, the economic value, potential impact, and complexity of this unique auction has generated a lot of interest in both the business and academic communities. At least three major mobile broadband providers, AT&T, Verizon Wireless and T-Mobile, have commissioned teams of researchers to perform speculative analyses of the auction [70–73].

Arguably the most novel and challenging feature of the incentive auction process is its intricate entanglement with the repacking problem, which concerns allocating a set of stations to a set of channels given constraints derived from the physical nature of the problem. This problem can be naturally posed as a large scale Boolean Satisfiability problem (see [74]) with tens of thousands of variables and hundreds of thousands of clauses in a typical instance.

The constraints themselves have been released on FCC's LEARN website [75], a website intended to help the public understand how the incentive auctions work. The data is in the form of two files. The first file lists, for each of the 2,173 repack-eligible TV stations, the list of channels to

which it may be assigned[4]. The second file contains 291,739 entries, each of which details the co- or adjacent-channel interference constraints between TV stations (e.g. "station A may not operate cochannel to station B on channel C").

A study of this repacking problem focusing on the computational difficulties and the performance of SAT solvers can be found in [74]. Further, the FCC has conducted a public workshop to help disseminate information about this complex problem, a webcast of which can be found online [76]. Recently, using the data released by the FCC, an analysis of the feasible repackings corresponding to a variety of contingent spectrum clearing targets was done in [70] (commissioned by AT&T). We use some of the techniques developed there as the starting point of our analysis.

6.1.2   Brief overview of our methods

To generate the various sets of data used in this chapter, we rely on data from the FCC's LEARN website [75], as described above. We then use PycoSAT, a Python wrapper for the well-known SAT solver library PicoSAT [77], in order to synthesize the constraints and output a feasible repacking. Two other studies on repackings have used PicoSAT [70, 72] and it was also featured in an FCC workshop on the topic of repacking in the incentive auctions [76].

Finally, we build on this data via our Whitespace Evaluation SofTware (WEST), an open-source toolbox for computing the amount of available whitespace described in Chapter 11. Complete methodological details are in Chapter B.

Note that because it is a combinatorial problem, there are many candidate station assignments that achieve the same goal (e.g. clear N channels by removing exactly M stations). For each possible scenario, we generate 100 candidate assignments which are later presented as aggregate statistics. Typically, we will use the median value (as taken over all assignments) as this is a standard and robust metric.

6.2   How many incumbents must be removed to free spectrum?

There's no such thing as a free lunch. However, there are good and bad ways of removing incumbents in order to clear new spectrum. A naive way is to remove precisely the incumbents which happen to be in the channels to be cleared, i.e. scenario 3 in Figure 6.1. However, this immediately leads to a few problems:

1. More incumbents will be cleared than necessary – compare the two lines of Figure 6.2. The difference between these lines represents the number of stations that can be "removed" via repacking rather than taken off the air.

---

[4]External factors, such as harmonization with Mexico and Canada, sometimes prevent the assignment of a particular station to a particular channel.

2. In some cases (e.g. with TV) it can be difficult to assess the value of the incumbent. So although we could consider invoking something akin to eminent domain, it's unclear what the fair market price would be. This means that any offered price would likely be challenged, delaying the reallocation of spectrum and creating uncertainty for all parties.

3. If only the incumbents which happen to be in the channels to be cleared participate in the market, the lack of competition could potentially lead to obvious problems such as holdouts (stations demanding unreasonable sums of money because they have a good bargaining position).

For these reasons it is important to any market-based clearing process that we have a means of substituting one station for another in order to foster competition. In the incentive auctions, this substitutability is facilitated by the ability to repack TV stations.
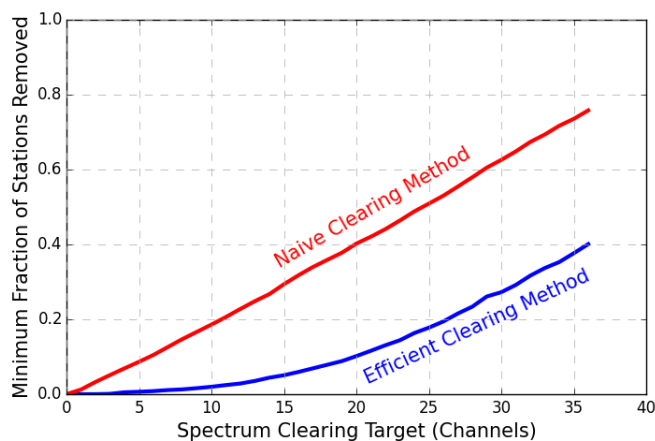
Figure 6.2: Minimum number of incumbents (TV stations) that need to be removed to meet each spectrum clearing target. In this figure we compare a naive clearing method (scenario 3 in Figure 6.1) and an efficient clearing method (scenario 4). Efficient clearing is accomplished via "repacking" (moving in frequency but not space) existing TV stations rather than simply removing them.

The blue line in Figure 6.2 shows the results of our computations for determining the minimum number of stations that *must* be removed in order to meet different spectrum clearing targets. The efficient-clearing numbers are substantially lower than the naive approach of removing all stations which happen to be in the desired band.

We perform these computations by building on the techniques in [70], where a similar computation was done with additional constraints that are motivated by those of the actual incentive auction (e.g. a station may only be reallocated to a channel near its original channel). We found that about 10% fewer stations need to be removed to meet the same clearing targets as compared to those reported

in [70]. This was partially a result of our exploiting certain properties of the SAT solver to improve its performance. The numbers could indeed be even lower; further improving the performance of SAT solvers on these problems is a topic for future work.

6.3   Television availability after repacking

Another way of looking at the impact of the results in Figure 6.2 is to examine which places lose access to TV. Figure 6.3 shows which places lose at least one TV channel under the naive clearing method (orange), the efficient clearing method (green), or both (blue) when 14 TV channels are to be cleared.

Our first observation is that there are almost no places where TV availability is impacted in the efficient clearing method but not the naive method. We further see that with the efficient clearing method, only stations in the most populous markets are affected. In contrast, the naive clearing method impacts a very large fraction of the population in addition to having a much greater impact on rural areas. Thus the ability to repack stations as opposed to simply removing them helps us confine the impact to only those places which are unavoidably impacted. We provide more detailed maps showing how many TV channels were lost in each market[5] in Chapter B.

Beyond answering the question of how much TV coverage will be lost, this figure is important because it gives insight into how the repacking will work. In particular, we see that the areas which currently have a lot of TV stations (e.g. New York City, Los Angeles) have so many that some must be removed rather than repacked. However, most of the country is not brimming with TV stations (as evidenced by the current amount of whitespace in these regions) and so no stations would need to be removed to meet most clearing targets.

6.4   Whitespaces remain after repacking

The key question regarding the incentive auction within the cognitive radio community is "will whitespaces still exist after the incentive auctions?" The answer to this is quite simply: yes. In this section we will look at the expected *minimum* amount of whitespace that will remain for each spectrum clearing target. As before, we will consider the case where the minimum number of incumbents are removed. If more incumbents are removed for the same clearing target, *more* whitespace will be made available[6].

There are several ways to measure the amount of whitespace that will remain after the auction, such as:

---

[5]In the efficient clearing method, San Francisco, Los Angeles, and the most populous portions of the east coast lose more than four channels while other areas do not. With the naive clearing method, a variety of regions (including e.g. Utah) lose 6 or more TV channels.

[6]Each repacked incumbent will be protected and hence will diminish the amount of whitespace available. Fewer incumbents means fewer restrictions on whitespaces.
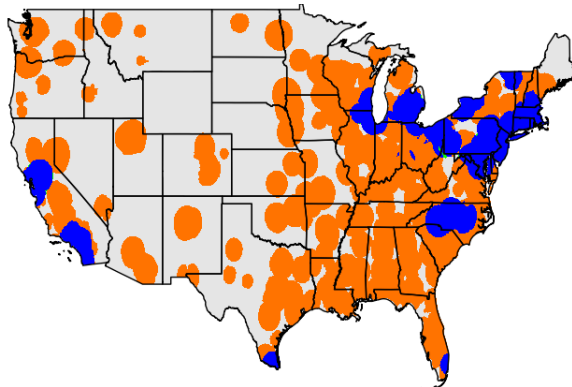
Figure 6.3: Map showing which locations in the United States may lose TV coverage after the incentive auctions if 14 TV channels are repurposed using the naive clearing method (orange), the efficient clearing method (green), or in both cases (blue). Gray denotes areas whose TV coverage is not affected.

1. Raw amount of whitespace after the auction

2. Change in whitespace due to the auction

Figure 6.4 shows the predicted number of whitespace channels available to fixed whitespace devices if the incentive auction clears 7 TV channels. We see that the major population centers are indeed feeling a bit of a spectrum crunch (as they always have) but that at least 30 percent of the population has more than 15 whitespace channels.

To get a better sense of how much whitespace is likely to remain after the auctions, we turn to Figure 6.5. Two sets of CCDFs are shown, one for fixed devices and the other for portable devices[7]. CCDFs for the current assignment of TV stations are shown in black for comparison and the other lines represent the expected minimum amount of whitespace that will remain with a variety of spectrum clearing targets. The median US citizen will have approximately 4 whitespace channels (24 MHz) available for fixed devices (and 14 channels—84 MHz—for portable devices) in an auction which clears 21 TV channels (the red line), one of the more optimistic auction outcomes in the FCC's eyes.

---

[7]The main distinction between fixed and portable devices is that fixed devices may not operate in a channel adjacent to a TV station whereas portable devices may do so (at a lower power). In this chapter we do not consider the standard operating-channel restrictions applied to portable devices because we expect those to change after the incentive auction [40]. It is also worth noting that fixed devices may be able to operate under portable-like rules (adjacent-channel but with lower power) in the future.
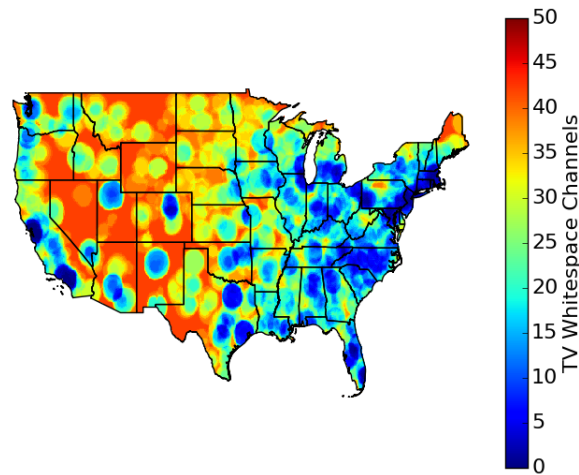
Figure 6.4: Number of expected whitespace channels for fixed devices if 7 channels are reallocated in the incentive auctions.

Next, we directly compare the expected amount of whitespace after the auction to the current amount of whitespace. This is shown via 2-D histograms in Figure 6.7 for portable devices and Figure 6.8 for fixed devices. Red indicates that many people fell into that particular category whereas blue indicates an event affecting fewer people. The two diagonal lines correspond to (1) having lost no whitespace (the top line) and (2) having lost an amount of whitespace which is equal to the number of channels cleared in the auction. The horizontal line indicates the maximum possible number of whitespace channels (channels above this line have been cleared and are not counted as whitespace). We describe the generation of these two-dimensional histograms in greater detail in Chapter B.

People who lose no whitespace in the auction are often in places where a TV station was removed in order to meet the clearing target, and so the channel that was cleared was never available as whitespace in the first place. In essence, *TV* channels were taken away, not whitespace channels. We can see from Figure 6.3 that this is predominantly in urban areas. People near the second diagonal line had enough free channels that all of their TV stations could be repacked, thus channels that were previously whitespace are the ones that have been cleared. Although this happens mostly in areas of low population density, the sheer quantity of them means that they represent almost 50% of the US population for the situation in Figure 6.3.

We note a few important features of these plots:

1. As is well-known, fixed devices in general have less available whitespace than portable devices (this is due to differences in the regulations for these devices).
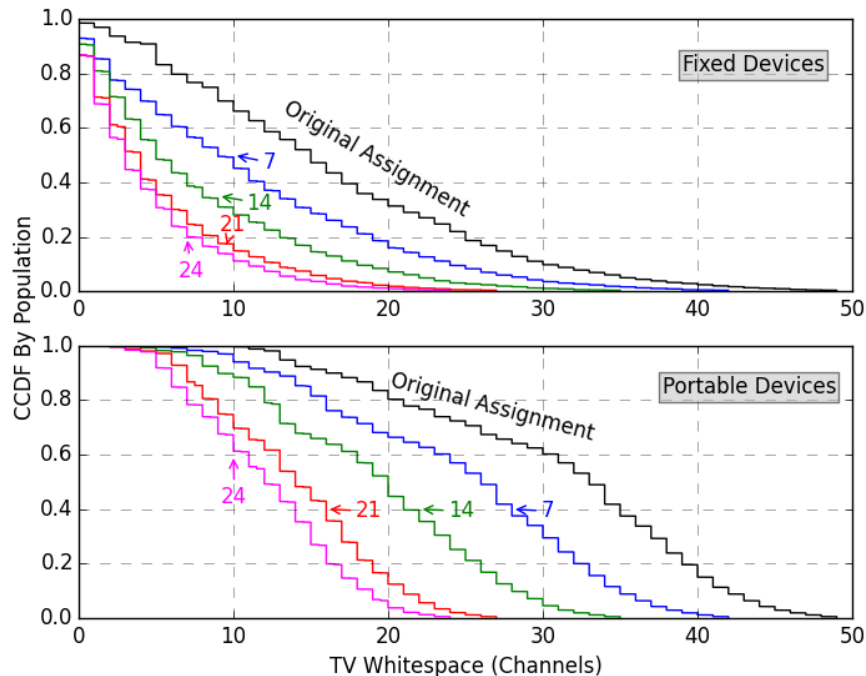
Figure 6.5: CCDF weighted by population showing the minimum amount of *TV* whitespace which is expected to remain after the incentive auction. Note that cleared spectrum is *not* included in these numbers.

2. When few channels are reallocated, the mass is mostly clustered around the lower diagonal line. The distribution of post-auction whitespaces becomes more variable if more channels are cleared.

3. In the case of fixed whitespaces, most people have few channels to begin with but tend to keep them.

4. Portable devices are more likely to experience a reduction in whitespaces. Essentially they "had more to lose."

Finally, Figure 6.6 compares the expected whitespace after a repack with N channels cleared (the y axis) to the amount of whitespace resulting from a naive reallocation which also results in N channels cleared (i.e. scenarios 3 and 4 in Figure 6.1). We see that the majority of the mass is falling below the $y = x$ line, indicating that whitespaces are more plentiful with the naive reallocation. The extra whitespace in the naive reallocation indicates an inefficient use of spectrum for the primary as compared to the appropriately-named efficient allocation.

6.5 Whitespaces vs. reallocation

As mentioned in the introduction and shown in Figure 6.1, there are several basic ways to create new spectrum opportunities. We explore these options in terms of the tradeoff between delivered services for the incumbent vs. a new spectrum-hungry device. We consider three types of devices: (1) those that want their own dedicated bands (e.g. LTE devices); (2) those that can operate only in the TV whitespaces; and (3) devices which are willing to harvest any spectrum possible.

Figures 6.9 and 6.10 both show this tradeoff in terms of the number of over-the-air TV channels the median US citizen could watch versus the amount of spectrum that the median citizen could access for different spectrum clearing targets. (We describe the process of construction of these figures in greater detail in Chapter B.) Each figure has two types of lines which parallel those in Figure 6.2:

1. Dashed: naive removal method (stations are removed from highest channels first; no assignments in uncleared spectrum are modified). This is scenario 3 in Figure 6.1.

2. Solid: minimal removal + optimized repacking method. This is scenario 4 in Figure 6.1.

In particular, Figure 6.9 shows the amount of completely cleared spectrum (in purple) as well as the amount of spectrum that could be obtained by an opportunistic portable device which can operate in either the TV whitespaces or the cleared spectrum[8] (in red).

When no TV viewership can be sacrificed (i.e. the top points on all three lines), being able to use whitespaces significantly increases the amount of available spectrum. However, sacrificing even one watchable TV channel (for the median US citizen) with efficient clearing gets us almost as much *cleared* spectrum as fixed whitespace. Portable whitespace will always outperform fixed whitespace because of the nature of the protections afforded to the incumbents which depend on the device type.

The most interesting curves are the right-most ones (in red) in Figure 6.9. These curves show that the total of cleared spectrum and portable whitespaces is essentially the same under the naive clearing method and the efficient clearing method. Moreover, that final curve has a clear slope of -6: to gain 6 MHz (one channel) of available spectrum, one watchable TV channel must be sacrificed. This shows that from the perspective of the TV-vs.-new-device tradeoff, the repacking method affects the *balance* between cleared spectrum and whitespace spectrum, *not* the total amount of spectrum available to an opportunistic whitespace-like device. In terms of total-spectrum, it is essentially a zero-sum game at the margin.

---

[8]The spectrum cleared in the incentive auctions will be available for use by whitespace devices subject to rules that protect the spectrum purchaser's rights as a primary user of the spectrum [13, ¶678]. Because it is difficult to predict what the distribution of new primary users will be or even what their protection criteria will be, we treat cleared spectrum as 100% available (but not whitespace) for the purposes of this chapter.

To understand this, it is useful to reorder the curves to focus on the whitespaces, as is shown in Figure 6.10. The red line (total spectrum) is repeated for illustrative purposes. Here, first notice that TV whitespaces and watchable TV channels shrink *together*. While this is not the normal look of the TV-vs.-whitespaces curve, the effect is explained by the fact that cleared channels take away from both TV and whitespaces. The behavior is intuitive because only the interstices within *TV spectrum* are deemed whitespaces and as the total size of the TV band shrinks, it is natural to expect the whitespaces to shrink along with it. Next, notice that for the same amount of watchable TV, the naive clearing method gives rise to far more fixed and portable whitespace than the efficient clearing method. This is because the efficient clearing method is in effect filling in the whitespaces with repacked stations that are being relocated down from the cleared spectrum instead of being evicted.

Whitespace can take advantage of assignment inefficiencies in a way that reallocation simply cannot, so in any practical scenario there will always be a significant whitespace gain. For example, even in the 24-channels-cleared plan that the FCC is considering to reallocate 144MHz of spectrum to LTE, we anticipate there will still be 75MHz of portable whitespace left in the remaining TV bands for the median US citizen.

6.6 Conclusions

This chapter has opportunistically used the upcoming TV spectrum incentive auction in the USA as a way to examine the tradeoff between clearing spectrum and enabling the use of whitespaces.

We saw that, with efficient clearing, over-the-air TV availability will be affected primarily in major metropolitan areas where it is necessary to evict TV stations in order to achieve most spectrum clearing targets.

Overall we conclude that:

1. Repacking enables clearing significantly more spectrum than does just removing incumbents.

2. The total amount of spectrum available for new uses is relatively insensitive to how incumbents are removed.

3. Efficient repackings basically trade whitespace spectrum for cleared spectrum.

4. Even the most efficient repackings leave plenty of whitespace spectrum — an amount that can be comparable with the amount of cleared spectrum.

This supports the idea that spectrum sharing between heterogeneous uses is essential for full utilization of spectrum [78].

(a) 7 channels reallocated

(b) 14 channels reallocated

(c) 21 channels reallocated

(d) 24 channels reallocated
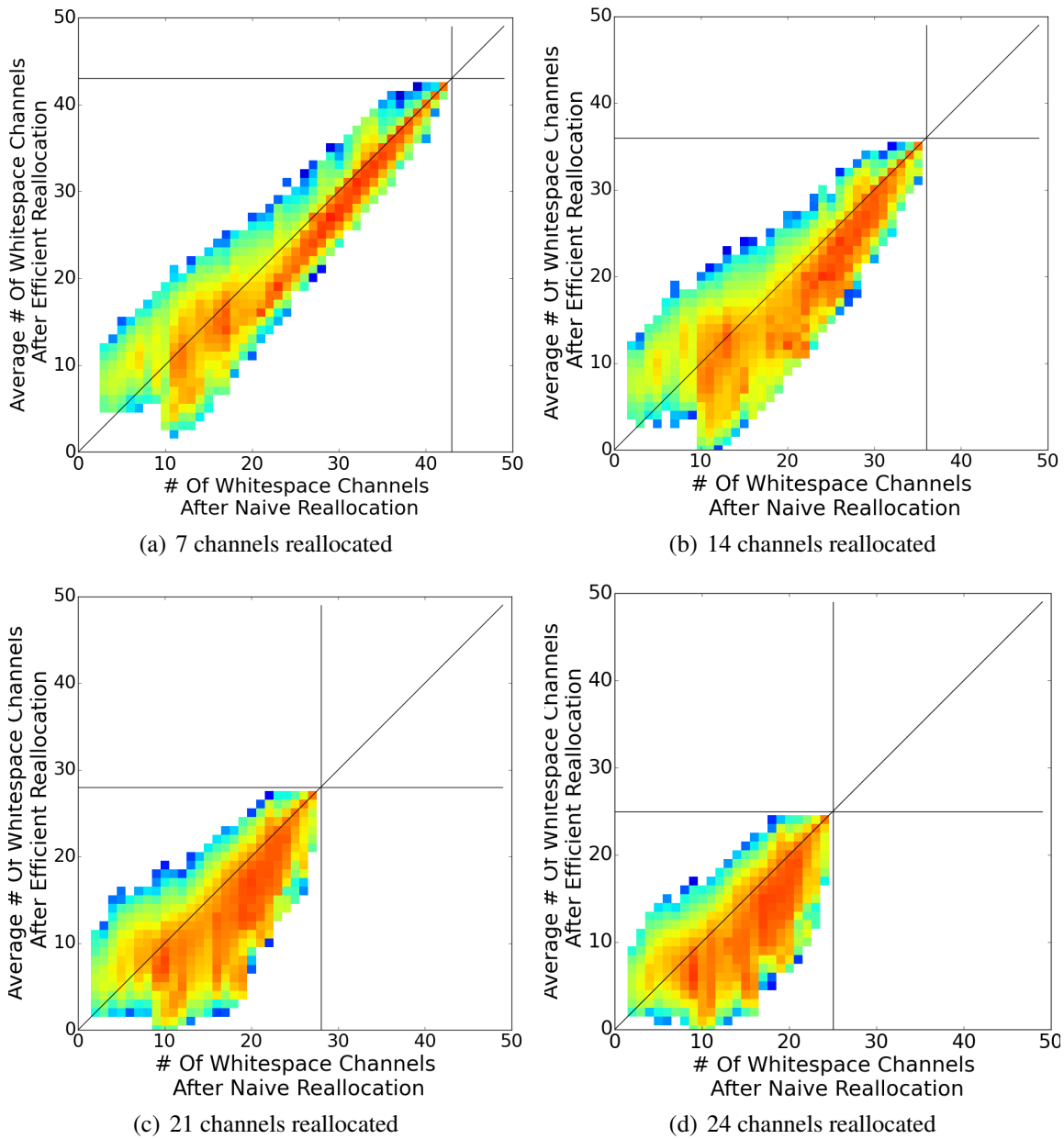
Figure 6.6: Direct comparison via 2-D histogram of whitespaces available to *portable* devices in scenarios 3 and 4 of Figure 6.1. Pixels are colored based on the number of people in the bin with red indicating the most people and blue the fewest.

(a) 7 channels reallocated

(b) 14 channels reallocated

(c) 21 channels reallocated

(d) 24 channels reallocated
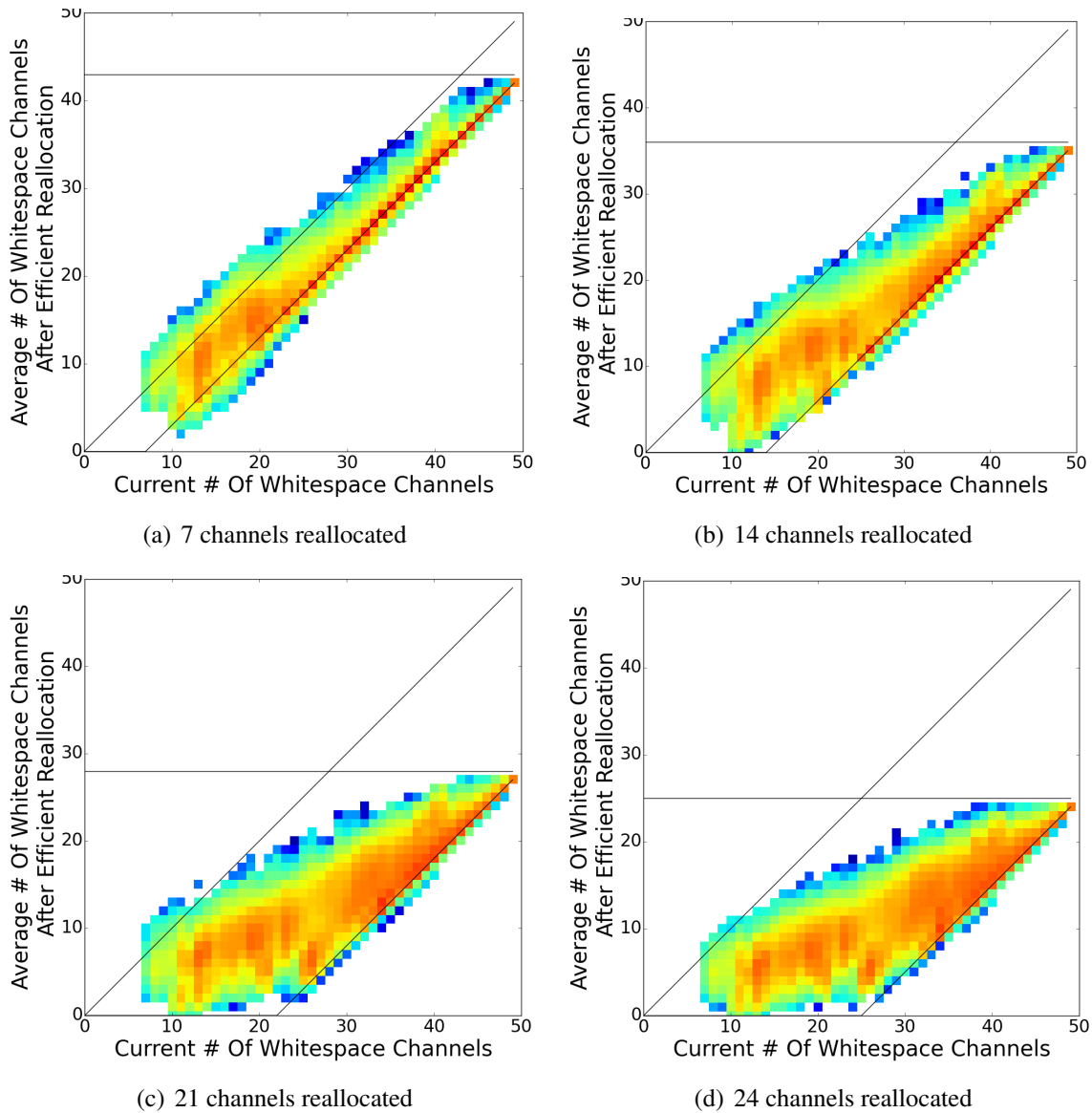
Figure 6.7: Direct comparison via 2-D histogram of whitespaces available to *portable* devices before and after the the auction. Pixels are colored based on the number of people in the bin with red indicating the most people and blue the fewest.

(a) 7 channels reallocated

(b) 14 channels reallocated

(c) 21 channels reallocated

(d) 24 channels reallocated
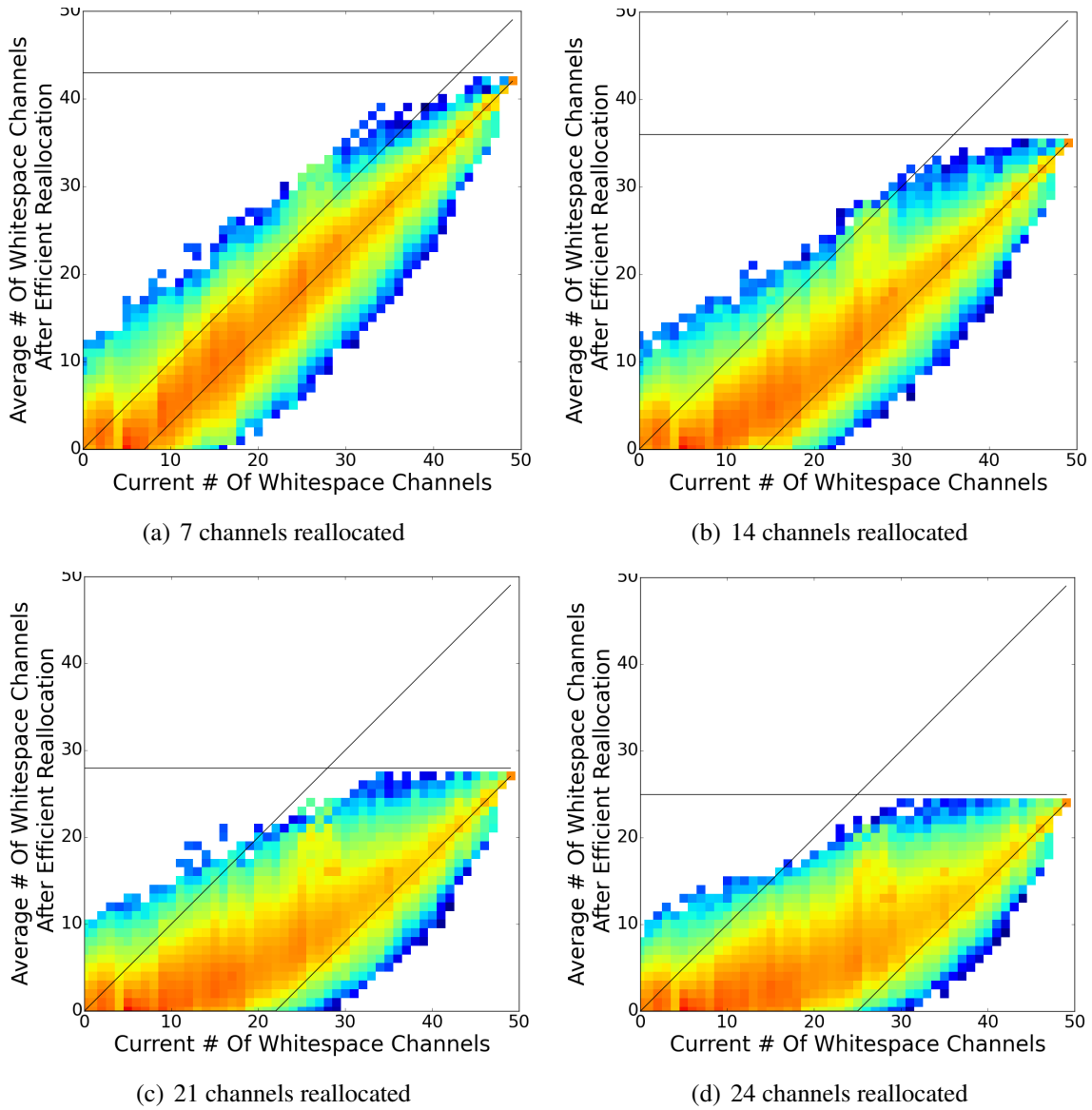
Figure 6.8: Direct comparison via 2-D histogram of whitespaces available to *fixed* devices before and after the the auction. Pixels are colored based on the number of people in the bin with red indicating the most people and blue the fewest.
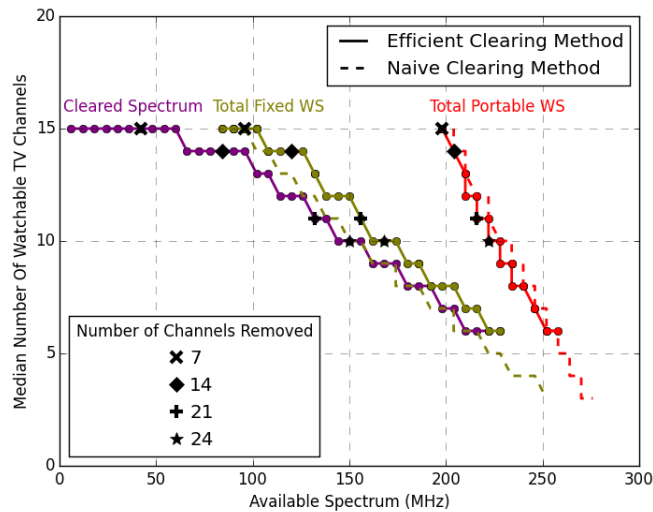
Figure 6.9: Tradeoff between over-the-air TV service and available spectrum. This takes a cleared-spectrum centric view with the first curve being cleared spectrum and the subsequent curves adding in the amount of fixed and portable whitespace, respectively. Note that the "Total Portable WS" line in this figure is the same as the "All Mobile Spectrum" line in Figure 6.10.
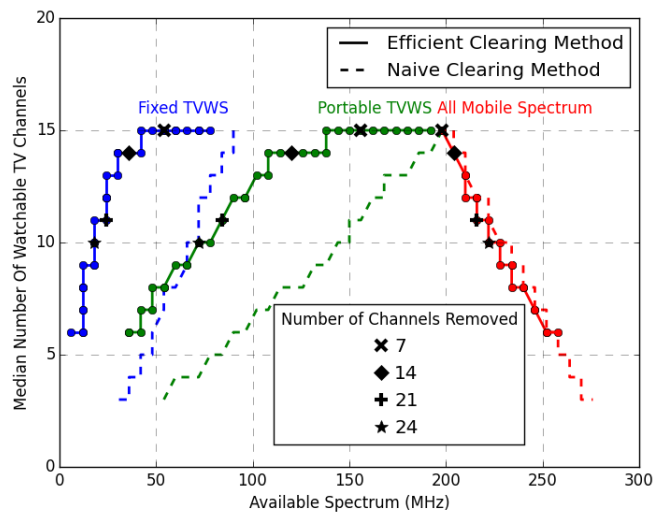
Figure 6.10: A whitespace-centric view of the tradeoff between over-the-air TV service and available spectrum. Note that the "All Mobile Spectrum" line in this figure is the same as the "Total Portable WS" line in Figure 6.9.

Part IV

Improving the whitespace ecosystem

Chapter 7 introduces a "refactored" whitespace architecture which is backwards compatible but more flexible than the current architecture. This architecture has positive implications for the certification process, making it more rigorous, predictable, and faster. It also has the potential to reduce the cost of whitespace devices and allow device manufacturers to individually choose points on the tradeoff curve of price and spectrum availability. Furthermore, it leverages existing infrastructure—rather than forcing it to be duplicated—to improve access to spectrum.

Chapter 8 explores the tradeoff between device design and spectrum availability. Specifically, it addresses the problem of weakly-localized devices (e.g. a GPS-enabled device operating indoors) which wish to utilize spectrum whitespaces. Under the current architecture, the only real option is to have a nearby device which *can* geolocate (e.g. a GPS-enabled device operating outdoors) which assists the weakly-localized device. We use this chapter to explore the alternative solutions that our proposed architecture opens up.

Chapter 7

Whitespace architecture

Some or all of the work in this chapter appeared in [79].

The PCAST report, released in July 2012, lays out a vision in which whitespaces refer not only to unused spectrum in the television bands but also to similar spectrum in radar-based and certain space-to-ground Federal bands. Because of the anticipated popularity in spectrum sharing, it is important to design good regulations which promote innovation.

While some whitespace devices will be self-sufficient ("masters"), others will rely on help from other devices in order to access the whitespaces ("slaves"). Currently, this help is provided by a single master device. In this chapter, we argue that (1) this assistance need not be provided by a single device and (2) the assisting device need not be a whitespace device. Instead, we can think of the "slave" as being helped by a whitespace device support network, i.e. a variety of devices which each supply a piece of the whitespace access puzzle.

We begin by identifying the three key components of a whitespace device support network. We describe each component in detail before giving example deployments which are only possible with a support network.. In one example, a smartphone plays the role of the "master" by providing the "slave" device with a location as well as a means to access the database.

Next, we remark on the advantages that this separation provides when it comes to certification. In particular, regulators can now perform unit tests to verify that each component operates correctly on its own, rather than certifying an entire device all at once.

We also discuss security concerns that may arise from our proposed architecture and argue that no intrinsic vulnerabilities are added as a result.

7.1 Introduction

It is not enough to simply open up spectrum for sharing: it must also be usable. Unnecessarily restrictive regulations (current and future) may inadvertently affect the design of the whitespace devices, thus artificially limiting their potential and blocking innovation. To that end, regulatory agencies have attempted to build a high degree of flexibility into their regulations.

Even as early as 2008 [1], the FCC recognized databases as an important method for discovering whitespaces[1]. It has since been shown that databases are actually the only certifiable method for discovering whitespaces [48], and the research efforts of the spectrum sharing community generally reflect that fact. However, the database method also requires some form of geolocation capability in order to accurately query the database. This means that devices lacking geolocation capability—either simple/cheap devices or GPS-reliant devices operating indoors—are left without any acceptable options.

Recognizing this limitation, the FCC (and Ofcom) introduced two modes of operation for devices: Mode I (client a.k.a. slave) and Mode II (master), saying "we believe that this approach will provide flexibility to permit a wide range of unlicensed broadband uses and applications while ensuring that the most appropriate and effective mechanisms are in place to protect TV and other licensed services" [1, ¶54].

This separation of responsibilities — geolocation and database access from actual operation — represents a significant step toward a more general architecture. The flexibility inherent in the existing regulations allows for a variety of devices that would have been completely out of the question with the naive approach; for example: low-power wireless sensor networks, all indoor devices, and widely-dispersed decentralized wireless control systems.

However, we argue in the rest of this chapter that the separation of responsibilities is not yet complete. In particular, we recommend that the essential components of whitespace systems be recognized as *individual* components which can be combined in a variety of ways. We do not propose to add new components to the architecture, merely to call out and separate the components which already exist in todays regulations. These components are:

1. The ability to determine the device's location

2. The ability to communicate with the database

3. The software-defined ability to assemble a database query and understand the response

---

[1]Using the database method, a secondary device intending to operate in the whitespaces must communicate its (approximate) location to a whitespace database (WSDB). The WSDB then calculates, using data and regulations provided by the regulator, the operating parameters which will be safe (from the incumbent's perspective) for the secondary device to use. These parameters, which include frequencies on which the device may transmit, are then communicated to the secondary, after which it may commence operation.

The critical observation is this: *these components need not be part of the same device, nor even part of a whitespace device*. We see several advantages to this proposal:

- Certification of individual components (and, separately, their interactions) is much simpler and more trustworthy than certifying an entire, complex device.

- Whitespace devices from all bands will have the opportunity to share a common support infrastructure in order to amortize costs and enhance scalability. This is consistent with the PCAST vision of widespread spectrum sharing[2]. In fact, our vision takes it a bit further: while PCAST envisions only a common database, we see a much richer common infrastructure emerging naturally [54].

Refactoring is a common software engineering practice in which old code is rewritten to improve design without changing external behavior. Code which is unnecessarily convoluted is often refactored into more general code which performs the same function. We see our proposal as a "refactoring of regulations" which reduces complexity while increasing performance and confidence. This refactoring does not change the protection guarantees offered to primary users or the political balance struck among non-technical social objectives.

The organization of this chapter is as follows. In Section 7.2, we describe the key components of whitespace devices and in Section 7.3 we give examples which motivate their separation. In Section 7.4, we connect this separation to the well-established software engineering principles of modularity, testability, and scalability. In Section 7.5, we discuss the fortuitous certification consequences of our proposed architecture. Section 7.6 briefly addresses the security concerns involved in distributed localization. Finally, we close by discussing some compelling byproducts of our proposed architecture, such as the economic and database-driven incentives.

7.2 Essential components of whitespace systems

A whitespace device (WSD) is a device that operates under particular time-space-frequency constraints designed to protect devices with higher priority. This definition is irrespective of the type of primary as well as the WSD's priority in the band. For example, in line with the PCAST vision, it applies equally well to a device operating with authorized shared access in a military radar band[3]. It is with this broad definition in mind that we proceed to identify the universal goals and components of whitespace devices.

---

[2]"This report argues that spectrum should be managed not by fragmenting it into ever more finely divided exclusive frequency assignments, but by specifying large frequency bands that can accommodate a wide variety of compatible uses and new technologies that are more efficient with larger blocks of spectrum." [54, page vii]

[3]The PCAST report suggests that there should be three tiers of spectrum access rather than the standard two that the FCC and Ofcom assume. In particular, it suggests the following tiers: "federal primary access," "secondary access," and "general authorized access" [54, §2.3].

7.2.1 Universal goals of whitespace devices

Throughout this chapter, we will assume that whitespace devices and their manufacturers have the following values:

- *Equipment should be (at least potentially) inexpensive.* In particular, regulations should avoid requiring that devices add components (cost) purely for the sake of regulatory compliance. If it becomes too costly to use the whitespaces, the ISM (i.e. internationally-available unlicensed) bands or even paid spectrum may be a better choice for many companies[4].

- *Whitespaces should be accessible/recoverable under reasonable conditions.* For example, devices should not be relegated to the ISM bands when indoors, nor should they by default lose out on many whitespace opportunities.

Our work strives to enable manufacturers to meet their goals in reasonable and flexible ways. We also aim to enable and promote application-agnostic regulations.

7.2.2 The whitespace access problem

Let's take a moment to think explicitly about the problem that needs to be solved by a device operating in the whitespaces using localization-based access. The primary goal of whitespace regulations is to avoid interfering with primary users and the secondary goal is to make as much dormant spectrum as possible available to secondary users. What capabilities do whitespace devices (WSDs) need to achieve these goals? Devices need to:

1. Gather location-relevant information ("localization")

2. Communicate location-relevant information to the whitespace database and to receive operating parameters in return (a "gateway")

3. Bridge the above services, resulting in a simple set of operating parameters ("packager") (this may not seem relevant right now but the importance will soon become apparent)

Under the current master-client paradigm, a device either has all of these capabilities (master) or it has none of them (client). Specifically, the master is assumed to have geolocation capability and access to a whitespace database (WSDB). The master also takes care of managing these tasks for the client, so that the client sees only a set of operating parameters and not all of the coordination required to get them.

We wish to consider the possibility that these responsibilities could be spread over multiple devices in a variety of configurations instead of using the restrictive master-or-slave dichotomy. We describe these components in detail in the following sections. For reference, Figure 7.1 lists the responsibilities and certification requirements for each component.

---

[4]An extra \$10/device will end up costing \$1B if 100M devices are sold.

Figure 7.1: Responsibilities of each component

7.2.3 Localization

The localization service provides information that can be used to approximate a client's location. We will elaborate on the idea of approximate location in Section 7.3.5, but localization equipment could take the following forms:

- A nearby whitespace device with its own geolocation capability (e.g. GPS).

- A transceiver on the roof of a building with a GPS unit installed[5]. This device could serve indoor master-type users as well as nearby client-type users.

- A beacon which uses time-of-flight measurements to determine the device's location relative to the beacon. (This would be used in conjunction with other localization equipment; see the example in Section 7.3.5.)

- A nearby smartphone running certified software (see Section 7.3.3).

- A nearby GPS navigator running certified software.

We wish to emphasize the fact that localization is a very generic service which is not specific to the whitespaces or even communications equipment.

---

[5]This is in some sense already allowed: "Personal/portable devices operating in Mode II [master] with radio-based geolocation capabilities in most cases will also need to be located outdoors to receive geolocation signals; we will allow these devices to receive such signals through a separate antenna that is located outdoors." [1, ¶91]

The important thing is that the protocol be able to provide the database with a set of *possible* locations (which we will call its "uncertainty region"). The protocol need only guarantee that the client is within that region.

The database would then return a list of the channels which are safe to use *everywhere* inside that uncertainty region[6]. It is not important that the uncertainty region be of any particular shape, or even that it be a single region. It is clear how the database should respond to a device which says "I'm either in New York City or San Francisco, but I'm not sure which." For the rest of the chapter, we will assume that it is sufficient for a device to provide an uncertainty region (or the raw data for computing one) to the database, and that the database can correctly respond to such requests. See Chapter 8 for more details on how the database could perform such a computation.

Clearly the number of available channels will decrease with a larger uncertainty region (see Chapter 8 for an in-depth discussion including plots of the tradeoff). Devices that cannot determine their location very precisely will have fewer channels available to them than those which are quite certain of their location. However, now a tradeoff curve can be established and manufacturers can choose their preferred operating point, rather than having it chosen for them by the regulator.

Notice that there is nothing inherently whitespace-y about a localization service. After all, location (relative or absolute) is a function of space, not frequency. This means that localization services could serve devices in a variety of bands and not just for whitespace access. Thus, as dynamic spectrum access becomes prevalent and more localization services are deployed, even old devices will enjoy an increased quality of service. So the amount of easily-recoverable whitespace increases with the need for whitespace.

For a lengthier discussion on a localization approach which uses sensing as a supplement, please refer to Chapter 8.

7.2.4 Gateway

A gateway should provide a way of securely communicating with the whitespace database (WSDB). It only needs to act as a relay, without performing any computations of its own. In fact, communications with the WSDB should be encrypted to prevent the gateway from tampering with the information therein. To reiterate, a gateway should act as a tunnel (in networking terms) between the WSDB and the packager.

Examples of potential gateways include:

- A nearby WSD with its own gateway capability.

- A smartphone running relay software (see Section 7.3.3).

---

[6]Under Ofcom rules, a safe power limit would also need to be computed and returned. As a first pass solution, the minimum power allowed in the uncertainty region would be safe.

- A WiFi router.

- A cell tower.

Although it may be built on complex lower-level protocols, a gateway actually provides only a very generic service: the relaying of (for example) IP packets. Database access through a gateway is a high-level protocol and there is no need to certify any of the lower-level protocols on which it is built.

7.2.5 Packager

The packager is a typically-software middleware that will sit between the client and the other two components (the gateway and the localization service(s)). It is responsible for translating requests, aggregating information, and providing a context for the information. Although the FCC and Ofcom mandate that the packager be part of the master[7], we expect that it will generally be found on the client because it is so lightweight.

The packager provides the following functionality:

- Receive information from localization services

- Create and sign a message including localization information and device ID

- Communicate with the client and gateway

Note that, for example, the packager is not *required* to interoperate with all localization services. Utilizing more services increases the quality of service for the client. However, in the worst case scenario when it cannot find or communicate with any localization services, it simply tells the client that there are no whitespace channels available at this time.

7.2.6 Interaction of components

To clarify our vision, let's take a moment to examine the interaction of the three components we highlighted. The interaction is shown in Figure 7.2. The steps, numbered identically in the figure, are as follows:

---

[7]According to the current regulations and consultations, the slave transmits its device ID to the master and receives a list of available channels in return. By definition the master is taking the role of packager in this case.
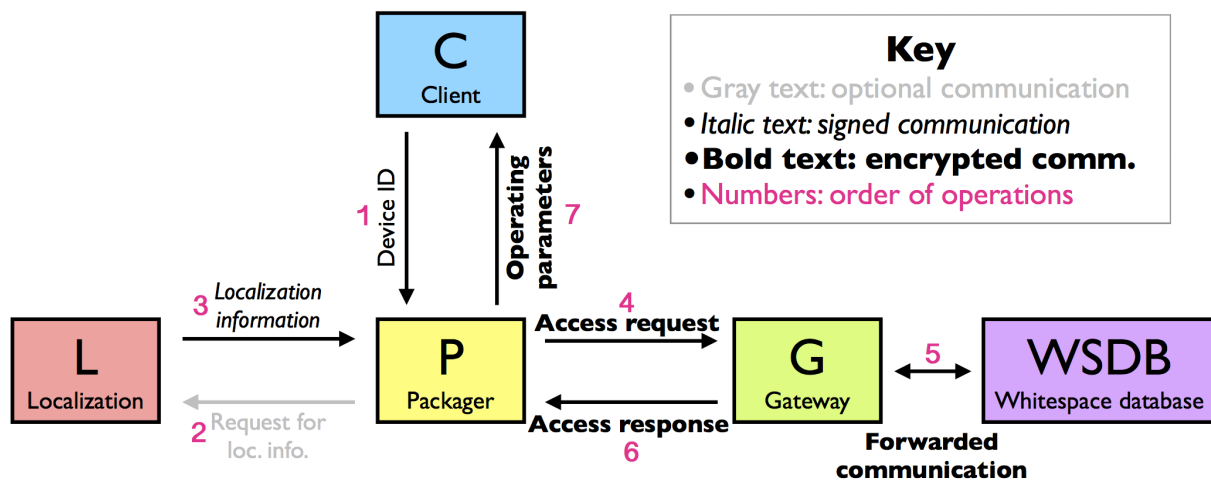
Figure 7.2: Interaction between components of our proposed whitespace architecture (see Sections 7.2.6 and 7.6 for details).

1. The client communicates its device ID (also known as the "FCC identifier") to the packager. The device ID is already required by the TV whitespace database access protocol [23, §15.711(b)(3)(iv)(A)] and we expect similar requirements in all shared-spectrum bands.

2. The packager optionally sends a request for localization information to a localization service. The necessity of this depends on the type of localization service used (e.g. broadcast vs. paired).

3. Localization information is digitally signed by the localization service and sent to the packager (see Section 7.6 for more details on this).

4. The packager assembles an access request consisting of the localization information and the client's device ID. This request is encrypted and sent to the gateway. Note that the packager may include multiple pieces of localization information.

5. The gateway forwards the encrypted access request on behalf of the packager. The whitespace database determines potential operating parameters for the client (e.g. a list of available channels), encrypts them, and sends them back to the gateway.

6. The gateway forwards the encrypted access response to the packager.

7. The packager forwards the encrypted access response to the client.

7.3 Examples of the separation of responsibility

So far, we have provided a description of the modular architecture that we are proposing. Under this proposed architecture, modules could be physically combined in a variety of ways, as shown in Figure 7.3. In this section, we motivate this modularity with a few examples.

(a) Master

(b) Master-client architecture as viewed by the client

(c) All separate devices

(d) Mostly separate

(e) Client and gateway on the same device, localization separate.

(f) Client and localization on the same device, gateway separate
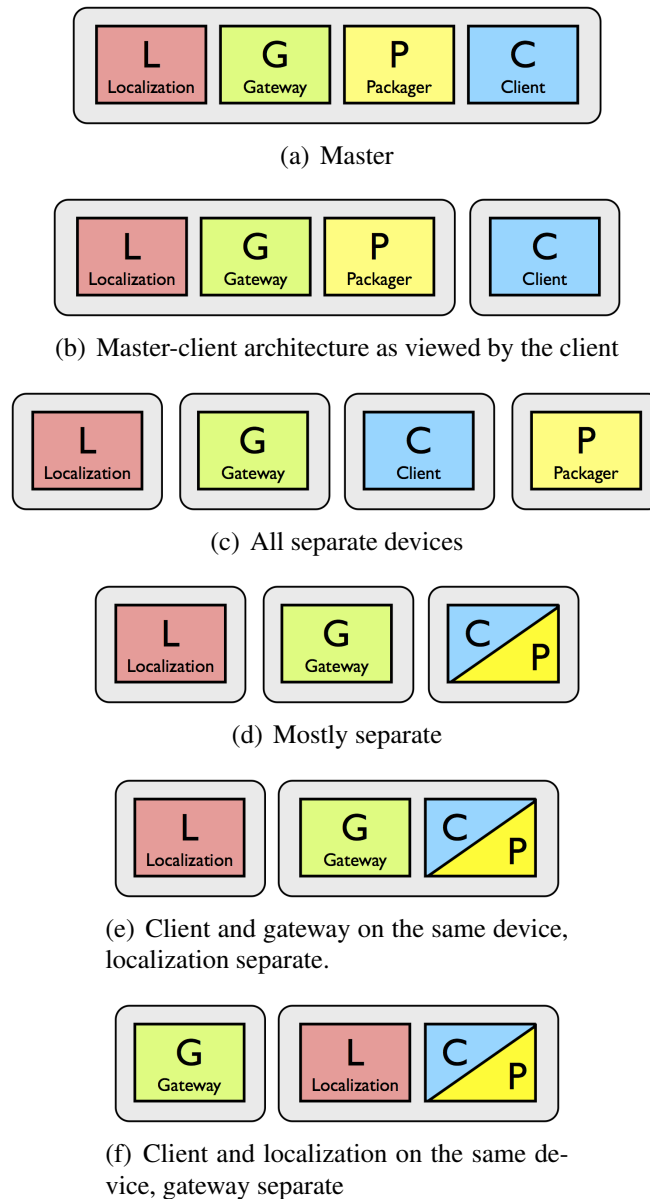
Figure 7.3: Different configurations for the components of a whitespace device support network. Colored boxes represent components, whereas gray boxes denote a device containing one or more of these components.

The standalone (a.k.a. Mode II, master) whitespace device that the FCC proposes is mandated to have all three capabilities listed above built-in. In the typical master-client paradigm, localization, the gateway, and the packager for the slave reside in the master device (also a whitespace device), as shown in Figure 7.3(b). However, this need not be the case: for instance, the client could contact a different device for localization,[8] and contact the whitespace database directly (shown in Figure 7.3(e)). It could also use another device to talk to the database for him (as shown in Figure 7.3(c)).

This section demonstrates the power of this proposed architecture via a series of real-world examples. Most examples will take the form shown in Figure 7.3(c) but the important point is that their form is unconstrained.

For the sake of brevity and clarity, we will assume that the client contains the packager in all but the first example below. We believe this will often be the case.

7.3.1   Example 1: the current master-client architecture

As shown in Figure 7.3(b), the current master-client architecture fits naturally within our new framework. The gateway, localization service, and packager are contained within the master device, while the client device is bare-bones.

Although this may seem like a silly example, we feel it is important to highlight that the hardware in existing and planned devices is already compliant with our proposed architecture. Moreover, the components of the master device are already highly likely to be separate; for example, manufacturers will probably use an off-the-shelf GPS module, etc.

7.3.2   Example 2: gateway for bootstrapping

Even the same physical equipment may take on several of these forms throughout its lifecycle. In the UK, slave WSDs may have geolocation capabilities but no gateway. The master continually broadcasts very restrictive[9] *generic* operating parameters to all slave devices in its service area. *If* these parameters allow for use of whitespaces, the slave may contact the master, providing its location in exchange for more permissive *specific* operating parameters. If not, there is no whitespace-based way for the devices to communicate.

In practice, it is difficult to obtain whitespace access via generic operating parameters alone due to their overly-restrictive nature. A proposed solution is the addition of a temporary gateway service (e.g. a nearby WiFi router or cell phone) for a geolocated slave device which allows it direct access to the specific operating parameters.

---

[8]Since this localization device is not communicating directly with the whitespace database, its location should be cryptographically signed to prevent tampering, forgery, or replay attacks. We will discuss this in greater detail later in Section 7.6.

[9] To protect the primary, the restrictiveness of the generic operating parameters increases with the service area of the master. This is because the regulations allow for operation only on channels which are available in the master's entire service area. The current form of the regulations means that often there is no whitespace available when using generic operating parameters.

7.3.3 Example 3A: smartphone as "master"

A key observation is that the equipment that provides these services doesn't need to receive or transmit on whitespace channels itself, and hence need not be (part of) a whitespace device. Of course, it must be able to talk to the client in some way but this is an existing requirement in the master-client paradigm and can be accomplished using a variety of standards such as WiFi and Bluetooth.

In this example, a smartphone running an regulator-certified application uses its built-in GPS to determine its own location[10]. It pairs with the client device using Bluetooth, which limits the distance between master and client significantly[11]. Thus the client's uncertainty region is roughly a 100-meter circle centered at the smartphone's GPS-determined location.

Now that the client has determined his uncertainty region, he can again contact the smartphone but this time use it as a gateway (again, through the certified application). Using the smartphone's data connection (WiFi or cellular), the client establishes secure communications with the WSDB and sends its uncertainty region along with other identifying information. In response, the WSDB securely replies to the client (through the smartphone) with a list of channels that are safe for the client to use.

The importance of this example relies upon the ubiquity of smartphones with data connections: most users already own a smartphone with a data plan. If the WSD is expected to be used by a person, it may be reasonable for a manufacturer to require customers to use a smartphone app with their device rather than spend the extra money required to give it its own geolocation and gateway capabilities.

Another critical feature of this example is that the smartphone is not itself a whitespace device. As mentioned previously, there is nothing whitespace-related about geolocation or data access, so the only requirement is that these services be able to communicate with whitespace devices. This means that supporting devices (such as smartphones) can be used with a variety of spectrum-sharing devices, not only those in the TV whitespaces. Thus the infrastructure costs can be amortized over many many devices instead of requiring a new infrastructure for each set of whitespaces. This idea of infrastructure reuse is hinted at in the PCAST report [54].

7.3.4 Example 3B: wireless sensor networks made more feasible by an smartphone

Imagine an ecology researcher who wishes to measure the rainfall in a particular area. She chooses to use devices which operate in the TV whitespaces due to their low-power-yet-long-range characteristics (a consequence of their low frequency). However, she requires inexpensive

---

[10]One could also consider allowing the smartphone to make use of other localization information, e.g. from WiFi fingerprinting or base-station triangulation [80].

[11]With many hacks, Bluetooth can transmit over distances of up to 2km. However, a more common operating distance would be maxed out at around 100 meters. http://en.wikipedia.org/wiki/Bluetooth

devices (i.e. client devices) because (1) funding is limited and (2) the risk that the devices will be stolen increases with their value.

With the existing master-client architecture, she must purchase a dedicated master device which will enable the cheaper client devices to access the spectrum. Under our proposed architecture, she could take her smartphone with her to the field to act as a "master" instead[12].

Moving the geolocation and gateway capabilities out of a dedicated whitespace device and into the already-purchased cell phone decreases costs enormously. This is especially true if the master device would have otherwise been purpose-built for whitespace sensor networks—or worse, specific sensors.

7.3.5 Example 4A: city-wide localization equipment

A city, wishing to provide services for its citizens and businesses alike, installs a beacon in a prominent location[13]. This beacon is a fixed device, its location is certified by a technician, and it regularly broadcasts its location[14]. This transmission may occur on any band, e.g. an open whitespace channel, an ISM band, cellular band, or even over FM radio. Any whitespace device that can hear the beacon is able to use its localization information to help determine its uncertainty region.

There are two important concepts in this example:

- The beacon may actually be a very cheap and simple device: it need not include a GPS unit nor does it even need an Internet connection. In principle, this beacon may be a simple box with an antenna which requires only (1) location-certified installation and (2) a power source.

- The load on the beacon is constant, regardless of the number of whitespace devices which benefit from its transmission. This is clearly preferable to the O(N) scaling demanded by the existing master-client architecture.

The idea to use beacons to assist with whitespace access is not new. In fact, the FCC considered using beacons in a slightly different way in the 2008 regulations:

---

[12]Note that the current FCC regulations require frequent contact with the database. Depending on the circumstances, one can imagine citizen involvement as a potential solution to this problem [81].

[13]It is not only public agencies who have motivation to provide such a service. Indeed, a beacon could be ad-supported. For example, the packager could require the user to view an ad in order to understand (decrypt, perhaps) the localization information.

[14]We will discuss the related security concerns in a later section but for now assume that the beacon's information is cryptographically-signed.

> A second method would be for an unlicensed device to receive information transmitted on a control signal by an external source, such as a broadcast station, CMRS base station, or another unlicensed transmitter, indicating which channels are available at its geographic location. Under the control signal method, **a device would be allowed to transmit only after it receives a signal with information that positively identifies which TV channels are available for use**. Control signals could be transmitted by any number of sources, such as broadcast TV and radio stations and/or licensed wireless communications providers. [1, ¶56] *(emphasis added)*

The FCC quickly abandoned this proposal due to a variety of reasons: (1) lack of interest by stakeholders, (2) the perception that this would require expensive infrastructure, (3) the "inherent conflict of interest for TV licensees" [1, ¶70].

However, our proposal is different in several key ways:

1. No special regulations would need to be written to allow the use of beacons: their functionality and corresponding certification process would already exist as generic "localization services."

2. The equipment need not be in the purview of traditional stakeholders (TV station operators).

3. The equipment would be certified to perform only specific and simple tasks, rendering tampering more difficult.

4. Our proposal does not require the beacon to have a connection to the whitespace database[15]. This substantially decreases the complexity and the cost of the beacon.

5. Our proposal, which uses simple equipment in the field and defers actual calculations to the WSDB, is more trustworthy. Rather than relying on the beacons to determine which channels are available for secondary use, the decision is made at the database. Thus when an update is required (e.g. if a bug is found) (1) only the WSDBs (rather than all of the beacons) will require updates, and (2) regulators can assume that the updates will be applied within a reasonable amount of time.

---

[15]Although not explicitly stated, it is implicit in the quotation above that the beacon would need a connection to the whitespace database. Without this connection, it would be unable to correctly determine which channels are *currently* available. Manual updates, while technically possible, would be inherently unreliable and out of date while imposing additional maintenance constraints on the owner of the beacon.

7.3.6 Example 4B: city-wide localization equipment combined with building-level localization equipment

This example builds on the previous example in that it assumes the existence of a city-wide localization beacon. However, here we also consider a building manager who has decided to enhance the whitespace access capabilities of devices inside his building by placing a second type of localization equipment on the roof of his building.

This second piece of equipment has two advantages:

- In the event that the city beacon's signal is too attenuated indoors, it can act as a secure relay[16].

- If the equipment has additional sensing capabilities, it may be able to use techniques such as location fingerprinting (e.g. with WiFi or with TV signals themselves—see Chapter 8 for more details on how this could be done) to narrow down the uncertainty region.

This additional equipment may be a portable device which only needs access to electricity (perhaps via solar cells). Its location will not be certified which saves the building manager money by avoiding professional installation. The installation would ideally be as difficult as installing a clock.
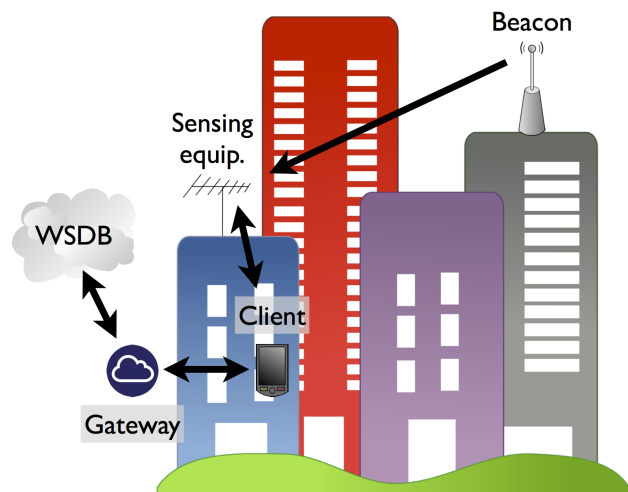
Figure 7.4: Potential whitespace support network for indoor client

This scenario is illustrated in Figure 7.4. Clients within the building communicate with the sensing equipment. Time-of-flight information may be used by the building's equipment to certify that

---

[16]We address the security concerns of this in a later section.

the client is within a particular small distance of the building's equipment. The clients then communicate the following information to the whitespace database through a gateway (which may or may not be part of the client):

- The beacon's location information collected by the building's equipment.

- Certification from the building's equipment that the client is within some range R_2.

- (Optional) Additional localization information (e.g. TV signal levels) collected by the sensing equipment.

In this scenario, the WSDB would calculate the potential locations of the sensing equipment based on the beacon's location. In Figure 7.5, the beacon is at known location L_1 and the green dotted circle represents the possible locations of the building's equipment. The database then adds a buffer of size R_2 to the green circle to account for the client's uncertainty relative to the building's equipment. The construction is shown via the light blue solid-line circles (dark blue, L_2, indicates the true location of the building's equipment). A list of channels safe for use in the final uncertainty region (shown in hatched purple) is computed by the WSDB and returned to the client.
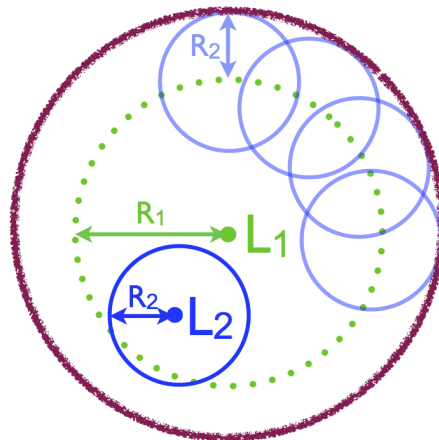
Figure 7.5: Illustration of a safe way to chain localization information. L_1 knows its absolute position and its service radius, R_1. L_2 (with service radius R_2) is within L_1's service area but it knows no more than that; for illustrative purposes, we have shown an example location for L_2. In order to be conservative, the whitespace database thus draws a circle (shown in purple) of radius $R_1 + R_2$ around L_1's location. This circle is the union of all possible service areas for L_2 (a few of which are shown in faded blue). In this way the WSDB is certain that a device within range of L_2 will be within this larger circle.

This example shows the potential for full separation of roles in order to inexpensively solve the problem of indoor whitespace access. A single device may serve multiple roles (e.g. the client could contain the gateway) but it is not necessary to do so. Breaking out these roles grants innovators greater flexibility in the design of whitespace networks.
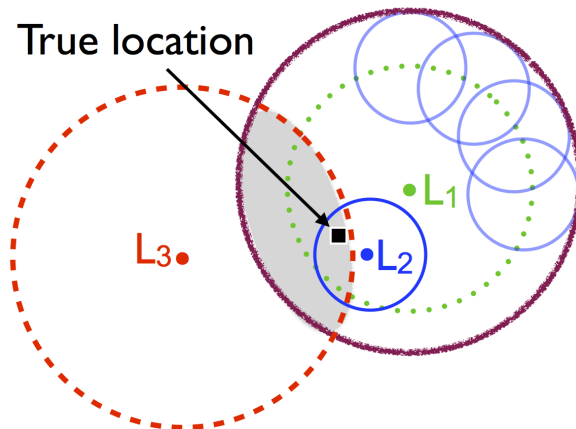
Figure 7.6: Synthesis of localization information based on the example in Figure 7.5. The addition of L_3 reduces the uncertainty region (shown in gray) from the entire interior of the purple circle to only its intersection with L_3's circle.

7.3.7 Example 4C: enhanced localization

Suppose the client is able to receive localization information from a third source (e.g. a beacon whose signal is strong enough to be heard indoors). Now the client must be somewhere in the intersection of the previous purple uncertainty region and the service area for the third source. This intersection is shown in gray in Figure 7.6. This example demonstrates that additional localization information can substantially reduce the size of the uncertainty region.

We generally understand how such systems work since we use similar procedures for determining the location of cell phones (e.g. for the E-911 service). Thus we can let innovators reuse some of the work from related fields to robustly solve this problem.

7.4 Connection to software engineering principles

In this section, we will show how our proposed architecture satisfies the modern software engineering principles of modularity, testability, scalability, and upgradability. Each of these principles is well-recognized not only within the field of software engineering but also in many other engineering fields, especially design for manufacturing and assembly.

7.4.1 Modularity

Modularity is a design paradigm which stresses the separation of responsibilities among components of a system. A modular system is comprised of components, each of which has a specific task and a well-defined interface to the rest of the system. It is possible, in principle, to replace each component without needing to redesign the rest of the system.

Within the field of computer science, the benefits of modularity are widely known. As Liskov says in her book:

> Modularity is the key to writing good programs. It is essential to break up a program into small modules, each of which interacts with the others through a narrow, well-defined interface. With modularity, an error in one part of a program can be corrected without having to consider all the rest of the code, and a part of the program can be understood without having to understand the entire thing. Without modularity, a program is a large collection of intricately interrelated parts. It is difficult to comprehend and modify such a program, and also difficult to get it to work correctly. [82]

The advantages of modular design have been long recognized in fields outside of computer science and electrical engineering. For example, the use of interchangeable parts was critical to the efficiency and success of assembly lines and other mass-manufacturing processes. The flexibility of being able to seamlessly substitute a new part for a broken one greatly decreased costs since now items could be repaired instead of replaced. Modular building designs have recently gained popularity due to their reduced costs and ability to be mostly-constructed off-site.

In the case of whitespace devices, modularity not only helps with testing (as argued in the next section) but also with giving devices greater flexibility. By not rigidly defining the location or owner of the modules, many more designs are possible.

7.4.2 Testability

One of the great benefits of modular design is testability. While it can be difficult to certify that an entire system is working correctly due to its inherent complexity, it is often easier to certify each of its basic components individually. After all, the components are by definition simple pieces with well-defined behavior.

For example, one wouldn't dream of certifying a whitespace device and a whitespace database together as one unit, even though they will operate together in the field. Instead, regulators should test each individually to certify that it is operating with the defined parameters[17].

Using modularity to improve testability is a common software engineering practice. Liskov's book separates testing into two components:

---

[17]However, the current FCC certification process [83] does not appear to be as rigorous as one might hope. For example, the device's geolocation capabilities are not directly measured in the current testing procedure, nor is it clear how the FCC will test that the device "cannot be made to operate on an unauthorized channel" [83, part 2, §2(j)]. In fact, much of the testing procedure appears to test the WSDB as much as the WSD itself. For example: "After receiving an available channel list, register a low-power auxiliary device on the TVBD operating channel... and confirm that the low-power device is accounted for in the schedule."

> Testing typically occurs in two phases. During *unit testing*, we attempt to convince ourselves that each individual module functions properly in isolation. During *integration testing*, we attempt to convince ourselves that when all the modules are put together, the entire program functions properly. [82]

What we are proposing is to apply this methodology to whitespace certification. Specifically, we propose that the components we identified be subject to unit tests. Manufacturers will naturally perform their own integration testing. Regulators can (and should) do integration testing for their regulations, as discussed in Section 7.5.5, but this is separate from the component certification burden imposed on companies.

One important consequence of performing unit tests rather than integration tests (as is currently the standard) is that now individual components can be patched or upgraded without requiring recertification of the entire system.

7.4.3 Scalability

As the number of networked devices increases at a dizzying rate, there is increasing emphasis on scalability in software design. Scalability refers to the ability of a system to handle additional requests (clients) without excessive overhead.

Computer scientists often refer to an algorithm's scalability in terms of the number of computations it performs based on the size of the input, N. O(N) means that each additional input adds a fixed cost whereas O(1) scaling indicates that the algorithm's completion time is independent of the input size.

In Ofcom's and the FCC's current master-client paradigm, the overhead for a master is typically O(N) where N is the number of clients it serves[18]. In our proposed system, the gateway will generally incur a cost which scales like O(N). However, the localization equipment can operate with a fixed cost regardless of the number of users via broadcasting its information.

The paradigmatic master's load scales poorly with the number of clients, incentivizing master devices to conserve their resources only for clients within their system. Reducing this burden means that manufacturers will be much more likely to build interoperable systems with reciprocal infrastructure-sharing agreements. In our proposed architecture, the localization equipment can operate with a fixed cost regardless of the number of users via broadcasting its information.

Looking forward, we and the PCAST envision a future in which the TV whitespaces are only one set among many whitespaces [54]. In this future, localization services and gateways could serve the same purpose for all types of whitespace devices with no modifications. As whitespace devices increase in popularity, the number of localization services is likely to increase. This will generally mean a reduction in the size of uncertainty regions, therefore increasing the amount of recoverable whitespace.

---

[18]The exception is Ofcom's generic operating parameters which are O(1) for the master.

7.4.4 Upgradability

In some sense, the current architecture is already remarkably modifiable, especially in contrast with previous spectrum allocation paradigms. For example, the following values specified in the regulations can be updated: coverage areas of primaries, separation distances, and power limits. Notice that all of these items are upgradable because the necessary data and computations reside *in the whitespace database*. This demonstrates the end-to-end principle[19] applied to dynamic spectrum access. The power of this model is shown by the rising popularity of software-as-a-service which enjoys very fast upgrade cycles since the software is hosted in the cloud.

Even in situations where providing a master device for the clients is not out of the question, clients will only see a performance boost when their specific master is upgraded. Under our proposed architecture, clients can more easily take advantage of upgraded hardware because of the greater likelihood for interoperability. Additionally, rather than needing to upgrade an entire master device, modules can be upgraded individually which makes incremental upgrades much more feasible.

7.5 Consequences for certification

One of the goals of our proposed architecture is to simplify and robust-ify the certification process for whitespace services. Due to the modularity of the system, components can be certified individually rather than as a combined product. This allows for simpler per-module tests which are easier to carry out and inspire more confidence.

We believe that every component we have identified as requiring certification also requires certification under the existing master-client architecture.

Streamlining and standardizing the certification process will decrease the cost (in dollars and time) of certification. This is much more attractive to manufacturers who may otherwise fear uncertain and lengthy procedures.

Perhaps most importantly, the certification of all three identified components—the localization services, gateway, and packager—is no longer a matter of interference. Naturally, all three must work in concert to ensure that the client does not operate illegally. However, the nature of the services they provide has nothing to do with interference. This makes the entire process less susceptible to concerns revolving around the vague term "harmful interference."

We will now describe the pieces of each component that need to be certified. Throughout the process, we will make references to the content of Figure 7.1.

---

[19]"The end-to-end principle states that application-specific functions ought to reside in the end hosts of a network rather than in intermediary nodes, provided they can be implemented 'completely and correctly' in the end hosts." — http://en.wikipedia.org/wiki/End-to-end_principle. This principle was first introduced in [84].

7.5.1 Certifying localization services

As mentioned previously and in Figure 7.1, the localization service has the following responsibilities that must be certified:

- *Determine absolute location (if applicable).* Existing devices such as GPS navigators are already certified (e.g. by the FAA for aircraft navigation) so a similar certification process can be developed. Furthermore, absolute localization is already required by the FCC for whitespace devices so no additional challenge is presented.

- *Determine bounds on distance to another device for chaining of localization information (if applicable).* This would be needed in systems such as the one in Example 4B. Again there are existing systems (e.g. the E-911 system) which provide such distance estimates.

It is evident many of the pieces in the localization service already have a candidate certification process. Those which do not are still easy to test. It would even be possible to create open-source government-blessed software packages which perform these tasks. Since the problem is no longer at all related to "harmful interference," the standard for such software could be reasonably outsourced to an organization such as NIST (National Institute of Standards and Technology).

7.5.2 Not certifying the gateway

The worst thing a gateway can do is fail to relay data. We assume that the traffic being routed through the gateway is appropriately encrypted, which prevents eavesdropping and tampering (more details in Section 7.6). Furthermore, we suggest a scheme in Section 7.6 that protects against replay attacks, man-in-the-middle attacks, etc.

Communications between the packager and the WSDB should be encrypted with the database's public key (thus the data will only be meaningful to the database the client intends to use). Replies to the packager should be encrypted with the client's public key (thus rerouting to another client would be useless as it would not be able to decrypt the data).

The bottom line is this: *regulatory bodies need not trust the gateway in any way.*

7.5.3 Certifying the packager

As with the localization service, the packager needs to be certified to create and encrypt messages containing the correct information (localization information, device ID). This can be tested in a similar manner via software unit tests.

7.5.4 Certifying the client

The pieces of the client requiring certification are already supposed to be certified by regulators under the existing architecture. For example, the client must already be certified to provide the correct device ID and to operate according to specified parameters. Thus no changes are needed here.

7.5.5 Not certifying the system

Certifying that each piece of the system operates correctly frees the regulator to think separately about how all of the pieces will interact. Regulators can create models and simulations of these components to examine their interactions in order to prevent any unfortunate combinations *before* they show up in an actual system. Using simulations allows regulators to carry out much more thorough and reassuring studies than could ever be done on a per-device or per-system basis.

Even in the event of a failure, the regulator maintains a high degree of control over the system via the databases. In most cases, we expect that system-wide flaws (e.g. an unforeseen interaction of components) would require only an update to the WSDB code. In the worst case, devices or components can simply be denied access to the whitespaces, either temporarily or permanently. This ability to control even deployed systems inspires confidence that WSDs will operate correctly.

7.5.6 Avoiding unnecessarily complicated regulations

In the quotation below, we see a specific unfortunate consequence of the structure of the existing FCC regulations. Here, a special allowance has been made for fixed devices which do not have a data connection.

> If a fixed TVBD does not have a direct connection to the Internet and has not yet been initialized and registered with the TV bands database consistent with §15.713, but can receive the transmissions of another fixed TVBD, the fixed TVBD needing initialization may transmit to that other fixed TVBD on either a channel that the other TVBD has transmitted on or on a channel which the other TVBD indicates is available for use to access the database to register its location and receive a list of channels that are available for it to use. Subsequently, the newly registered TVBD must only use the television channels that the database indicates are available for it to use. A fixed device may not obtain lists of available channels from another fixed device as provided by a TV bands database for such other device, i.e., a fixed device may not simply operate on the list of available channels provided by a TV bands database for another fixed device with which it communicates but must contact a database to obtain a list of available channels on which it may operate. [9, §15.711(e)]

Using a modular architecture like the one we propose would avoid such complicated exceptions by providing this flexibility by default.

7.6   Security concerns

The FCC currently requires that all communications between the slave, master, and WSDB be secured [9, §15.711(f)(ii)-(iii)]. Ofcom has a similar requirement [85, ¶5.68-5.70]. Naturally, these requirements should extend to communication between modules as well.

In Figure 7.2, we have elected to recommend that most of the inter-component communication be either signed or encrypted. Now we will briefly expand our description of these interactions, now focusing on the security aspects. As before, the numbers in the list below correspond to those in the figure.

1. The device ID may be transmitted in the clear[20]. Because the response from the database will be encrypted with the public key associated with the provided device ID, tampering with it would simply make it impossible to decode the resulting operating parameters.

2. This request can be sent in the clear as it is only a "please talk to me" request.

3. This information needs to be signed by the localization service using a service-specific key for verification purposes. Furthermore, the messages should contain the following information:

   - Unique localization service identifier (to help identify compromised or broken equipment)
   - Timestamp (to help defend against replay attacks)
   - Localization information (e.g. GPS coordinates)
   - (Optional) Transmission power (used by the WSDB to estimate the coverage area)

4. This should be encrypted with the WSDB's public key to prevent tampering.

5. Since information is only being relayed, no extra logic is required here.

6. The access response should be encrypted with a key generated from the device ID. It will contain a timestamp, a list of available channels and (if applicable) corresponding emissions limits.

7. The packager simply forwards the access response to the client. The client uses its hardcoded device ID to decrypt the access response and extract its operating parameters.

Although the interfaces between components will be more exposed in our proposed architecture than in the current one, we believe that appropriate security measures can reduce the associated risks.

---

[20]Encryption may be useful to guard against an eavesdropping packager.

7.7 Consequences

We have identified the essential components of a whitespace device support network:

1. Database gateway: this equipment is responsible for contacting the database on behalf of the client.

2. Localization service: this may take many forms, including a fixed beacon, a beacon with GPS capabilities, or a beacon plus sensing equipment.

3. Client packager: this service is responsible for aggregating, coordinating, and disseminating information on behalf of the client.

Separating out these components has some consequences which were not addressed above. We now briefly address them.

7.7.1 Economic incentives

With this new perspective on the minimal requirements for the whitespace access problem, we see that *existing* devices such as smartphones are only a software upgrade away from being whitespace-enabling devices. By building on top of existing technology, our proposed architecture enables accelerated growth within the WSD market.

Furthermore, the burden for whitespace access can be creatively distributed as makes sense for the problem at hand.

All of this means that client devices—which are much cheaper than the paradigmatic master devices and have the added benefit of actually working indoors—are now truly within reach. The bottom line is this: *there is an economic incentive to adopting this potential separation of responsibility.*

7.7.2 Lighter regulations

Although our proposal has major beneficial consequences for devices and device manufacturers, the regulatory changes needed are actually minimal. Changes to existing devices are also minimal since, as mentioned earlier, even the current master-slave setup fits within our proposed architecture.

With the exception of the chaining of localization information—which is optional and need not be implemented in the first round of regulations—our proposal essentially requires more general wording within the regulations but no new ideas. This general wording pays off quickly when one considers the reduction in special-case regulations, such as the one mentioned in Section 7.5.6. We wish to emphasize that our proposal is actually for rules which are *lighter* than the existing regulations, not more burdensome.

7.7.3 Added value in whitespace databases

Databases could potentially compete on how much and what kinds of localization information they can safely fuse.

7.7.4 Market penetration

As the number of whitespace-related devices increases, the probability of being unable to access any particular component of a whitespace device support network decreases. Thus as the market penetration increases, devices will be able to more precisely locate themselves and thus recover more whitespaces. Until infrastructure is widely available, devices may use cheap techniques with poor-but-usable service and plan to opportunistically use current and future infrastructure.

7.7.5 Enabling the one-more-band model

A very real use case is the one in which an existing wireless device wishes to expand its capabilities slightly using whitespaces. Enabling cheaper devices makes this a possibility, as we will see in the brief example below.

Consider the case of a cellular provider whose towers are saturated due to the ever-increasing use of mobile data by consumers. A popular strategy is to use WiFi offloading to reduce the load on the cellular towers. At the same time, the TV whitespaces are being heralded as "Super WiFi" because of their potential for long-range wireless data links. A natural fit would be to use a whitespace device to handle traffic on behalf of the cellular tower.

However, such a device would need to determine its (approximate) location in order to access the whitespaces. A natural fit is the cellular tower: the cell tower is professionally installed and its location has already been determined for other purposes. Furthermore, our proposed device would already be communicating with the cell tower in the cellular bands in order to negotiate the handoff of mobile devices.

Thus the cell tower can easily provide localization and gateway services for the whitespace device. In exchange, the whitespace device handles data requests on behalf of the cell tower once it has established a connection to its backhaul provider. At this point, the whitespace device can provide its own gateway services. In this way, the cell tower receives large returns on a small investment in the whitespace device.

7.8 Conclusions

In this chapter we have proposed a modular approach to the design and regulation of spectrum-sharing devices. At a high level, we have borrowed principles from the established area of "design for manufacturing and assembly" which seeks to reduce product lifecycle costs by considering manufacturing constraints in the design phase. In general, as radio regulations have an increasing

technical/architectural footprint, this same mindset needs to be brought into the design of regulations.

We identified and described the three key components of WSDs (the localization, gateway, and packager services). We then provided example configurations of these modules to motivate their separation. One of our key observations was that although the components identified above need to be certified for compliance with whitespace rules, *they do not need to be part of a whitespace device*. For example, a smartphone could easily and reasonably provide all three services to a client whitespace device via a regulator-certified app. This reduces manufacturing and deployment costs while increasing flexibility for users and designers of whitespace devices.

Next, we argued that certification of these components, for the most part, *should already be part of the certification process of a whitespace device*. Unit and integration testing (as opposed to integration testing alone) has long been recognized as a critical piece of reliable testing and certification.

We briefly discussed the security implications of our proposed approximate-localization architecture and compared it to existing solutions. We concluded that our proposal introduces no new vulnerabilities.

Finally, we remarked on some consequences of our proposal. In particular, we noted that our proposal promotes lighter and less-complicated regulations while also promoting diversity in whitespace device and system design.

7.9 Future work

We conclude with a few notes on some remaining work. These concerns were uncovered as a result of this refactoring exercise, but they apply even to the current approaches.

7.9.1 Security concerns

The protocol we proposed is a first attempt at designing such a system. History has shown very clearly that even seemingly simple and secure systems can have unanticipated flaws. Before such a system is adopted, this problem should be thoroughly studied by parties more qualified than we are.

7.9.2 Certification procedures

Formal certification procedures need to be proposed for each of the identified components. Furthermore, a testbed for the interaction of modules needs to be created so that we can study the behavior of the components as an entire system.

7.9.3   A chicken-and-egg problem: how does the client contact the packager?

Under the current FCC regulations, a slave device may contact the master using "an available channel used by the [master] or on a channel the [master] indicates is available for use by a [slave]" [23, §15.711(b)(3)(iv)(D)]. Ofcom proposes a similar approach in which the master broadcasts "generic operational parameters" to all nearby slaves [85, ¶5.40].

We believe that an Ofcom-like approach of broadcasting conservative operating parameters would be safe. However, this is actually another (virtual) component in the whitespace device support network whose role and operation needs further study. In reality, we expect that this component would exist on the same device as at least one of the other components (most likely with the packager or gateway).

7.9.4   Caching ability

In both the FCC and Ofcom regulations, the master is allowed to cache operating parameters for the slave (in a sense). The communication load on the master on behalf of the slave is therefore greatly reduced.

FCC: Currently, the master's communication with the WSDB on behalf of the slave is limited to verifying the slave's device ID. Once complete, "a [master] must provide a list of channels to the [slave] that is the same as the list of channels available to the [master]" [23, §15.711 (b)(3)(iv)(B)]. This reduces the per-slave overhead of the master.

Ofcom: Similarly, the master receives "generic operational parameters" from the WSDB which are used (initially[21]) for all slaves [85, ¶5.38].

Under our proposed architecture, one has to wonder where such caching should occur. For example, it could happen at the packager or even an WSDB proxy but work is required to determine how to reliably certify this functionality. Our discussion above and Chapter 8 reveal that the FCC's existing approach to the caching problem is fraught with problems.

---

[21]The slave also has the option of requesting "specific operational parameters" after its initial association with the master, which increases the burden on the master and presumably prevents caching.

Chapter 8

Solving the weakly-localized device problem

Some or all of the work in this chapter appeared in [86].

In this chapter, we consider the problem of granting whitespace access to devices which have neither certified sensing capabilities nor a means of direct geolocation. These devices, called "slaves," use a nearby "master" device to assist in determining which channels are available for secondary use. Such devices must be supported since even a "master" device which uses GPS for geolocation will need to become a "slave" when operating indoors where GPS operation is notoriously poor.

The two regulatory bodies that are most active in this space, the Federal Communications Commission (FCC) in the United States and Ofcom in the United Kingdom, have similar yet slightly different approaches to the problem. While in the US the slave is directed to use the channels which are available at the *master's* location, slaves in the UK are given operating parameters which should be reasonably safe anywhere within the master's coverage area. We demonstrate in this chapter that the first approach is too permissive while the latter is too conservative.

Ultimately, we believe that the problems with these approaches are due to the misconception that whitespace devices need to determine their locations. In truth, the *actual* goal is to determine a set of channels on which it is safe for the device to transmit. For example, it is clear how a whitespace database (WSDB) should respond to a weakly-localized device which can reliably say "I am located either in New York City or San Francisco, but I don't know which." In this case, the WSDB should compute the set of channels which are simultaneously available for use in both NYC and SF, then direct the device to choose from that set.

To demonstrate the power of this shift in perspective, we propose an enhancement to Ofcom's necessarily-conservative approach which *safely* increases the number of whitespace channels available to slave devices via simply sensing for *strong* signals. However, our actual goal is much larger than a particular method. As discussed in Chapter 7, we believe that regulators should certify "localization" in a broader sense of the term, and this work simply serves as a proof-of-concept/need for that argument.

8.1 Introduction

8.1.1 Realistic whitespace devices: master and slave

The unlicensed use of the whitespaces enables new niches in which wireless products can be created without having to pay exorbitant and prohibitive amounts of money for spectrum in which to operate. This makes pay-once products possible (as opposed to subscription-based services like cell phones).

Beyond opening up spectrum for use by innovative companies, it is important to enable them to make devices at a reasonable price[1]. The standard method for accessing whitespace spectrum is to use GPS to determine the device's location and subsequently contact a WSDB for operating parameters which depend on that location. However, adding GPS capabilities to a device is expensive in both dollars and power.

Moreover, unassisted GPS performs notoriously poorly indoors to the point of being useless. The FCC, recognizing these limitations, created two classes of devices: master and slave[2]. Master devices have actively functioning geolocation capabilities (i.e. they will know where they are to within 50 meters [23, ¶48]) but slave devices need not. The slave communicates its device ID to the master who subsequently contacts the database on behalf of the slave. The database then returns the list of channels available *at the master* (recall that the slave's location is unknown but is assumed to be near the master). Thus the slave receives a list of channels on which it is allowed to transmit and commences operation.

Ofcom allows for a similar architecture, but has opted for an additional safely feature. Under the proposed Ofcom regulations, the coverage area of the master device is estimated and only general operating parameters which are safe for the entire area are returned[3] [87, ¶5.38.2].

8.1.2 Overview of chapter

In Section 8.2, we describe the FCC's approach to the master-slave architecture in greater detail. In doing so, we also identify a potential flaw in this approach: the distance between the master and slave devices is not bounded nor accounted for in the operating parameters. We describe scenarios in which this leads to the slave operating improperly and we quantify the probability of this occurrence using real-world data.

---

[1]Even when spectrum use is on a licensed basis, cost is important. Any unnecessary cost burden on individual devices trying to use spectrum functions as either a tax on the user or as revenue that will not be obtained by the government in an auction (e.g. for secondary rights). The cost need not be dollars but could also be battery life.

[2]The master is called Mode II and the slave is called Mode I in the FCC rules. However, we feel that the master-slave nomenclature is clearer and therefore we use that convention.

[3] After receiving generic operating parameters from the master, a slave device may also request that the master provide it with specific operating parameters, which are typically more permissive since the database now has additional information about the specific slave device. In this case, the problem is the same as in the US.

In Section 8.3, we discuss Ofcom's solution to this problem. Briefly, the database estimates the potential locations of the slave device and allows it to operate only on channels which are available at all potential locations. Although safe for primaries (incumbents), this approach is necessarily conservative, as described in Section 8.4.

Section 8.5 proposes an enhancement to Ofcom's approach which drastically improves the quality of service for slave devices while preserving the quality of service of primary systems and preserving the ease of implementation for whitespace devices. This approach, a variant of location fingerprinting, is formally evaluated in Section 8.6.

Implementation details and data sources are given in Section 8.10. Source code for all simulations is at [6].

8.1.3 Related work

While there are many papers on sensing [88], cooperative sensing, location fingerprinting and other related topics, there does not appear to be any work which seriously looks at the master-slave issues we have identified nor any which addresses the inherent location uncertainty of the slave. Thus the work done by the FCC and Ofcom constitutes the majority of the prior work on this topic.

For example, [89] use a more restrictive form of location fingerprinting which includes a beacon in the primary's signal, requiring coordination and expenditure on the part of the primary. This lack of incentive alignment caused the FCC to drop the idea of beacons entirely [1, ¶70]. Moreover, the authors of [89] were primarily concerned with localizing the secondary device rather than accepting and addressing the location uncertainty.

A similar methodology which used TV signals to carry information about nearby used channels is considered in [90], but the authors missed the crucial step of using communicating with the whitespace database to reduce uncertainty.

8.2 The problem with the FCC's approach

Under the current FCC rules, a slave device employs a master device to help it access the whitespaces. In particular, the slave communicates its device ID to the master who relays this to the whitespace database (WSDB) along with the master's location. After verifying that the device is allowed to operate (based on its device ID), the WSDB signals that the master may share its list of available channels with the slave.

The assumption that is implicit in this mechanism is that *the slave is close enough to the master that the list of available channels has not changed.* However, a simple calculation[4] shows that even a pair of 40-mW transmitters can hypothetically communicate at more than 1 kbps at a range of 100km. While this is not an acceptable data rate for most purposes, it *is* enough to obtain the information which allows the device to access the whitespaces since the amount of master-slave communication needed for this is quite minimal[5].

The FCC makes this assumption more explicit in its 2012 comments on which fixed devices may act as master devices:

> **We are prohibiting fixed devices with an HAAT greater than the current maximum of 106 meters from providing channel lists to Mode I personal/portable [slave] devices.** This action is necessary because a Mode I device, which does not incorporate a geo-location capability, obtains a list of available channels from a fixed or Mode II device that is determined by the geographic coordinates of those devices. Under the current 106 meter limitation, **the communication distance between a Mode I device and the fixed or Mode II device that provides a channel list is relatively short**, and thus there is a low probability that a Mode I device will operate at a location where its channel list is not valid, i.e., does not meet the minimum separation distances from co-channel and adjacent channels TV stations or other protected services. However, **if the fixed device that obtains the channel list for a Mode I device operates with greater HAAT than the current rules permit, the Mode I device could operate at a greater distance from the coordinates of the fixed device where the available channel list was calculated. This will increase the chance that the Mode I device could operate at a location where the channel list is not valid.** We will therefore require that the TV bands database not provide channel lists for Mode I devices through fixed devices with an antenna HAAT of greater than 106 meters. [23, ¶19] *(emphasis added)*

---

[4]Here we use the theoretical Shannon capacity of a channel, pathloss from the ITU propagation model, the bandwidth of a TV whitespace channel in the United States (6 MHz), a 10 meter HAAT, and a noiseless channel (thermal noise only).

$$\begin{aligned}
\text{rate} &= W \cdot \log_2\left(1 + \frac{\text{signal power}}{\text{noise power}}\right) \\
&= W \cdot \log_2\left(1 + \frac{(\text{transmit power}) \cdot (\text{pathloss})}{\text{noise power}}\right) \\
&= 6 \times 10^6 \cdot \log_2\left(1 + \frac{(.04) \cdot (8.8 \times 10^{-17})}{2.4 \times 10^{-14}}\right) \\
&= 1274 \text{ bps}
\end{aligned}$$

[5]We believe that another implicit assumption is that the slave device will be regularly communicating with the master device (i.e. they are part of the same system) and thus such a low rate would be unacceptable. However, this need not be the case: a slave may "associate" with any master with whom he can communicate.

At this point the problem has been tacitly recognized but it is not well understood. Television service areas (whose borders often mark a change in channel availability) are often—and usually correctly—thought of as large areas. For example, Figure 8.1 shows the approximate service areas of the television stations on channel 10. From this view, the FCC's lightweight approach to the problem appears reasonable.
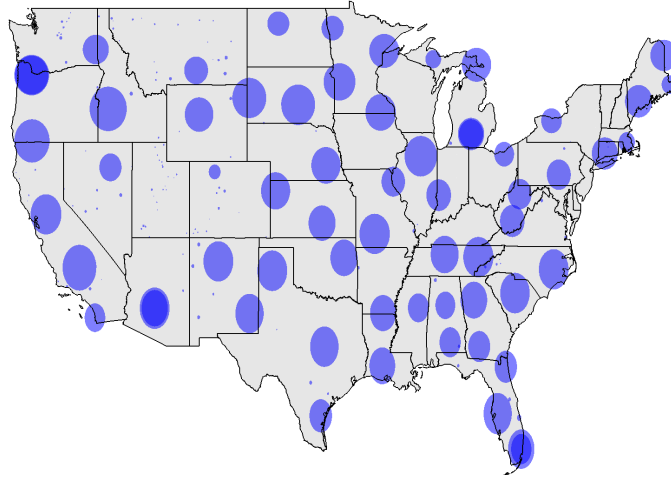
Figure 8.1: Service areas for TV towers on channel 10.

Even when we consider the aggregate effect of the television stations on all channels, the spatial variation appears smooth and predictable. For example, see Figure 8.4(a) in which the number of available whitespace channels is plotted.

However, results such as these lull us into a false sense of security. Consider Figure 8.2 which highlights the variation in the list of available channels across the United States. Each unique list list of available channels[6] is mapped to a unique color, which is then plotted on a map. As can be seen in Figure 8.2(b) (a zoomed-in version for clarity), the list actually changes quite rapidly. This variation, especially the concentric circles, results from two things:

1. The co-channel and adjacent-channel excluded areas for the same protected region will differ slightly (the co-channel exclusions extend about 10 km further than the adjacent-channel exclusions). This leads to some of the concentric circles[7].

---

[6]It may be helpful to think of the list of available channels as a binary vector, with each entry answering the question "is this channel available for secondary use?"

[7]Of course, the real world will differ from the figures shown here as we are using the simple ITU propagation model [29] which does not take terrain into account. Incorporating terrain would mean that service areas would change from being perfect circles to being slightly (or significantly, depending on the variation in terrain) misshapen. (This can be seen by looking at the actual protected contours for TV stations, available on the FCC's website [41].)
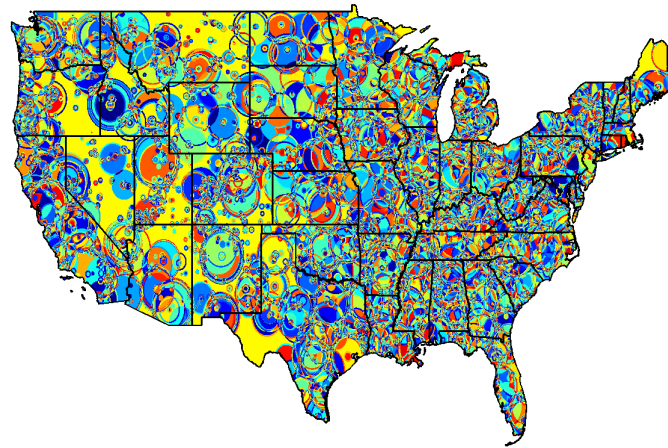
(a) Full map
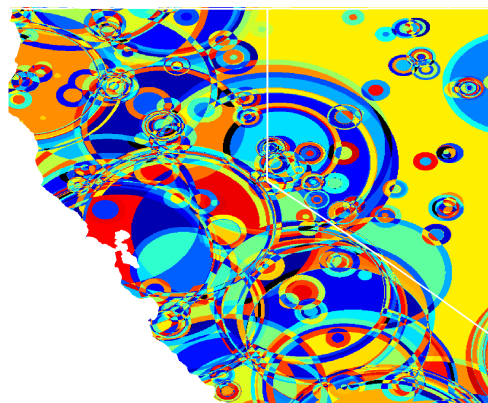
(b) California-Nevada region (zoomed version of Figure 8.2(a))

Figure 8.2: Color-coded maps showing the variation in the list of available channels (a) across the United States and (b) in the California-Nevada region. Each color represents a unique list of available channels.

2. In many cases, the same tower will transmit on different channels (e.g. Sutro Tower in San Francisco broadcasts on over 10 TV channels [91]). Each channel has different propagation characteristics as well as a potentially-different transmit power. This means that although they may aim to serve the same market, each station will end up with a slightly different service area.

Note that we have not said anything about the size of the difference between the various lists, only that they are different. In the same way that relying on the "number of whitespace channels" map in Figure 8.4(a) is extremely optimistic, relying only on this very colorful map is extremely pessimistic. A more reasonable metric is the probability of wrongful transmission, where "wrongful

transmission" is defined as a transmission which is against the spirit of the regulations (e.g. inside the service area of a TV station).

In Figure 8.3, we show the CCDF of the probability that a slave will transmit on a channel which is actually not available at his location, despite being assured by the WSDB and the master that the channel is available[8]. We have plotted several scenarios, each representing the slave's maximum distance from the master. For example, we can see that even if the slave is within 10km of the master, about 40% of the population has at least a 4% chance of wrongful transmission.
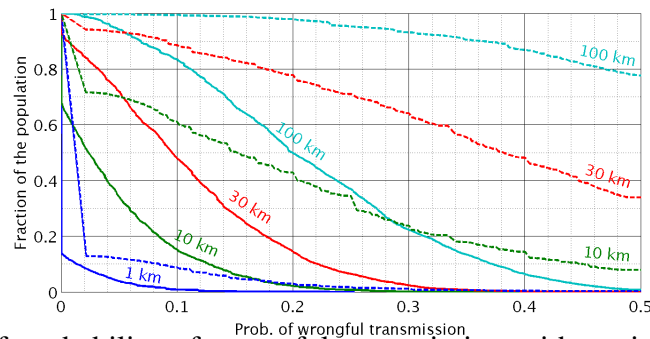
Figure 8.3: CCDF of probability of wrongful transmission with an ignorant (solid lines) and opportunistic (dashed lines) secondary transmitter. Details on this calculation can be found in Section 8.10. Note that the jump in the opportunistic lines from 0 to 0.02 occurs because the probabilities are discrete with minimum step size 0.02.

There are a few points which are important to keep in mind when interpreting these results:

- While 10km (and especially 100km) might seem extreme, note that (1) the master is potentially quite long-range (it may, for example, be a fixed device with height up to 106 meters above average terrain [23, §15.711(b)(3)(iv)(C)]) and (2) the slave has little to no incentive to choose a nearby master if its first goal is to obtain more spectrum. We assumed uniform distributions for an "ignorant slave" (solid lines) but a slave could easily choose to "game the system" by choosing the master which gives the most favorable results ("opportunistic slave," dashed lines).

- The number of channels available to such an opportunistic slave device[9] is shown in Figure 8.4(c) (we assume it can only use masters within 25km). The corresponding probability of wrongful transmission (assuming the slave chooses the most favorable master but then chooses uniformly among the channels "available" to it) is shown by the dashed lines in Figure 8.3. The probability of wrongful transmission is alarmingly high with an optimistic 10km limit on the master-slave range.

---

[8]Details on this calculation can be found in Section 8.10.

[9]Although we tend to think of a model in which the master and slave devices are tightly coupled (e.g. via proprietary protocols which vary by manufacturer), we envision a much richer and liberal ecosystem in which, for example, master and slave devices interoperate naturally.

- The slave need not contact the master via a whitespace channel. In fact, he is not even limited to communicating with the master wirelessly. He thus is not limited to whitespace emissions limits which potentially increases his range (e.g. via ham radio frequencies[10]). See §15.711(b)(3)(iv)(D) of the 2012 FCC rules [23] for more details on the master-slave communication requirements.

- Once the initial exchange is complete, bidirectional communication is no longer required (see [23, §15.711(b)(3)(iv)(D)]). In this manner, a slave device could initiate contact with a master while he is nearby, then move to another location and simply receive periodic updates from the longer-range master. We believe this is an unintentional loophole in the regulations.

8.3   Ofcom's approach to the master-slave problem

The next step to solving this problem has been taken by Ofcom. We will describe this solution and its shortcomings in this section.

The UK's Ofcom has partially accounted for this problem in their rules, using the service area of the master to estimate the maximum distance between the slave and the master [87, ¶5.11].

> For this, the WSDB will use the TVWS availability data obtained from Ofcom and the channel usage parameters received from the master WSD (see 5.35.5) to calculate the coverage area in which slave WSDs are likely to operate. It will then calculate the generic operational parameters that apply within the coverage area based on a number of default (conservative) device parameters. [87, ¶5.38.2]

They then appear to use this estimated distance to create a set of locations which is highly likely to contain the true location of the slave. Using these potential locations, they can calculate which channels[11] will be safe for the slave to use as long as he is at one of those potential locations.

Although there are some subtleties to estimating the master-slave distance (e.g. using the antenna gain of the slave device [87, ¶5.85]), this solution can certainly be made conservative enough to adequately protect the primary's systems. One ridiculous example of this is to assume that the slave is always within 1000 km of the master: while it is an extremely conservative bound, it has a very high probability of being correct.

---

[10]He could technically communicate with the master over the Internet, but we believe this is an unintentional loophole.

[11]Note that part of the motivation for Ofcom's regulations are its vision of variable power limits which depend on the device's location. To understand the subtleties of this, please refer to Chapter 3.
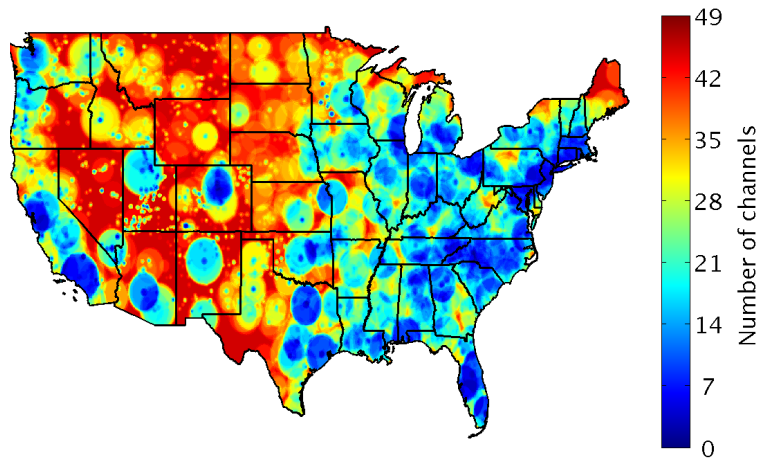
(a) Number of channels available to a device with perfect geolocation
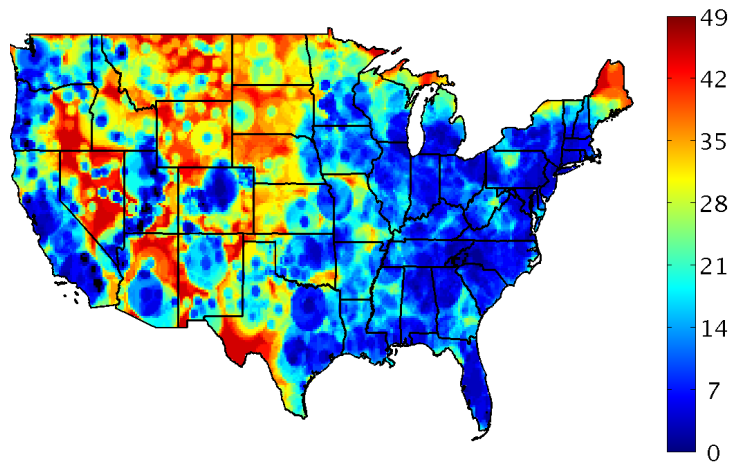
(b) Number of channels available to a slave which knows only that it is at most 25km from the master
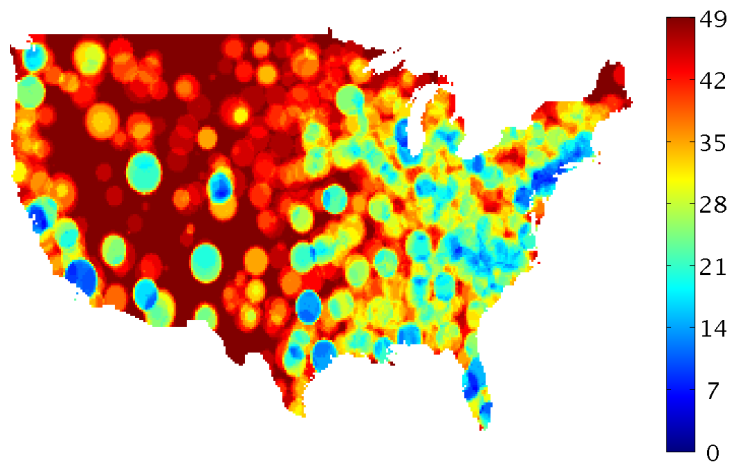
(c) Number of channels available to an opportunistic slave which contacts the "most desirable" master (i.e. the one which reports the most available whitespace channels) within 25km

8.4 The problem with Ofcom's approach

The problem with Ofcom's approach is that a tradeoff has been established: protecting primaries vs. providing reasonable opportunities to slave devices (as compared to devices with geolocation). For example, we see in Figure 8.4(b) that assuming a maximum master-slave distance of 25km drastically reduces the number of whitespace channels available to slave devices in the United States. We have already argued in Section 8.2 that even 25km is a conservative estimate, so it is providing poor service to both primaries and secondaries.

For simplicity, we will assume a fixed master-slave maximum distance (i.e. that the slave is within R km of the master) as opposed to one which varies based on the characteristics of the master and slave devices. Figure 8.5 quantifies this problem for various values of R in the context of the United States.

In Figure 8.5(a), an empirical CCDF (complementary cumulative distribution function) weighted by population shows the raw number of channels lost to the slave due *solely* to his lack of geolocation capability. As we expect, the number of channels lost increases with R. For example, at $R = 2$ km, the median person is losing about 1 channel as opposed to 2 channels at $R = 10$ km.
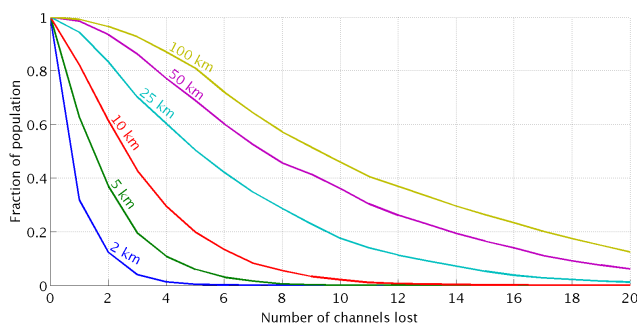
(a) Number of channels lost using Ofcom-like regulations
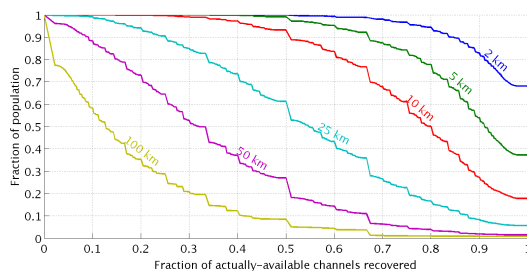
(b) Fraction of actually-available channels which are available to slave devices using an Ofcom-like regulations

Figure 8.5: CCDFs (by population) showing the impact of Ofcom-like regulations

However, channels have different value in different places depending on their scarcity: in urban areas where whitespace channels are sparse, each channel is worth more. In contrast, whitespace

channels are less valuable (individually) in rural areas because they are relatively abundant. Thus in Figure 8.5(b) we look at what fraction of the actually-available (i.e. available with perfect geolocation) whitespace channels a slave device would be able to use under Ofcom-like rules. Again, with $R = 2$ km we see that most people will be able to use most of the actually-available channels whereas with $R = 10$ km the median person will only recover about 80% of their potential channels. This metric is quite useful as it provides a number-of-channels-agnostic approach to quantifying these rules.
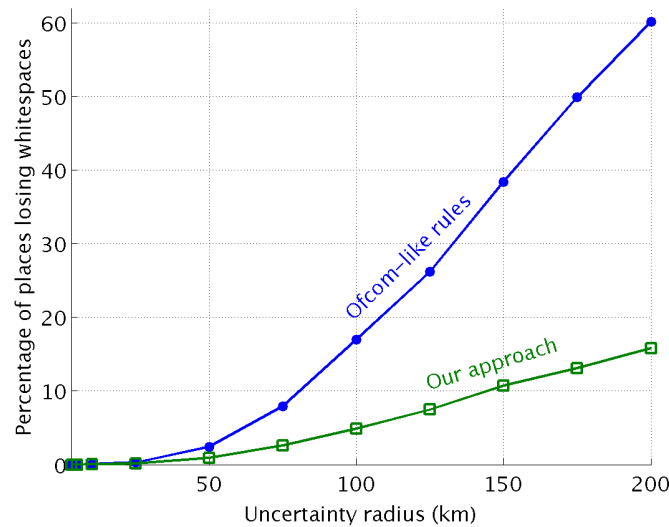
Figure 8.6: Percentage of places that lose access to the whitespaces under various approaches.

Figure 8.6 shows the percentage of places which, purely as a result of location uncertainty, lose *all* access to the whitespaces. For now, it is sufficient to note that Ofcom's approach scales very poorly with the uncertainty radius: at $R = 100$km, 15% of the country loses access to whitespaces that would have been available with geolocation capabilities.

We wish remind the reader that in order to be completely safe, the regulator must choose a distance beyond which they think it will be *impossible* for a slave to communicate with a master. Consider for a moment that a 100-mW transmitter (100 mW is the maximum EIRP for personal/portable devices) can theoretically communicate at 3.2 kbps to another device 100 km away[12]. This means unnecessarily barring slaves from a great number of channels which otherwise could have been safely used[13]. Keep in mind that even a device which actually has geolocation capabilities will be forced to operate as a slave if it cannot successfully locate itself (e.g. GPS fails indoors).

---

[12]Assumptions: channel 2, 10.1m HAAT, clean channel. Even a 40-mW device could communicate at almost 1.3 kbps.

[13]An alternative is to mandate a minimum spectral efficiency for master-slave communications, e.g. 2 bps/Hz.

In the next section, we will suggest a simple enhancement to Ofcom's approach which will greatly increase the opportunities for slave devices while adequately protecting primary systems.

8.5 Proposed enhancement to Ofcom's approach

Let's summarize what we have learned so far:

1. A slave device could conceivably be located tens or hundreds of kilometers from its master.

2. The list of available channels can vary rapidly with location.

3. Using the master's location for the slave is inherently unsafe due to the first two points.

4. Ofcom's solution of using only those channels available in the slave's set of potential locations drastically reduces the slave's opportunity, forcing a choice between primary and secondary quality of service.

We believe that this tradeoff is actually a symptom of the way that people are thinking about solving this problem, rather than an unavoidable fact of nature. Instead of constraining the set of possible slave locations to being a circle, why not allow it to take any shape at all? For example, it is clear how the database should respond to a slave device which can reliably say "I'm in either San Francisco or New York City, but I don't know which." The size or shape of the set of potential locations (which we term the "uncertainty region") is actually *completely irrelevant*.

The question now becomes: how can we meaningfully reduce the size of the uncertainty region? We have previously established in Figure 8.2 that the list of available channels has high spatial variation[14]. There is an entire field which capitalizes on spatial variation in radio signals: location fingerprinting.

Location fingerprinting in conjunction with radio environment maps helps a user to pinpoint his location by correlating signal measurements at his location with those in a pre-populated database[15]. For example, WiFi fingerprinting uses the names and signal strengths of nearby wireless access points to estimate a user's location. This is done by Apple in its iOS when GPS is not available [80].

---

[14]Moreover, its variation is actually correlated to the phenomenon we'd like to detect (i.e. available channels).

[15] **Location fingerprinting**: Location estimation is a general problem which has been addressed extensively. This work has been done with an eye toward sensor networks. In this application, many sensor nodes are deployed to gather information about the environment. The nodes may be statically configured or may be moved from time to time. Determining the exact position of each sensor requires significant effort and potentially hundreds of nodes may be deployed, hence automatic position-finding significantly reduce the costs of sensor networks.

For example, [92] and [93] consider the problem of determining the relative position of nodes in a network. [94] considers a similar problem where a small portion of the nodes know their absolute locations; RSS and TOA are used to estimate distances.

Note that our approach is to actually allow the consideration of a *simpler* problem than the traditional localization problem. We relax these problems by *not* requiring full localization and instead attempt to simply reduce the uncertainty in a device's location. This uncertainty—however large or small—is acknowledged by the whitespace database in that it will only permit the device to transmit on channels which are available for secondary use *everywhere* the device *could* be located. In this way reduced uncertainty will increase the whitespace opportunity for that device but any amount of uncertainty can be accommodated. Thus we can simply "solve" the location fingerprinting problem to the desired fidelity and not worry about fully solving these problems.

We strengthen our argument with the following facts:

- Most devices are transceivers rather than transmit-only devices. Thus they have the prerequisites for rudimentary receiving (i.e. sensing) capabilities.

Another technique for location estimation is location fingerprinting. Location fingerprinting (LF) is a localization technique which uses pre-existing (computed or measured) signal strength maps—typically of WLAN signals—in conjunction with measurements from the device itself. In particular, the measurements are checked against the maps to determine locations in which the measurements are self-consistent.

The accuracy of LF depends heavily on the map data as well as the usage environment. For example, WLAN LF works quite well in urban areas where WiFi hotspots are in abundance (and the maps are likely to be of good quality and recent) whereas it performs poorly in rural areas.

Note that the TV bands (as well as AM and FM radio bands) have ideal qualities for use in LF: (1) stations are relatively static over time; (2) stations are at fixed locations; (3) signals in these bands propagate well.

LF requires map data which is extensive and accurate. Thus, the biggest challenge in LF is creating and maintaining the database of map data [95]. There are two main techniques for database creation: (1) radio environment mapping and (2) use of propagation models to predict the signal strength at various locations. While (1) is more accurate (assuming the landscape doesn't change significantly), it requires substantially more data-gathering effort.

Radio environment maps: The now-conventional approach to whitespace databases assumes that ground-truth (regarding which channels are free to use and where) is obtained via registration and reliable propagation models. However, while registration is easy to assume for TV transmitters in first-world democracies, many other systems and countries cannot rely on such information being easily, publicly accessible.

There is extensive research on how to create and maintain radio environment maps (REMs). Radio environment maps synthesize information from many sensors to create an estimate of the signal strength at a wide variety of locations. REMs are typically used for location fingerprinting, network planning, and signals intelligence.

Creating an accurate REM is a difficult problem for several reasons:

- Gathering accurate data is a costly proposition as it requires many measurements, particularly in fast-changing environments (e.g. cities) [95].

- Synthesizing such large amounts of information may be quite difficult and time-consuming, depending on the collection method [96].

- If sensing information comes from multiple types of devices, the different device characteristics (e.g. sensitivity) must be accounted for.

- Depending on the application, the signal strength may vary with time (e.g. in public safety bands which are used intermittently), further compounding the problem.

- Although sensing for the *absence* of TV signals is quite hard [97], it is in fact relatively easy for a device to determine if it could *receive* TV. TV signals are at or above 15dB SNR inside the service area, something that is quite easy to sense[16].

- The list of channels available for secondaries has a lot of variation, as seen in Figure 8.2(a). Similarly, the list of channels on which users can successfully watch TV also varies a lot (omitted for brevity).

When we combine the three facts above with location fingerprinting, a potential solution is quite clear: a slave can reduce the size of its uncertainty region using an easy-to-generate report on which TV channels it receives[17] (i.e. someone could watch TV on that channel there).

So how does it work? Let us first explain our process with a toy model which will also be used later. We will then formalize our proposal.

8.5.1   Illustrative example

A simple example is illustrated in Figure 8.7. This example shows how the uncertainty region of a secondary device changes as it learns more about its environment.

Figure 8.7(a) simply sets up the example. It shows the service areas of towers on two channels, a red channel and a blue channel. True to real life, these service areas may or may not overlap.

Figure 8.7(b) shows in gray the initial uncertainty region of a particular secondary device. The size of this initial region is determined by the device's estimated maximum distance to the master and, since no other information is known yet, takes the shape of a circle. The secondary's true location is marked by a star. Notice that TV signals for both channels will be strong at this location.

In Figure 8.7(c), we see the initial reduction in the size of the uncertainty region. Having sensed a strong TV signal on the red channel, the secondary can use this information to exclude areas where TV signals on the red channel are expected to be weak[18]. Notice that the resulting uncertainty region is now disconnected (which is *not* a problem) and, more importantly, reduced in size.

The final figure, Figure 8.7(d), shows the secondary's uncertainty region after sensing a strong TV signal on the blue channel. Again, regions where blue-channel TV signals are expected to be weak are excluded from the uncertainty region. We see now that the uncertainty region is just a fraction of its initial size, demonstrating the power of our location-fingerprinting-like technique.

---

[16]At first glance, it may appear as though this solution applies particularly well only to the TV whitespaces. However, it is not unreasonable for a device to use TV signals to reduce its location uncertainty in order to gain access to a *different* set of whitespaces. After all, the PCAST report [54] suggests that all whitespaces should use a common database.

[17]Notice how we took a somewhat useless or disappointing result in the sensing-only world—"oh no, that channel's in use... I guess I'll try another one"—and turned it into useful information!

[18]In this chapter, we consider the sensing output to be either "can receive TV" or "cannot receive TV." In reality, at minimum a third option—"cannot tell"—should be allowed. Conceivably, more precise outputs could be used.

(a) The service areas of towers on two channels, red (dotted lines) and blue (dashed lines), are shown.

(b) Uncertainty region of secondary is shown in gray. Its true location is marked by a star.

(c) Uncertainty region after the secondary senses a strong TV signal on the red channel.

(d) Uncertainty region after the secondary also senses a strong TV signal on the blue channel.
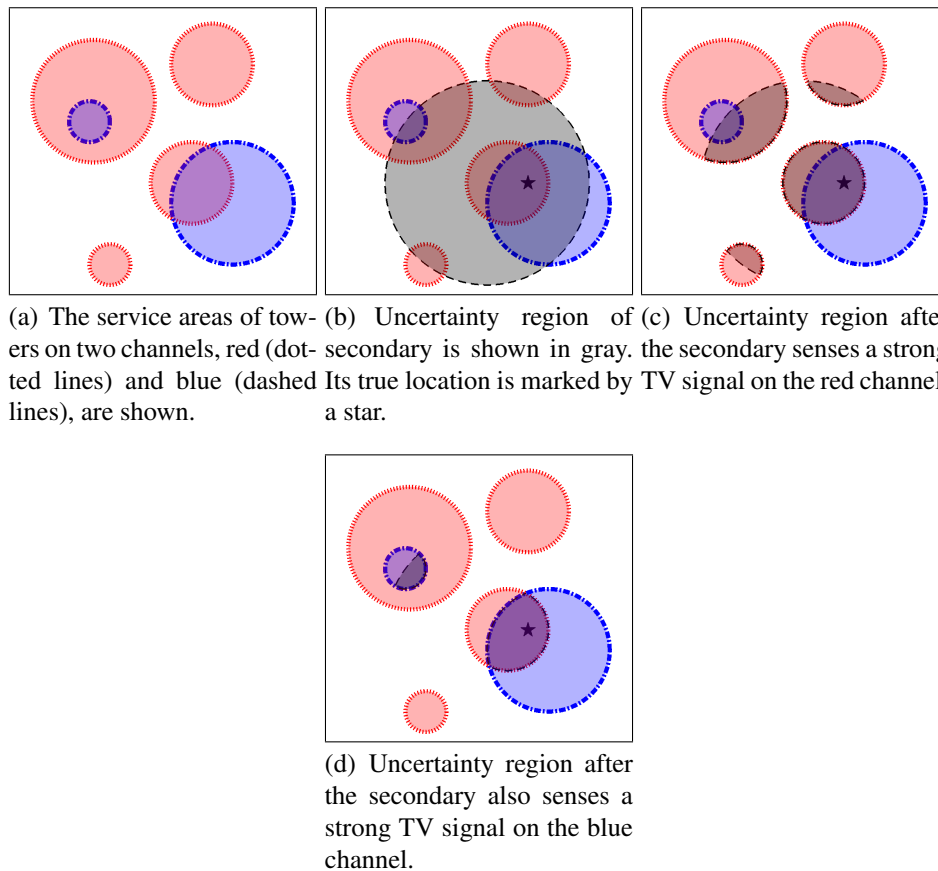
Figure 8.7: To illustrate our approach, we present a toy example as a series of figures. In (a), the service areas of TV towers on two channels, red and blue, have been marked. (b) shows in gray a sample "base" uncertainty region for a secondary device which knows only its distance from the master. A star marks the secondary's true location. After sensing a strong TV signal on the red channel, the secondary's uncertainty region can be reduced, as shown in (c). Notice that the region is noncontiguous but reduced in area. The secondary subsequently senses a strong TV signal on the blue channel which further reduces its uncertainty region, as shown in (d).

A side note: our assumption is that the slave simply notices when it receives TV on a particular channel (thus it is in the service area for *some* tower on that channel). However, since it will be within the service area, one might suggest that it actually decode the TV signal and determine the station ID, further reducing its uncertainty. While this is possible in theory, it might not be practical: including an ATSC decoder module might be too expensive to justify for devices which are not already intended for watching TV—for the same cost, a GPS unit would solve the same problem (when outdoors).

8.5.2 Details of our proposed solution

We propose that the slave be allowed to use sensing techniques[19] to determine the channels on which it can receive TV in order to reduce the size of its uncertainty region. Although the slave would also benefit from identifying those channels on which it cannot receive TV, a device which can do this reliably may as well be a sensing-only device. Our solution hinges on the fact that the presence of a signal is much easier to detect reliably than the absence of one.

Once the device has identified some channels on which he receives TV, it communicates these to the WSDB (through the master). The whitespace database uses this information in conjunction with its radio environment maps and the slave's maximum distance to master in order to narrow down the set of possible slave locations. The WSDB does so through Algorithm 1 which follows the approach described in Section 8.5.1.

Algorithm 1 Potential WSDB algorithm to calculate list of truly available channels for slave device

---

Require: T: list of channels on which TV is received at the slave ($|T|$ may be 0)
Require: L: precise location of the master device
Require: R: estimated maximum distance between slave and master ($R \geq 0$)
Ensure: Slave device transmits only on channels which are truly available
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ Determine the uncertainty polygon
1: Uncertainty polygon, P, initialized to a circle of radius R centered at L
2: **for all** t in T **do**
3: $\quad$ Let A_t represent the coverage area of t
4: $\quad$ $P \leftarrow P \cap A_t$
5: **end for**
$\qquad\qquad\qquad\qquad\qquad$ ▷ Determine the channels available everywhere in the uncertainty polygon
6: List of channels available for slave use, C, initialized to the empty set
7: **for all** whitespace channel, W **do**
8: $\quad$ **if** W available for whitespace use everywhere in P **then**
9: $\quad\quad$ $C \leftarrow C \cup \{W\}$
10: $\quad$ **end if**
11: **end for**
$\quad\quad$ **return** C

8.6 Evaluation of proposed solution

First, we compare our proposed approach with the Ofcom-like method from Figure 8.5 and the results are shown in Figure 8.8. For high values of R, learning the channel states is incredibly valuable (e.g. the median person goes from recovering just over 10% of his actually-available

---

[19]We recognize that this is actually not as simple as it sounds. For example, the slave will need to determine that he hears a TV signal as opposed to noise from other whitespace devices. However, in most cases he should receive either TV signals *or* WSD signals, not both: recall that WSDs cannot transmit within the service area of a TV station.

channels to recovering over 65% of the same channels in the $R = 100$ km case). Naturally, the differences are smaller for small values of R since there was less room for improvement with the original Ofcom-like method.
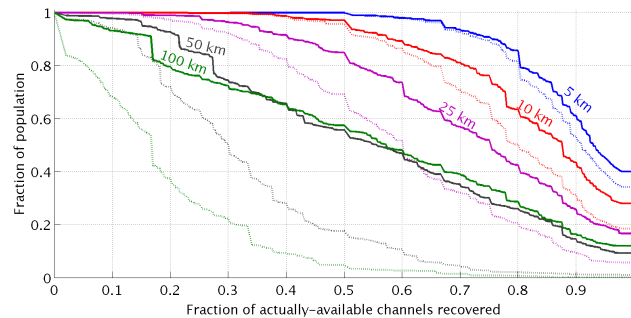
Figure 8.8: CCDFs (by population) comparing our approach (solid lines) to an Ofcom-like approach (dotted lines) for each uncertainty radius R.

Now we've shown that there are real gains to be had from using this not-very-complicated system, let's see what's going on in more detail. We performed a similar test (via Monte Carlo simulations) on real-world data[20] and the results are shown in Figure 8.9. Each subfigure shows a different original uncertainty distance so as to be comparable to the results shown earlier. Each subfigure shows the results for slightly different scenarios:

- **LrxO** ("learn rx [received channels] only," blue solid line): learn about channels on which TV is received[21]. If TV is not definitely received on a particular channel, we conclude nothing after "learning" its channel state. (Note: it would be more accurate but less useful if this line simply stopped after we had learned all of the received channels. However, we extend it horizontally for the purposes of comparison with the other lines.)

- **LrxF** ("learn rx [received channels] first," red solid line): learn first those channels on which TV is received (it's faster to sense these channels so in reality even a sensitive slave would do this) and then confirm that the other channels do not receive TV.

- **LA** ("learn all," purple solid line): learn all channel states in a random order.

The x axes of the plots represent the number of channel states which have been learned by the slave. In other words, $x = 0$ is equivalent to Figure 8.7(a), $x = 1$ is equivalent to Figure 8.7(b), etc. The y axes list the percentage of channels that can be recovered (as compared to the number of channels recoverable with perfect geolocation).

---

[20]Information on our methods can be found in Section 8.10.

[21]Within this set of channels, channel states are learned in a random order which varies for each point in the Monte Carlo simulation.
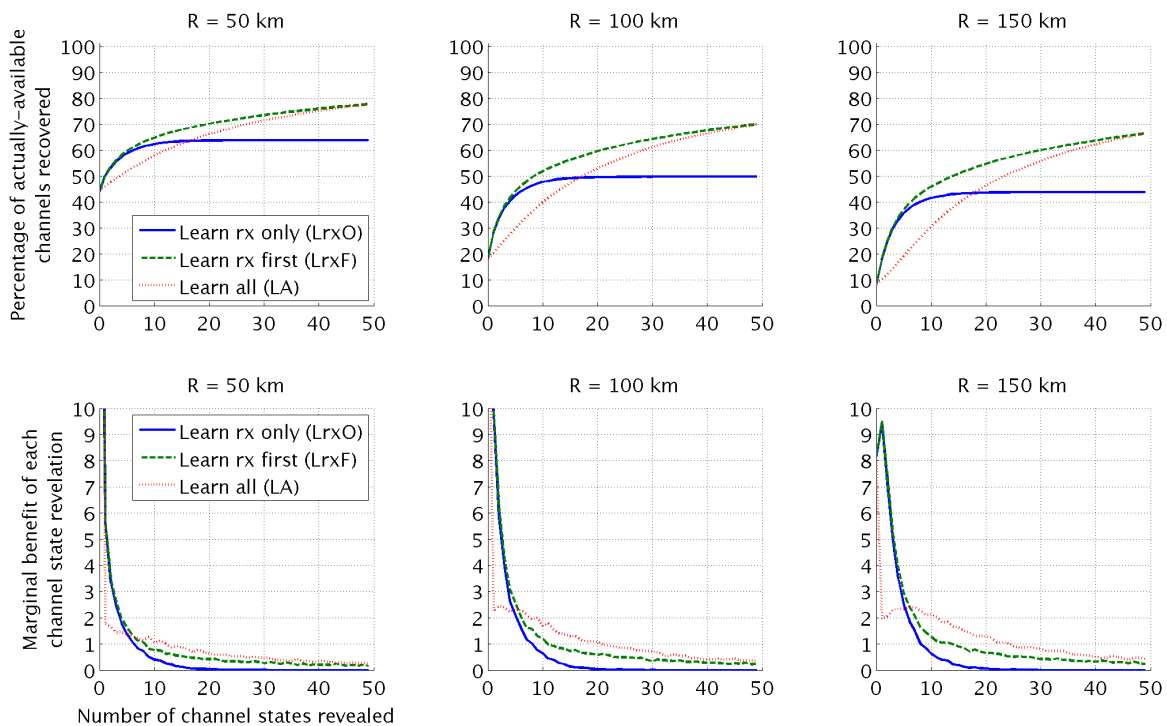
Figure 8.9: Percentage of recovered channels as a function of the number of channel states revealed under various schemes. Solid lines represent absolute recovery percentage while dashed lines indicate the marginal benefit of learning a new channel state.

The solid lines represent the absolute recovery percentage, which is naturally quite interesting. However, it is also interesting to see what the marginal benefit of learning each channel state is and this is shown by the dashed lines (colors match). Let's notice a few things about these plots:

1. As expected, the absolute recovery percentage (top plots) decreases as the uncertainty radius R increases.

2. LrxO (the solid blue line) provides the most benefit (as compared to LA, the dotted red line) for awhile until there are no more channel states to learn. At that point, continuing to learn anything (as opposed to receiving no new information) obviously wins out.

3. If we were allowed to continue learning channels after we ran out of channels with the LrxO method (shown by LrxF, the dashed green line), we would continue to do better than LA. It appears that more information is revealed if you know you're within a TV's service area than if you're not (comparing LrxF and LA). This may be because being inside of a service area is a low-probability event (look at the sparsity of towers on each channel). It may

also be because the state of the adjacent channels can be more easily predicted (there is a high probability that they are also unavailable for secondary use due to adjacent-channel exclusions).

4. This last conclusion is also supported by the graphs of the marginal benefit of learning each channel state (bottom plots). The marginal benefit of learning the state of a random channel (LA, the dotted red line) is initially lower than that of learning a channel on which TV could be received (the other two lines). As more channel states are revealed, this ordering switches. This is because (1) the LA approach is still learning a mixture of received and not-received channel states while (2) the LrxF approach is learning only not-received channel states. The marginal benefit of the LrxO approach obviously goes to zero because no further channel states are revealed once all TV-received channels are known.

8.6.1 Connection to location fingerprinting

Location fingerprinting (LF) is a technique used to determine a device's location based on its knowledge about the environment. This is achieved by comparing its observations with global data while attempting to find a unique location at which its observations are consistent.

Our proposed approach is very similar to location fingerprinting in that it also uses local observations matched against global data. There are two main differences, though: (1) we already have the global data in the TVWS databases and (2) the *problem* is different. The goal of LF is to uniquely pinpoint a device's location whereas our approach is to identify a set of channels which are safe to use at any of the potential locations. It is in fact possible to find a common list of safe channels despite having very distant potential locations.

The key point is this: location uncertainty is *not* the same as uncertainty in the channel list. Heretofore we have discussed the location uncertainty as if it is the only important thing. However, what one *actually* cares about[22] is the uncertainty in the channel list: if we know which channels are safe to use, it doesn't matter where the slave is located. This is important because it is possible to have a high location uncertainty but have a good idea of which channels are safe to use. So note that we will not necessarily know where the slave is even if we know which channels it can safely use. In fact, even in the basic case where the slave knows only its maximum distance R to the master, it can still recover some channels without knowing its location very precisely, as shown in Figure 8.5(b).

8.7 Consequences: added value for databases

In this section, we discuss some of the opportunities for database providers to add value under our proposed system.

---

[22]It is certainly possible that there are reasons beyond safety for wanting to locate the slaves. However, we take the viewpoint that safety and deployability are the main concerns of the regulators.

8.7.1 Channel state discovery

Naturally the order in which channel states are revealed will have an effect on how quickly and efficiently the uncertainty can be reduced. Databases seeking to add value could interact with the client (rather than just receiving a list of channel states) to tell him which channel state to sense next (e.g. "can you definitely receive TV on channel C?"). This would reduce the discovery cost (in time and energy) for the slave by telling him where to concentrate his efforts.

8.7.2 Types of fingerprint information

Databases could also compete on how much and what kinds of information they can safely fuse. For example, databases might contain additional radio environment maps (e.g. for air traffic controller frequencies or cellular bands) which would let slave devices use fingerprint information from other bands. If we accept the vision of the PCAST report (shared whitespace infrastructure), it is likely that databases will already contain such information.

8.7.3 Types of location information

In future work, we plan to discuss the idea of using multiple services for localization. For example, a slave device might find himself in the following situation:

- Within R_1 km of master 1.

- Within R_2 km of master 2.

- Able to detect strong TV transmissions on channels $C_1, ..., C_n$.

A database which can safely fuse a wide variety of location information would clearly be of use to slave devices.

8.8 Conclusion

This chapter has described in some detail the FCC's and Ofcom's approaches to safely enabling slave devices (i.e. devices which do not have geolocation capability). We have described the most glaring problem with the FCC's approach, namely its indifference to the distance between the master and slave. We have shown quantitatively that this is potentially a significant problem, especially if the whitespace device chooses its master opportunistically.

We have also described Ofcom's approach to this problem, in which they estimate the distance between the master and the slave. We have shown a major downside to this approach: it necessarily establishes a tradeoff between primary protection and secondary quality of service.

We went on to propose an enhancement to Ofcom's approach which leverages existing capabilities in secondary devices to safely allow them to recover much of what was unnecessarily lost with Ofcom's approach. We showed qualitatively that our proposed approach could provide gains over existing approaches. This lessens the tension between primary and secondary because now the distance can be very conservatively estimated without significantly negatively impacting the secondary device.

Although we have proposed an enhancement that appears to provide gains for all parties, the approach itself is not actually the most important part of this work. Instead, we wish to convey the message that localization services can take many forms, as seen, for example, in the iPhone which aggregates data from GPS, WiFi fingerprinting, and cellular positioning. In fact, our proposal is very similar to assisted GPS as seen in cell phones. Localization is an integral part of whitespace devices, and constraining it to a narrow definition (e.g. 50-meter accuracy) is unnecessarily restrictive to innovation.

In the regulatory sense, localization should be loosely defined as "the ability to provide information from which a set of *potential* locations can be computed." Databases can then choose if and how to fuse this information, potentially with guidance from a regulator or standards body. Since the databases are software-upgradeable and overseen by the regulator, any bugs or loopholes can be quickly patched. In the worst-case scenario, a particular type of information can be deemed unsafe and hence ignored.

8.9   Future work

8.9.1   Consequences of incorrect channel state estimates

We have not analyzed the consequences of incorrect channel state estimates, though we have tried to minimize the likelihood of such an event by suggesting that the slave identify only those channels on which it can definitely receive TV (where the chances for a false positive are quite low due to the high SNR conditions). However, the potential for harmonics in nearby channels complicates this as well.

Incorrectly identifying a channel state will, with our simple proposed algorithm, certainly result in an incorrect location estimate because of the nature of the algorithm. However, if the slave identifies enough channel states, an inconsistency may eventually be discovered (e.g. "there are no places where you can watch channel 2 and channel 5 so one of your states must be wrong") so that at least the error can be recognized. Future work in this area may draw on results in the location fingerprinting literature to address this problem in a robust way.

8.9.2   Errors caused by propagation model inaccuracies

It has frequently been noted [98] that propagation models are inaccurate when compared to the real world. Hence the database may incorrectly estimate the service areas of the TV towers (though

it could be seeded with the predicted contours from the FCC) and thus incorrectly estimate the slave's list of available channels.

However, slave devices need not give the databases a binary answer to "do you receive TV on channel C?". Instead, they could feed back the signal strength sensed on that channel and the database can apply the threshold. Since the slaves will also be transmitting their device IDs which link them to a specific make and model of device (to check that they are allowed to transmit), the database could also use information on the sensitivity of their sensing to build up a map of signal strengths throughout the nation. In fact, there is already work in this field [96, 99]. Thus over time we can refine the information in the databases to minimize this problem. An important thing to note is that the quality of our estimates will improve with the number of devices, thus improving the accuracy at the same rate as accuracy becomes important.

8.10 Methods

This section contains details on the methods used to produce the results in this chapter. While we believe that these details are important, we expect that they will not interest the average reader, hence they are situated at the end. All relevant code can be found after publication at [6].

We use three sources of data for our models:

- Population data from the US 2010 census [10, 11]

- ITU propagation model [29]

- TV station assignment data from the FCC [28]

We also assumed for simplicity that WSDs can transmit on any of the whitespace channels, rather than being artificially limited to, e.g., channels above channel 20. The current FCC rules for TV whitespaces allow only fixed WSDs (i.e. not slaves) to transmit on channels 2-20.

Probability of wrongful transmission

1) Ignorant secondary: The probability is calculated thusly:

- Suppose the master is at location L and that we know the slave is within distance R of the master. Draw a circle of radius R around L. This represents the set of possible slave locations, S.

- For each channel C that is *available* for secondary use at location L, calculate the fraction of the area of S in which channel C is *not* available for secondary use. Call this fraction $P_{wrong,C}$. This represents the probability that the slave will wrongfully transmit given that it is operating on channel C and assuming that his location is uniformly distributed within S.

- Average $P_{wrong,C}$ across all channels which are available at the master. This is the probability that the slave transmits wrongfully given that it chooses a channel uniformly from the list provided to it by the master.

- The above process calculates the probability of wrongful transmission at a location L. Repeat this process for many locations across the United States, then calculate the CCDF weighted by population. This is what is represented by solid lines in Figure 8.3.

2) Opportunistic secondary: For the probability of opportunistic-and-wrongful transmission, we create an "opportunistic channel list" (call it O) by marking a channel as available for secondaries if it is available *anywhere*[23] within S. The slave is assumed to choose uniformly among the channels in O, so the probability of wrongful transmission in this case is simply

$$\frac{|O| - |C|}{|O|}$$

Note that since $|O| \leq 49$ (there are 49 TV channels), the minimum nonzero value of this fraction is $\frac{1}{49}$ which accounts for the initial jump in the dashed lines in Figure 8.3.

Evaluation of proposed enhancement

Our real-world simulations are Monte-Carlo style for tractability reasons. In general, we use a discretized map of the United States, as illustrated in Chapter 10. All data (e.g. available channel lists, TV service areas, population data) is discretized before being processed. Typically we use 3200 x 4800 points to cover the entire US, which means that the typical point represents roughly one square kilometer of area. In some cases, we evaluate on a much smaller grid (e.g. 200 x 300) for tractability but the underlying data (e.g. channel lists) comes from the higher-resolution map.

---

[23]For practical reasons such as numerical accuracy, we require that the channel be available for secondary use in at least 1% of S.

Part V

Future whitespaces

Chapter 9

GSM whitespaces

The work in this chapter was presented at IEEE DySpAN 2014 [39] and won the best student paper award for the policy track. This work was done in collaboration with Shaddi Hasan, Kurtis Heimerl, Kashif Ali, Sean Roberts, Anant Sahai, and Eric Brewer[1].

The TV whitespaces are an excellent proof of concept for the idea of dynamic spectrum access and sharing. However, as the PCAST [54] report acknowledges, we cannot rely on the TV whitespace alone. It is important to provide unlicensed spectrum across a range of frequencies and for a variety of purposes.

This chapter shows how the concept of spectrum whitespaces can be extended to the GSM band. Unfortunately we are unable to estimate the amount of GSM whitespace due to a lack of data (carriers typically do not make the locations of their cellular towers public) but field measurements suggest that the opportunity is quite large. If and when the data becomes available, integration with WEST would be straightforward.

This chapter also highlights another key point: having a flexible regulatory architecture allows for creative and efficient solutions to problems. For example, we propose using existing handsets to do a form of collaborative sensing for the incumbent, using mechanisms that have been widespread since the early days of cellular phones. This "hack" allows for additional incumbent protection that would have otherwise required dedicated and cost-prohibitive sensing hardware.

9.1   Introduction

Cellular networks are the largest communication systems on Earth, with billions of people relying on them for basic communication services. These networks have positive economic impacts ([100,

---

[1] My contributions were primarily in the form of high-level ideas (e.g. nomadic operation) and the necessary background on whitespace policy and stakeholders. The other authors contributed the test bed and the work to retrofit it for the purposes of this work. The writing was done primarily by the other authors, although I helped edit the manuscript.
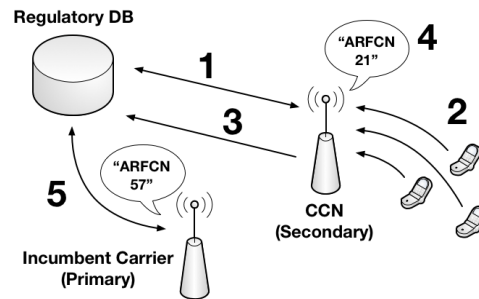
Figure 9.1: Overview of Nomadic GSM. (1) The secondary initially queries a regulatory database for available channels in its area. (2) The secondary gathers measurement reports from its subscribers' phones. (3) Secondaries report spectrum usage (both their own and measured) and service usage (e.g., number of calls and SMS) to a database on a regular basis. (4) Secondaries use measurement report data and guidance from the reg. DB to pick future non-interfering channels to use, and regularly change channels. (5) Optionally, primaries update the regulatory database with their own spectrum usage and query it to obtain reports on what spectrum in use by secondary operators.

101]), and spending on telecommunication matches that of a necessity among the poor [102]. As such, providing universal access to cellular service is an important policy objective, with universal service obligations (USO) and subsidies being the primary policy mechanisms for accomplishing this. Although these programs are important and have been widely adopted, they have failed to provide truly universal service: hundreds of millions of people worldwide live outside the coverage area of existing cellular carriers. These people largely live in rural areas, where traditional telcos are unable to operate profitably due to high costs and low subscriber density. Even with USO subsidies, carriers are inherently disinclined to invest in rural infrastructure, which is marginally profitable at best, particularly given the opportunity cost compared to investing in lucrative 3G and 4G infrastructure in urban areas.

Fundamentally, USOs assume a *top-down* model of cellular deployment where large-scale organizations bring cellular coverage to remote rural areas for subsidies. Historically this made sense; there were only a handful of organizations capable of building and operating cellular networks in any particular country. Yet recent innovations invalidate this assumption. The cost and complexity of building and maintaining a GSM cellular network has decreased to the point where individuals or rural communities can create their own micro-scale cellular networks [5]. These *bottom-up* "community cellular networks" (CCNs) have demonstrated that local communities can operate their own telecommunications infrastructure and provide essential communication services.

Unfortunately, despite the potential benefits of community cellular networks, regulatory barriers exist. GSM uses licensed spectrum, and gaining access to long-term spectrum licenses is almost impossible for micro-scale rural operators. To solve this, we argue for spectrum sharing in *GSM whitespaces* to provide GSM service in rural areas. GSM whitespace refers to licensed GSM

spectrum that is unused in a particular geographical area[2] and thus could be re-used by a secondary operator without interfering with the primary license holder. By allowing CCNs to operate in GSM whitespaces, regulators would empower rural communities to build infrastructure appropriate to their own needs, without waiting for incumbent carriers to begrudgingly allocate resources their way.

To enable this, we propose Nomadic GSM (NGSM), a hybrid sensing and database-driven approach for GSM spectrum sharing (Fig. 9.1). NGSM takes advantage of the fact that GSM handsets continually measure spectrum occupancy in their neighborhood and report these measurements back to their serving cell. By doing so, we can sense for potential interference at the *client* device, avoiding the hidden node problem. Although certain edge cases necessitate a spectrum occupancy database, NGSM enables secondary operators like CCNs to share licensed spectrum without requiring cooperation or participation from existing licenseholders. NGSM works with existing, unmodified GSM handsets. As such, it is deployable today, and we demonstrate this with a prototype deployment in a CCN in Papua, Indonesia.

The contribution of this chapter is as follows. First, we define GSM whitespaces and describe Nomadic GSM, a scheme for dynamic spectrum sharing in GSM whitespaces that enables secondary operators—community cellular networks—to provide service without interfering with each other or with primaries and that does not require explicit cooperation or engagement with primary license holders. Next, we consider the opportunities and risks spectrum sharing presents to major stakeholders and how NGSM addresses these. Finally, we demonstrate the feasibility of our proposal by building, deploying, and evaluating a prototype implementation of NGSM that is compatible with existing, unmodified GSM handsets. We close with a discussion of why the whitespace approach works better than the obvious market-based alternatives and a path forward for regulators.

9.2 Related Work

9.2.1 Policies for Rural Service

One policy mechanism for bringing coverage to rural areas is a universal service obligation (USO) [103]. USOs, originally developed for postal service, refer to a requirement for a baseline level of service to every resident of a country. An example is the US Telecommunications Act of 1996 [104], whose goals were to promote the availability of quality services at reasonable rates, increase access to advanced services, and provide these services to all consumers, including low-income or rural people. Similar regulations exist in many countries, including Indonesia [105], where we deployed our pilot system.

Despite these lofty goals, hundreds of millions of people in the world remain without basic telecommunication services. The reasons for this are fundamentally economic; operators would prefer

---

[2]Note we do *not* refer to the space between adjacent GSM channels.

to work in areas where they are profitable without the headache of dealing with USOs [106]. Researchers have attempted to address some concerns with USO systems through competitive means [107], including USO auctions [108]. This work argues for a fundamentally different model of rural access; one owned and operated by rural entrants and communities themselves. This would free traditional firms from USOs while providing coverage in underserved markets.

9.2.2 Locally-owned Infrastructure

Local or community ownership or development of critical infrastructure has a long history. A well-known concept is coproduction [109], targeting infrastructure such as irrigation [110]. There is a similar history of small-scale cooperative or locally owned telephony networks [111]. Modern cellular networks have largely ignored these models in most of the world, focusing instead on nation-wide networks and coverage. The Village Phone [112] was a system where "phone ladies" would buy handsets and sell use, similar to a phone booth; while the network infrastructure was owned by a nationwide carrier (Grameen), local entrepreneurs provided access to the network to their community. Galperin et al. [113] proposed running cellular networks as small-scale cooperatives, using evidence of earlier cooperative telephony networks in Latin America as a motivating example. Elgar made similar arguments for the viablity of "bottom up" telecommunications [114]. However, only recently has cellular equipment become inexpensive enough for these models to be economically feasible.

One example of cheap cellular equipment is OpenBTS [115]. OpenBTS is an open-source GSM base transceiver station (BTS) implementation which has enabled a wide range of projects aimed towards building small-scale "community cellular" networks [4]. Heimerl et al. demonstrated the viability of independently run, locally operated cellular networks [5]. Similarly, Rhizomatica has deployed several community-run cellular networks in Oaxaca, Mexico [116]. Zheleva et al. [117] deployed a similar system for purely local communications in Zambia. Of these networks, only the Oaxaca network has a short-term experimental spectrum license; the rest operate without licenses. This reality motivates our desire to develop a mechanism for effectively licensing and regulating spectrum access for community cellular networks.

9.2.3 Cognitive Radio

The literature on cognitive radio, whitespaces, and dynamic spectrum sharing is vast; while most work in the space focuses on TV whitespaces (TVWS), our work is more closely related to work on re-use of cellular spectrum. Sankaranarayanan et al. [118] propose reusing timeslots in a GSM cell for adhoc networks during periods when the GSM cell is lightly utilized. Buddhikot et al. [119] describe a system for indoor femtocells to dynamically share spectrum with incumbent carriers by operating over ultra wide bands. Yin et al. [120] proposes a similar system and provides measurement results which indicate that unused spectrum (i.e., whitespace) exists even in a dense, urban environment (Beijing). The assumption in the community, however, seems to be that cellular spectrum is efficiently used and that finding GSM whitespace is challenging.

In contrast to these, we focus on reusing GSM whitespaces to provide GSM service by means of macrocells in rural areas. Moreover, rather than relying on fine-grained spectrum sharing, we rely on spatial separation to provide coarse-grained sharing at the level of full GSM channels. This high margin for error—due to the large distance between primary and secondary networks—along with our novel sensing strategy is likely to be more appealing to incumbents.

9.3 Community Cellular Networks

Historically, cellular networks have been expensive to build and complicated to operate; this is particularly the case for rural cellular networks [121]. A single rural GSM macrocell can cost upwards of US$500,000 to build, not including the supporting network core infrastructure that the network operator must already possess. Macrocells have high power consumption, and in areas without reliable grid power must rely on diesel generators; the fuel for these generators is a major ongoing operational expense and target for theft [122]. These factors have created a situation where only a handful of entities, primarily large corporations or governments, are able to operate cellular networks. Spectrum licensing compounds this: not only must an organization who wants to obtain a license pay large amounts of money, they also must understand how spectrum is regulated, how and when auctions take place, and how to participate in those auctions, all factors which raise the barrier to entry for small organizations.

Recent technological innovations—notably, low-cost software defined radios and open-source software such as OpenBTS [115]—have challenged this status quo. A rural community can build and operate their own cellular network for under $10,000 in capital expenditure [5]. Low-power equipment can be operated using solar panels, dramatically reducing operational expenses. These networks rely on voice over IP (VoIP) technology and can thus use any available Internet backhaul to connect to the global telephony network), including satellite or fixed wireless broadband.

These advancements have enabled a new model, the *community cellular network* [4]. Community cellular networks are locally owned and operated, and they consist of at most a handful of BTS sites. Such networks exist in Papua, Indonesia [5] and Oaxaca, Mexico [116]. Not only are these networks effectively serving rural communities where incumbent carriers have failed (or even refused) to do so, they are financially sustainable for the local operators. The Papua network, for example, generates a revenue of around US$1,000 per month, which while minuscule by traditional telco standards represents a good business opportunity for a local entrepreneur. Moreover, both of these networks were built and are operated without any involvement or coordination with existing operators.[3]

Compared to traditional cellular networks, the core advantage of CCNs is that they enable local independent entrepreneurs to solve their own communication problems. There's no reason existing telcos cannot take advantage of low-cost equipment targeted towards CCNs to build out

---

[3]Indeed, the network in Papua is operating without a license, though it has not received any complaints.

rural infrastructure, but access to low-cost equipment isn't enough to ensure sustainable operation in rural areas. A key finding from prior work on community cellular networks is that locally operated microtelcos have the flexibility to make decisions that traditional telcos cannot. In the example of the Papuan CCN [5], service was *coproduced* [109, 110] with the local community: pricing decisions were made locally, and electricity and backhaul were sourced from a school in the community. The microtelco in Papua was also able to set prices that were appropriate for their own community and costs, thus ensuring sustainability. A large-scale telco does not have this flexibility—the overhead of managing small, potentially informal, relationships with many widely distributed partners is prohibitively expensive and time consuming. Yet these relationships and the understanding of local community structure and norms are the key advantages of local entrepreneurs.

Beyond simply being more affordable, CCNs also have inherent advantages for providing rural service. Although other technologies and spectrum bands (e.g., WiFi) could provide rural communications services, using operating GSM base stations in spectrum traditionally used for GSM networks leverages the wide installed base of billions of existing handsets with existing charging, repair, and distribution infrastructure. Inexpensive and ubiquitous, existing GSM phones ease adoption by providing a familiar experience for end users. People want to be able to use their existing phones, and it's unlikely any manufacturer will produce a cheap, durable phone just for rural areas using a novel protocol.

CCNs put operating cellular network infrastructure within reach of individuals. It is technically and economically feasible for individuals to deploy this infrastructure for their communities on their own initiative, as many already do with WiFi infrastructure. The primary obstacle is access to spectrum: unlike WiFi, devices for cellular networks operate in licensed bands. Removing this barrier is vital to widespread deployment of community cellular networks, and their unique strengths argue for policy mechanisms to support their growth. GSM whitespace presents an opportunity to resolve this tension.

9.4 GSM Whitespaces

9.4.1 Defining GSM Whitespace

GSM whitespace refers to spectrum that has been licensed to carriers for GSM networks but is unused in a particular geographic area. As defined, GSM whitespaces are incredibly common worldwide: due to exclusive licensing of 2G GSM spectrum, any areas that are unserved by telcos are guaranteed to have unused spectrum in the 2G GSM bands.

Consider the case of Indonesia. Fig. 9.2 shows the national cellular coverage map for Indonesia.[4] Although the entire GSM900 and GSM1800 bands have been licensed to carriers in Indonesia (Table 9.1), vast swaths of the nation remain without any coverage. The largest provider, Telkomsel,

---

[4]Obtaining accurate data on what areas are actually served is very difficult. The data from this figure comes from the map of international roaming coverage published by AT&T. It generally matches self-reported tower locations found in annual reports [123] and crowdsourced coverage maps [124].

Uplink (MHz)	Downlink (MHz)	Licensee
890.0 - 900.0	935.0 - 945.0	Indosat
900.0 - 907.5	945.0 - 952.5	Telkomsel
907.5 - 915.0	952.5 - 960.0	XL

Table 9.1: Bandplan for the GSM900 band in Indonesia [14]. The entirety of the band has been granted to these three carriers under nationwide licenses.

Figure 9.2: Indonesian cellular coverage. Wide swaths of sparsely populated parts of the country lack any cellular coverage, which includes at least 10 million people. The red star on the right marks the location of the Papua CCN.

claims to cover "over 95%" of the population as of 2013 [125], meaning close to 10 million people live outside of coverage in Indonesia alone. In contrast, the GSMA suggests [126] this number could be as high as 90 million. The number of people living outside of coverage (and hence in areas with ample GSM whitespace) could exceed a billion in developing countries alone.

Exclusive licensing of GSM spectrum creates significant amounts of unused spectrum. Regulating spectrum in rural areas in the same way as urban areas inflicts a significant social cost: although low potential revenue makes it difficult for incumbent carriers to justify providing service in remote areas, exclusive license agreements prevent any others from offering service. Licenses have traditionally been offered in this manner because there was no local competition for the rural spectrum and it was easier for carriers to plan their networks assuming an exclusive license. We recognize the latter reason as valid, but the rise of CCNs puts the former out of date.

9.4.2 Spectrum Sharing in GSM Whitespaces

Our proposal to resolve this disconnect between spectrum licensing and rural service is simple: allow CCNs to utilize spectrum available in GSM whitespaces. Although we can draw some lessons from work on TVWS, the opportunities presented by GSM whitespaces have fundamental

differences. Most importantly, our usage scenario is far simpler than those envisioned for TVWS. Our proposal aims to broaden access to basic communications services, not to maximize spectrum utilization. We are only interested in enabling a single type of service in the whitespace, GSM cellular service, and this service has well-defined and minimal spectrum requirements (each channel is 200kHz wide). We are also primarily concerned with operation in rural areas with ample available spectrum. Finally, the economics of CCNs suggest that few secondary operators will coexist in the same area at the same time; we stress again that the localities CCNs are designed to serve are *unprofitable* for traditional telcos. This constrained design space simplifies our task.

Our goals for GSM whitespace spectrum sharing are:

1. *Safety.* Secondary operators should be able to provide cellular service in unused spectrum in standard GSM bands without interfering with primaries or other secondary operators.

2. *Independence.* Primary operators should have no new burdens restricting their usage, and should not need to cooperate with (or be aware of) secondary operators. Similarly, secondaries should not require special permission from or coordination with a primary.

3. *Verifiability.* Regulators and primaries should have visibility into what spectrum secondaries are using, and they should be able to verify that secondaries are actually providing service.

4. *Spectrum flexibility.* Secondary users should not be able to claim that use of any particular channel is necessary for their operation.[5]

5. *Backwards compatibility.* Existing, unmodified GSM phones should work with secondaries' networks.

We can achieve *safety* and *independence* by demonstrating a robust and reliable mechanism for detecting spectrum usage of other nearby operators, both primary and secondary. By reporting spectrum utilization measurements and usage by subscribers to a regulatory database, secondaries can provide *verifiable* rural coverage. *Spectrum flexibility* comes from ensuring secondaries have an actively and often exercised mechanism for frequently changing their broadcast channel without compromising their ability to provide service. By only leveraging existing mechanisms in the GSM specification, we can do all of this while maintaining *backwards compatibility*.

---

[5]This idea was advanced in a public conversation by John Chapin during the 2012 ISART workshop in Boulder, CO.

9.5 Nomadic GSM

The linchpin of our proposal is the feasibility of implementing a GSM base station that can achieve our goals for sharing spectrum in GSM whitespaces; this is Nomadic GSM (Fig. 9.1). NGSM is able to:

- quickly detect when it may be causing interference to a primary or another secondary operator (*safety, independence*);
- rapidly and frequently adjust its frequency usage to avoid causing interference (*spectrum flexibility*);
- accurately report its own frequency usage, as well as the frequency usage of other users in its area, to a regulatory database (*safety, verifiability*);
- and achieve the above without requiring modifications to existing client devices or significant interaction with existing license holders (*backwards compatibility*).

In this section, we describe the mechanisms by which NGSM meets these goals. We discuss the first three points in turn while continuously addressing the fourth.

9.5.1 Interference Detection

A key issue for dynamic spectrum sharing schemes that rely on sensing is the hidden node problem [127]. By definition, interference occurs at a receiver, so two transmitters may be interfering with each other even if they are unable to detect each other's transmissions by sensing the medium.

One solution to this problem that has been proposed for TVWS is a regulatory database of frequency usage. A similar database-driven approach to spectrum sharing also fits GSM whitespace. By their nature, GSM base stations will be connected to the Internet in order to provide service to their users; a local-only GSM network is only useful in limited cases. For example, in the Papua network roughly 66% of traffic is outbound [5]. We can report frequency usage and information on unused channels in the BTS's area to a database using this Internet connection. We assume secondary GSM operators will be willing to accept new regulatory requirements, such as registering their spectrum usage with a regulatory database. However, it is impractical (and contrary to our goals) to assume incumbent operators will accurately register their systems to a database; in effect, they will not be cooperating with secondary operators. We need a system to enable non-cooperative base stations to coexist with cooperative ones; this is a form of coexistence-based spectrum sharing [128].

NGSM leverages part of the GSM standard to overcome this challenge [129]. Every GSM BTS operates on one or more channels, known as ARFCNs (Absolute Radio Frequency Channel Number); because GSM employs frequency-division duplexing, an ARFCN specifies a particular pair of frequencies used for downlink (from the BTS to phones) and uplink (from phones to the BTS). In order to support handover of a phone between cells, base stations provide a list of frequencies for

up to six "neighbor" cells (the "neighbor list") to phones that are camped to (i.e., associated with) the base station. Since BTSs initiate handover, phones regularly scan each of these frequencies and report back the received signal strength (RSSI) for each, along with one for the current base station. The report also contains network and base station identification codes for each active ARFCN discovered.

By intelligently selecting the neighbor list at the BTS, NGSM can induce phones to report usage on frequencies of our choosing, without any modifications to the phones. Suppose we wanted to monitor whether ARFCN 20 is in use. NGSM would add this ARFCN to its neighbor list and then wait for measurement reports from handsets. If ARFCN 20 were not in use, handsets would report back as such. However, if another provider was actively using that band handsets would detect the other signal and inform our base station of its use. Importantly, this approach solves the hidden node problem by measuring interference at handsets, rather than at the BTS. However, all new logic required by NGSM is implemented at the BTS, ensuring backwards-compatibility with existing handsets. While conceptually similar to sharing spectrum sensing results as proposed in CORVUS [130], backwards compatibility with unmodified devices sets NGSM apart.

Monitoring the BTS's current ARFCN is slightly more complicated. Measurement reports are ambiguous in this case: if a handset reports a high RSSI for our ARFCN, it's impossible to know if that reading is due to the handset being near our tower or because we are interfering with another tower. Fortunately, there is a simple solution: configure our base station to use two or more ARFCNs simultaneously, rather than one. This is a common and well-supported configuration for GSM base stations, since a cell's capacity is directly related to the number of ARFCNs it supports.

NGSM handles this case as follows. First, we ensure that the neighbor list transmitted by the BTS contains on each of its ARFCNs contains both of the BTS's ARFCNs. Next, we *alternate* between each ARFCN, turning one completely off. Because the phones continue to receive both ARFCNs in their neighbor list, however, the BTS continues to receive measurement reports for both ARFCNs. If a primary user operates on the same ARFCN as one of our two ARFCNs, phones will continue to report the ARFCN is in use, even during periods when we have turned that ARFCN off, allowing us to detect which of our ARFCNs are no longer safe for use. The faster the rate at which we switch ARFCNs, the sooner we are able to detect potential interference.

Finally, we note that we can set the threshold for considering a channel occupied quite low since (1) switching to another frequency is easy to do and (2) there are likely many GSM channels available. Note that this technique can work with any number of ARFCNs per BTS, not just two, by always leaving one ARFCN off.

9.5.2 Changing Frequencies

The secondary's BTS changes its frequency use in three cases. First, to avoid causing interference: once a BTS detects that it may be causing interference, whether via measurement reports from handsets or the regulatory database, it needs to be able to quickly modify its frequency usage.

Second, a secondary needs to cycle through different frequencies on a regular basis. Doing so prevents secondary operators from claiming a particular frequency is essential for their operation, thus protecting primaries from spectrum squatting (Section 9.6.1). Finally, the BTS must switch between two channels during regular operation in order to detect interference on its own channels. The final two cases differ in timescale: while changing frequencies once per day may be sufficient for the former, in the latter case we want to be able to switch between channels quickly, on the order of minutes or even seconds.

What mechanism should we use to change channels? A naive solution would be to simply change the ARFCN on which the secondary's BTS operates. From the perspective of a phone, this is equivalent to shutting off the BTS on the old ARFCN and bringing up a new BTS on a different ARFCN. However, this approach has a serious downside: phones will have to re-associate with the BTS after each channel switch, causing downtime for users (phones take up to two minutes to reassociate [121]). Active calls would also be disrupted during an ARFCN switch. Given one of our primary design criteria is compatibility with existing, unmodified handsets, there is a tension between frequency agility and system usability.

We can address this concern in part by only cycling frequencies while the BTS is not being used (i.e., no active calls, SMS or data transfers). Rural cellular networks tend to be lightly utilized, especially during off-peak hours [121]. By cycling frequencies only when the BTS is not in use, we can avoid interrupting ongoing calls while minimizing perceived downtime to users. However, we don't actually reduce the amount of time that the BTS remains out of service.

We can take this one step further by leveraging the GSM handover mechanism and the fact that our BTS operates as two cells (i.e., already operates on two ARFCNs for the purposes of detecting interference on our own ARFCNs, as described earlier). Handover is designed to move a handset between cells of GSM network during a call and is instigated by the network infrastructure (the base station controller in a traditional GSM network) when call quality degrades. In this model, once the BTS decides to change one of its two ARFCNs, it first initiates handover for all phones camped to that ARFCN and moves them to the other ARFCN. Once all phones have camped to the new ARFCN, the BTS can safely turn off that channel or tune it to a new ARFCN. Phones will experience no downtime; even in-progress calls are not interrupted.

Importantly, GSM handover is universally adopted and widely used functionality for all GSM networks globally and as a result is widely implemented and tested in client devices. For example, handover allows users to make uninterrupted calls while in a moving vehicle. While our technique for determining when to perform handover is novel, the mechanism by which we would move clients from one frequency to another is completely standard.

9.5.3 Policing and Reporting Usage

We've described mechanisms for detecting interference by leveraging reports from phones and changing channels frequently without significantly impacting users. We now turn to reporting

usage and policing spectrum use. As discussed earlier, we assume that all secondaries' BTSs will have Internet access; this is reasonable given such access is essential to provide service to the public telephone network. Given this, these systems have two unique capabilities. First, these BTS units measure actual spectrum usage in their service area. Measurement reports gathered by phones can be used to determine ground-truth regarding spectrum usage in an area. This applies to both the spectrum the CCN gathering the reports is using as well as others in the area, enabling secondary users to "police" their area and report the existence of nonconforming operators. Secondly, CCNs know actual aggregate usage statistics about their users, such as number of calls or SMS served per day. Reporting both of these measurements to a database would give regulators insight into the scale and nature of rural service, and provide an effective mechanism for policing compliance with regulations on usage of GSM whitespaces. Incumbent licenseholders can also benefit from this data by using it to plan their network expansion into rural markets or for determining what portions of their spectrum are being used where by CCNs to obtain credit towards fulfilling USOs (Section 9.6.1).

Unlike the TV whitespaces (and others), we believe that the core role of a GSM whitespace database is to enable reporting, rather than to guarantee noninterference by appropriately herding devices (in frequency). This is important because it means that the GSM database does not require any action on the part of the incumbent before systems can *safely* begin using the GSM whitespaces. However, regulators can also respond to actual interference events by using the database to rapidly direct CCNs away from the frequencies on which interference is being perceived (a "frequency kill switch" of sorts).

9.6 Opportunities, Risks, and Incentives

9.6.1 Already Licensed Carriers

Although GSM frequency bands are heavily utilized in urban areas with high subscriber densities, spectrum is plentiful in unserved rural areas. We argue that sharing this rural spectrum imposes little if any cost to incumbents. For example, the Papuan CCN [5] operates in a frequency that has been licensed to Telkomsel, the largest Indonesian carrier. The network is serving a village four hours by car from the nearest place with cellular coverage. Although not legal, strictly speaking, the Papua CCN is isolated and does not impact the licenseholder's operations. It's even plausible that Telkomsel could provide service concurrently with the CCN if it decided to serve the same village; due to low subscriber density, the CCN is able to effectively serve its community with two ARFCNs (0.5MHz), under 7% of Telkomsel's GSM900 license allocation. That leaves most spectrum available to Telkomsel if it ever decides to serve the same area, even though their spectrum needs would be similar to that of the CCN.

Carriers need more than just assurance that sharing their spectrum won't impose any direct costs. We consider how to mitigate some potential risks and outline benefits spectrum sharing provides incumbent carriers.

9.6.1.1 Benefits and Incentives

GSM whitespaces offer several potential benefits for incumbent operators.

- **Fulfilling universal service obligations.** Sharing spectrum with CCNs could serve to fulfill a carrier's universal service obligations. Whether service is provided by a carrier or a CCN is functionally similar: rural customers receive access to communications service in both cases. Allowing carriers to take "credit" for CCNs operating in their spectrum for the purposes of demonstrating providing rural service would provide a strong incentive for carriers to support policies that enable spectrum sharing in GSM whitespaces. Requiring secondary users to report their spectrum usage and subscriber activity to a regulatory database (per our goal of verifiability) enables a simple way for carriers and regulators to determine such credit in a trustworthy way. Carriers would be able to automatically generate reports from the regulatory database to learn what areas and how many people CCNs that use the carrier's spectrum are serving. Since these reports come from CCNs, which are unlikely to have formal arrangements with carriers, "faking" this data will be difficult for carriers without directly supporting CCNs or providing service themselves. We envision the database being public, allowing civil society to call out untruthful providers.

- **Opening up new rural markets.** By their nature, CCNs open up new markets for cellular service in rural areas. These markets start small, but grow as the community and infrastructure expand. The CCN's presence encourages local investment in cellular phones and businesses to adopt the technology to improve their processes. Eventually, these markets may become economically viable for incumbents, and the CCN's presence has prepared the community for their arrival.

  Incumbent carriers could take advantage of this progression in more immediate ways as well. One example could be entering into a partnership with the local CCN where the carrier captures the CCN's customers as its own when these users travel into the carrier's network coverage. This approach preserves the autonomy of the independent CCN operator and has low overhead for the carrier while providing a channel for the carrier to acquire new, otherwise hard-to-reach, customers. When the incumbent eventually enters the rural market, those customers are immediately available.

9.6.1.2 Mitigating Risks

At the same time, spectrum sharing carries significant potential risks, the most significant of which is "spectrum squatting". NGSM's spectrum flexibility mitigates this to an extent, but GSM whitespaces offer inherent protections for primary users as well.

The "Grandparent Problem". The "grandparent problem" is a potential risk that carriers face when they allow another entity to provide service in spectrum they own. If the carrier ever wants to reclaim that spectrum from the other entity (e.g., once the agreement expires), customers of the

secondary entity may lose service and be upset at the carrier for being the "cause" of their service disruption. If those customers are a politically important constituency, such as a grandparent who is no longer able to communicate with grandchildren, the carrier may find itself in the crosshairs of negative public opinion and under pressure from policymakers and regulators to continue allowing the secondary entity to provide service in their frequency band. This isn't a concern with GSM whitespaces. Spectrum flexibility ensures secondaries can easily switch frequencies and continue providing service to their users even if a few GSM primaries decide to put a portion of their spectrum to use. Low population density means rural markets have minimal spectrum requirements, providing plenty of room for secondaries to coexist with primaries. Moreover, since all users of the band would be using GSM technology, the customers of secondaries could easily switch to the new provider by changing a SIM card.

Avoiding enabling new competitors. Another significant concern carriers are likely to have is that by sharing spectrum with CCNs they are enabling new competitors. At a high level, CCNs are not competing with incumbent carriers in the most significant markets, urban areas. There could be situations where a CCN and an incumbent carrier try to serve the same area. Competitively, this could resolve in two ways. First, the CCN could have intrinsically lower costs of operation than an incumbent carrier. If this is the case, the incumbent should simply use the CCN as a roaming partner, an established model in the cellular business, that would allow the carrier to receive the network connectivity they need at a lower price. On the other hand, if the CCN has higher intrinsic costs than the incumbent, the CCN will be unable to compete when the incumbent begins providing service, and competition will force them out of the market. Traditional telcos will almost certainly be able to undercut CCNs in the most lucrative markets, urban and suburban areas, protecting their own business.

Finally, because NGSM checks with a database, regulators have control over where secondary users can operate. Spectrum sharing could be explicitly disallowed in "non rural" areas. This level of control and oversight should further mitigate competitive concerns from primary users.

9.6.2   Other Stakeholders

The two other major stakeholders with an interest in the regulation of GSM whitespaces are community cellular network operators and regulators.

9.6.2.1   Community Cellular Operators

Community cellular networks have a range of options for how they deal with spectrum licensing besides GSM whitespaces. Small, isolated networks may choose to operate as pirates under the expectation that they are beyond the reach (or interest) of regulators; the Papuan CCN took this route. The Oaxacan CCN operates under a two-year experimental license, but while this enables legal operation in the short term, they have no guarantee their license will be extended in the future. Their license also restricts them from making profit with their network, which hinders their efforts to provide sustainable and reliable service to their users. While we are unaware of any

examples, a CCN could also partner with a carrier and operate under their license or simply buy a commercial spectrum license outright. As a small entity, obtaining a carrier partnership or a standard commercial license is out of the reach for most CCNs. Without proper incentives and risk mitigation factors like those outlined in Section 9.6.1, carriers have little reason to cooperate with CCNs; even finding an audience with a carrier to discuss a partnership is challenging for small entrepreneurs in rural areas. Commercial licenses can cost millions of dollars, well beyond the budget of a CCN.

Our proposal for spectrum sharing in GSM whitespaces represents a middle ground for CCNs with a range of attractive properties. Unlike pirate operation, it would allow CCNs to operate "above the board", reducing risk to both the operator and the long-term sustainability of a CCN while at the same time maintaining the flexibility and independence of pirate operation. Experimental licenses have similar drawbacks to pirate operation—while legal, CCNs still face the risk of shutdown should their temporary license not be renewed. Explicitly supporting commercial operation indefinitely is necessary for incentivizing local entrepreneurs to operate CCNs so they can confidently plan to recoup their initial investments.

9.6.2.2 Regulators

First and foremost, by enabling the operation of CCNs regulators can expand access to service in rural areas. This fits the social mission of regulatory bodies to ensure that telecommunications access is available to their nation's citizens and that spectrum, a vital public resource, is used equitably and efficiently.

GSM whitespaces give regulators a fundamentally new tool by which to achieve this mission. Today, decisions to build out rural infrastructure rest solely with incumbent license holders, and regulators are only able to indirectly influence these decisions through mechanisms like universal service obligations. Spectrum allocated for GSM networks is poorly utilized in rural areas: beyond the coverage area of existing cellular carriers, exclusively licensed spectrum is simply not used. Telcos are inherently disinclined from serving rural areas due to high costs of service and low revenue potential (due to low subscriber density), so this spectrum lies fallow. Existing mechanisms for incentivizing carriers to serve rural areas, such as universal service obligations, have high overhead and are fraught with political baggage.

In contrast, light regulation of GSM whitespaces, as we propose, allows local entrepreneurs to operate small-scale community cellular networks without requiring regulators to engage in expensive oversight of these operations. Moreover, since pirate operation is currently among the most attractive "licensing" options for CCN operators, providing a low-touch mechanism for these operators to register and regularly report on their spectrum usage gives regulators control over an emerging trend: rural, isolated communities have a strong demand for cellular communications, and it's foolish to hope that existing penalties will prevent them from building their own infrastructure if they are able.

9.7 Evaluation

In Section 9.5, we claimed that NGSM achieves all five goals for spectrum sharing in GSM whitespaces (Section 9.4). In this section, we justify that claim by implementing and evaluating NGSM. In addition to testing in a controlled environment, we also deployed NGSM into a real-world, operational CCN in Papua, Indonesia.

9.7.1 Implementation

We implemented NGSM as a software control layer based on OpenBTS [115], which uses a flexible software-defined radio and a commodity PC to implement a GSM base station. We support both dual-ARFCN, with one software radio per ARFCN. We also support single-ARFCN operation, with the limitation that a single-ARFCN BTS is unable to detect other users of its own channel. Our implementation of NGSM monitors all control traffic between OpenBTS and phones, including measurement reports, and configures the ARFCN and neighbor list used by OpenBTS as appropriate.

In particular, we randomly select 5 ARFCNs for phones to scan every N hours. All ARFCNs are initially considered "unsafe", save the ones initially in use. In our implementation, we randomly pick initial ARFCNs, though we expect a wider deployment might also be able to use a database query to pick the initial ARFCN. Once we receive measurement reports indicating an ARFCN is not in use, we consider the ARFCN "safe". An ARFCN remains safe as long as we receive no more than K reports indicating an RSSI[6] on that ARFCN exceeding R. Once these thresholds are exceeded, the ARFCN is demoted to being unsafe; in our implementation, once an ARFCN was demoted to being unsafe it remained so for at least $4N$ hours, at which point it could be scanned again and marked safe. ARFCNs that had been used by the BTS were similarly marked as unsafe once they had been used to ensure the BTS would use a different ARFCN each time it switched channels.

In our implementation, we chose a cycle length N of four hours to allow the BTS to scan a quarter of the GSM900 band every day, though this was itself chosen arbitrarily. We set $K = R = 0$ to be as conservative as possible in detecting other users of the band. These values are essentially the sensing threshold for the system; we leave full consideration of how to set these values appropriately to future work.

Given our scanning results, we select a pair of safe ARFCNs for the BTS to use. We alternate use of each ARFCN every $T = 90$ seconds by adjusting the TX attenuation on each software radio between 0 and 100 dB. One ARFCN operates without attenuation, while other operates at high attenuation. To change which ARFCN is active, we gradually increase the attenuation of the current ARFCN while reducing the attenuation on the inactive ARFCN to 0 dB. As the attenuation

---

[6]RSSI is defined in GSM 04.08; specifically, we use the RXLEV-NCELL value, defined from 0 (-110 dBm) through 63 (> -47 dBm).

Variable	Value	Purpose
R	0	For ARFCNs *not* in use by the secondary, the RXLEV threshold for determining whether an ARFCN is in use.
K	0	Number of reports exceeding R needed to declare an ARFCN unsafe.
T	90	Number of seconds between ARFCN switches for the two ARFCNs in use by the BTS.

Table 9.2: Parameters for NGSM.

on the former ARFCN increases, phones automatically handover to the latter; from the phone's perspective, we've simulated moving away from one cell and towards another.

If any of our in-use ARFCNs become unsafe, we immediately cease use of that ARFCN and switch to different safe one. This doesn't result in any service disruption since we can only detect use of one of our own ARFCNs while that ARFCN is fully attenuated. To remain spectrum flexible, we select a new pair of ARFCNs to use every night. Additionally, if the BTS is restarted for any extrinsic reason (such as a power failure), we also pick a new pair upon restarting.

We deployed our implementation of NGSM in both a controlled environment in our lab for testing, as well as the Papua CCN referred to earlier. We began running NGSM in Papua on October 14, 2013. We ran the system for testing for one week before we began collecting data, which we did from October 22 through November 1, 2013.[7]

As stated before, we used two independent software radios for dual-ARFCN operation. This was solely due to time constraints, and we stress that there is no fundamental reason our approach requires two radios: although implementing support for multiple channels on a single radio would require more engineering effort, it's within the hardware capabilities of existing software radios. We are currently in the process of adding single-radio operation to our prototype and see no major engineering obstacles ahead of us. An unfortunate consequence of this is that due to hardware limitations in the Papua CCN, we were not able to deploy the full dual-ARFCN version of NGSM and were constrained to a single ARFCN. This also means we were unable to detect potential interference on the ARFCN currently in use by this BTS. We were, however, still able to collect measurement reports and scan the band for other users on that network, and we still changed the ARFCN the BTS used at least once per day.

---

[7]NGSM is still running today; we would plan to include a longer dataset in a final camera-ready submission.
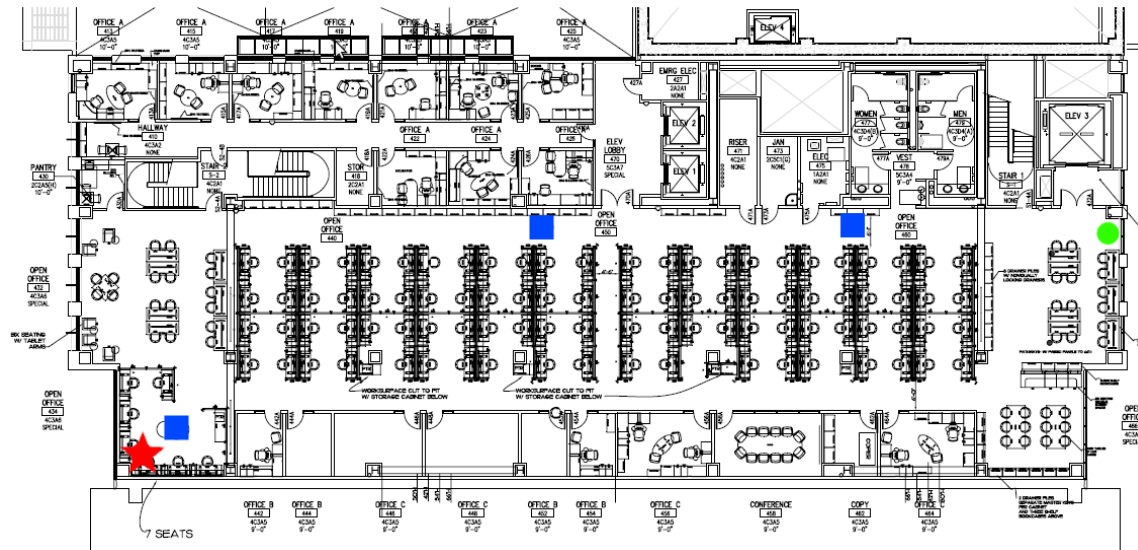
Figure 9.3: Floor layout showing the primary (green circle, right), CCN (red star, left), and three cellphones (blue squares). The monitoring site was co-located with the middle phone.

Finally, although all of our experiments were conducted in the GSM900 band, our implementation could be easily modified to operate on any other band used for GSM. Our implementation of NGSM is available at http://github.com/shaddi/gsmws.

9.7.2   Coexistence

The ability to detect and respond to potential interference is a crucial requirement for NGSM. To demonstrate our ability to do this, we set up two BTS units. The first was a standard, unmodified GSM BTS, configured to simulate a "primary" user's BTS broadcasting on a single ARFCN. The second ran NGSM with two ARFCNs as outlined in Section 9.5, simulating a BTS run by a "secondary" user, a CCN. We also configured three phones as customers of the secondary BTS. The primary BTS used the same ARFCN as the secondary, but its other parameters (such as network ID) were distinct from those of the secondary: to phones, the secondary and primary BTS units appear to belong to two completely separate network operators. Each BTS was configured to transmit at 100mW per ARFCN. Fig. 9.3 shows the layout of the two BTS units and the 3 phones on in our testing environment, a single floor of an office building. Additionally, we placed two spectrum analyzers next to the middle phone, tuned to both the downlink and uplink bands used by the three BTS in this experiment.

We started NGSM on the secondary BTS and began alternating between its two ARFCNs. We waited for the three phones to camp to the secondary BTS and begin transmitting measurement reports, simulating a CCN operating in a steady state (i.e., with phones camped to its tower, but not necessarily in use). One phone, the middle one, had an ongoing call to the BTS. We then turned on the primary BTS to simulate the appearance of a primary in the vicinity of the secondary.
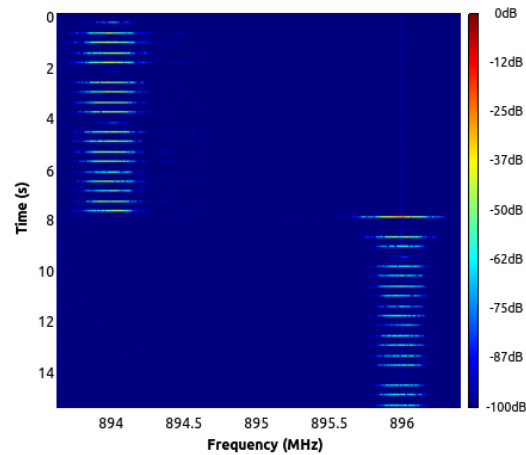
Figure 9.4: Spectrum usage of a handset during a call in the uplink band (i.e., from the phone to the BTS). The phone switches ARFCNs without interrupting the ongoing call.
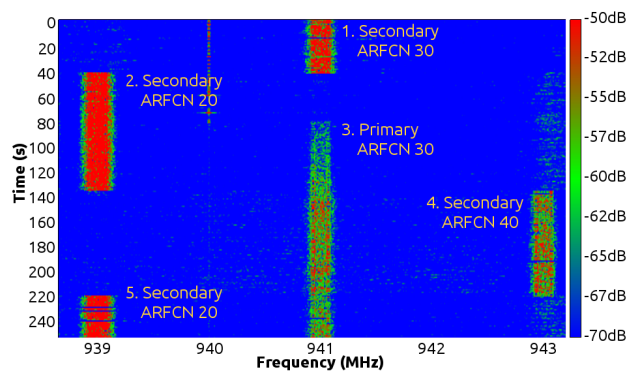
Figure 9.5: Spectrum usage during simulated arrival of a primary into the secondary's service area. Initially, the secondary is alternating between using ARFCNs 20 and 30 (939MHz and 941MHz, respectively). When a primary BTS appears on ARFCN 30, the secondary detects its presence and switches to using ARFCN 40 instead (943MHz). Thereafter, the secondary alternates between ARFCNs 20 and 40.

Fig. 9.4 shows the results of this test in the uplink band. This figure shows the usage of the phone on a call while the BTS alternates between ARFCNs. As expected, the phone completes handover successfully and the call continues without interruption. In Fig. 9.5, we see the spectrum usage on the downlink band during a simulated appearance of a primary user. Initially, the secondary BTS is alternating between ARFCNs 20 and 30. The primary appears on ARFCN 30 halfway through the experiment. Detecting this, the secondary BTS picks a new, unused ARFCN to use instead of ARFCN 30 (in this case, ARFCN 40). The secondary then begins alternating between ARFCNs 20 and 40, while the primary continues operation on ARFCN 30 without interference.
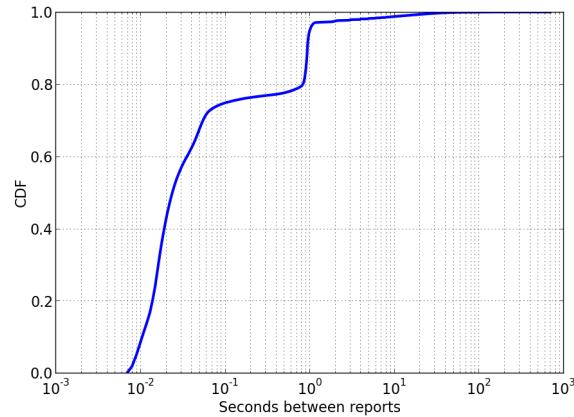
Figure 9.6: CDF of measurement report interarrival times received during daytime hours (7AM–12AM). Night time hours are not included since the CCN was typically powered off at night. $\mu = 0.64$, $\sigma = 4.02$. Note logarithmic scale.

Although this test may seem simple, it demonstrates a few important points. First, using handsets to detect interference is possible, without making any modifications to existing GSM infrastructure like phones or network infrastructure of existing carriers. Second, even when serving a handful of phones, the secondary BTS is able to detect interference and adapt its usage accordingly within a matter of minutes. Finally, it shows that secondaries can rapidly change channels upon detecting interference with minimal service disruption.

9.7.3 Measurement Reports

The time a secondary takes to detect a primary is inversely proportional to the frequency of measurement reports. Although phones constantly send measurement reports when in active use (e.g., during a call or when receiving an SMS), they only do so once every six minutes otherwise. Thus, measurement report frequency is directly related to the number of users a CCN has and how active those users are. In other words, we have measurements exactly when we need them: the potential for harmful interference is higher from an active network with many handsets, which in turn will have more frequent measurements.

We evaluate this empirically with our deployment in the Papua CCN, which has over 200 subscribers, more than 70 of which are active each day. Although we could not directly evaluate speed of detection—the operators of the CCN don't have the equipment necessary to replicate the previous experiment, namely a second BTS unit with which to simulate the appearance of a primary—we can evaluate the frequency of measurement reports the Papua CCN BTS receives. To do this, we logged every measurement report received by the Papua CCN from October 22 through November 1. A small, consistent interarrival time is valuable because it allows us to put bounds on how quickly a secondary can detect the primary.
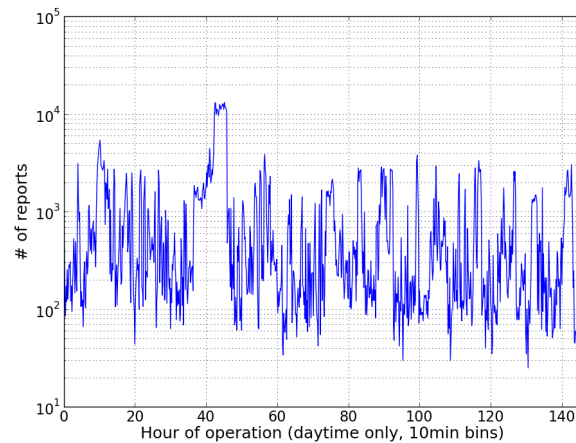
Figure 9.7: Time series of number of measurement report received per hour during daytime hours (7AM–12AM). Night time hours are not included since the CCN was typically powered off at night. Note logarithmic scale.

We received approximately 846,000 measurement reports during our 10 days of operation. Of these, we only consider those received during "daylight" hours between 7AM through 12AM; because the CCN typically turns off at night, outside these hours we do not receive measurement reports. This removes about 12,000 reports (1.5%) from our analysis. With this in mind, Fig. 9.6 shows the distribution of interarrival times between measurement reports received over the course of our two week deployment. The maximum spacing between measurement reports seen during daytime hours was 11.7 minutes; the 99.9th percentile was 56 seconds. This result suggests real-world CCNs will enjoy faster detection times than we observed in our lab.

Finally, measurement reports arrive consistently while the BTS is operational. Fig. 9.7 shows the number of measurement reports received during the operational hours of the CCN, collected in 10 minute bins. The minimum number of of reports received in any 10 minute window was 25; the median was 300 (obviously, when the BTS was off or not in use at night there were periods in which no reports were received). Combined with the previous distribution of report interarrivals, this demonstrates that we can rely on receiving regular measurement reports, placing an upper bound on the time to detect a primary's BTS on the order of minutes during normal usage.

9.7.4 Deployment

NGSM operated as expected when we deployed it onto the Papua CCN. Fig. 9.8 shows the measurement results from the deployment. In this figure, the in-use ARFCN is blue, while ARFCNs considered "safe" or "unsafe" are colored green and red, respectively. During the experimental period, the operator's primary source of electric power failed, causing several prolonged outages. Nevertheless, the CCN switched ARFCNs frequently, as designed. We were also able to verify through measurement reports that many ARFCNs were available for use around the Papua CCN, even when using the most sensitive detection thresholds ($K = R = 0$).

Metrics	Pre-NGSM		With NGSM	
	Mean	Deviation	Mean	Deviation
Calls	95.1	60.5	138.4	65.4
SMS	656.5	113.5	633.7	147.6
Active users	62.8	4.5	62.4	7.4

Table 9.3: Usage per day in the Papua CCN, before implementing NGSM (09-09-2013 to 10-09-2013) and after implementing NGSM (10-10-2013 to 11-08-2013). Deploying NGSM did not significantly impact usage.

Despite these frequent channel changes, we observed no negative impact on network usage after deploying our system. Table 9.3 shows the distribution of network usage metrics per day before and after the deployment. We only consider calls and SMS initiated by users of the CCN; incoming communication is not included in these statistics. Active users refers to the number of subscribers who initiated either a call or an SMS that day. A number of factors—the aforementioned power failures, natural variation in usage (e.g., people travelling), etc.–preclude statistical testing, but we observe that usage remains roughly the same in terms of active users and SMS, and actually increases for number of calls. This is not a surprising result—we designed our system to only change ARFCN during periods of little or no activity to avoid impact on usage. Nevertheless, it shows that a CCN can operate effectively even when it changes its ARFCN relatively frequently.

We expected the GSM900 band to be completely unused around the Papua CCN, as the network operators informed us that the nearest cellular service was almost 30 kilometers away, beyond several mountainous ridges (the Papua CCN itself is located in a small valley). In general, we found this to be the case, but there were a few interesting exceptions.

Somewhat surprisingly, we detected a usage of several ARFCNs during our deployment, many of which were licensed to carriers who do not provide any service in the Papuan highlands. For example, on October 26 the BTS received 19 reports over a 2 hour period indicating ARFCN 50 was in use. The Papua CCN's BTS performed as designed and did not use those ARFCNs going forward. Unfortunately we have no way of knowing what may have caused these reports, nor can we necessarily discount the possibility they were simply spurious reports. However, this highlights a crucial point: spectrum regulations might already be flouted in rural areas, and regulators (and licenseholders) have no way to detect these violations until they actually interfere with operations.
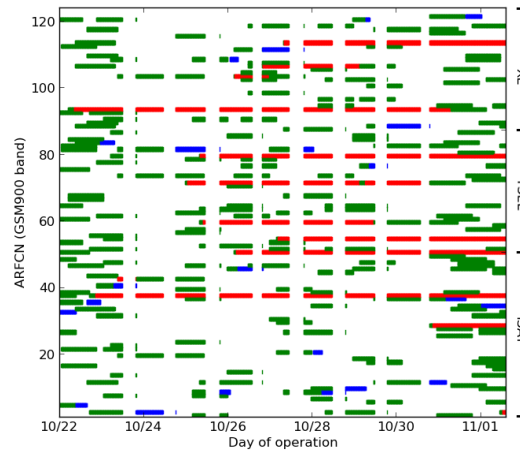
Figure 9.8: Spectrum usage and reports for the Papua CCN from Oct 22 through Nov 1. Blue represents the ARFCN in use by the CCN. Green represents an ARFCN that was scanned and considered "safe" to use, and red represents an ARFCN which was scanned but appeared to be in use, and was thus "unsafe". Unsafe ARFCNs could become safe ARFCNs once they had been scanned again and found to not be in use. Empty columns represent times when the CCN was out of service; the CCN operator's primary source of electric power failed during the evaluation period, causing prolonged nightly outages. Carrier allocations are shown on the right axis; the system jumps between all three.

9.8   Discussion

9.8.1   Market alternatives to GSM whitespaces

The obvious market-based alternatives to GSM whitespaces is to have the CCN operator enter into a contractual relationship with the license-holder of the spectrum (as a franchisee of sorts or an alternative local brand) or to engage in a transaction in secondary spectrum market to obtain local usage rights in that area. While conceptually tempting and simple, there are problems. The empirical evidence of such transactions not actually happening in the real world suggests that something must indeed be wrong with this approach.

9.8.1.1   Individually-negotiated contracts

The theory of Coasian bargaining says that the problem must be in transaction costs. Indeed, it is hard to imagine a carrier engaging a lawyer and engineers to travel to a remote area to negotiate and implement/verify a contract to split some profits that amount to a mere $1000 per month, even without factoring in the uncertain future squatting risk that they might feel they face from having the CCN operate solely in their spectrum. The NGSM approach to GSM whitespaces here eliminates that transaction cost and the uncertainty.

9.8.1.2 Standardized markets

In principle, a clean online secondary market could also eliminate some of these transaction costs. However, there are subtle issues here. First, what is being traded? There are two possibilities: spectrum (where CCNs pay spectrum license holders to be able to deploy their systems) or USOs (where CCNs bid to accept USOs in some area).

Second, once there are market transactions in an asset or liability, the asset/liability can be quantified in dollar terms. Hence by accounting principles, the entire asset or liability must be quantified in dollar terms on the firm's books. For spectrum, this could challenge the (speculative) high valuations that firms carry on spectrum to serve as collateral for loans, etc. For USOs, it would suddenly cause USO obligations to show up as dollar-valued liabilities as opposed to vague risks.

The net effect on the books of having such secondary market transactions is likely to be negative and so the managers of firms are disinclined to explore such markets/transactions. The NGSM approach to GSM whitespaces avoids having any dollar transactions by directly getting USO credit from the regulator for the actions of a third party.

9.8.2 Application to urban and marginal coverage areas.

Our GSM whitespace scheme is designed for rural areas. It assumes large amounts of unused spectrum and relies upon spatial separation to avoid interference. We do not believe it is a good fit for situations where spectrum is scarce and highly utilized. This is not a problem in our minds: existing spectrum allocation policy has proven to be adequate for ensuring widespread cellular coverage in urban areas.

Compared to areas completely beyond coverage, an even larger portion of the planet's population likely lives in areas with marginal cellular coverage: while they may be able to access a cellular network, coverage may be sporadic or otherwise spotty. These areas, like the completely unserved areas beyond them, are likely to have GSM whitespaces available, but the potential interactions between primary and secondary license holders will be more complex. We leave a full consideration of these issues for future work. In the meantime, the simpler case of spectrum sharing in completely unserved areas is tractable and should be considered by policymakers in the short term.

9.8.3 Distribution of clients

Our proposed interference detection mechanism relies on reports from handsets and thus is sensitive to their geographic distribution. In the degenerate case, all phones could be clustered in a small area and unable to detect other networks within range of our BTS site. This concern is mitigated partially by the fact that both primary and secondary users will be trying to serve the same people in a rural area, so the physical distribution of a secondary's user base is likely to be correlated with that of the primary. Additionally, the GSM standard provides a mechanism for obtaining geolocation information from phones, the radio resource location services protocol (RRLP). This

information could be used in conjunction with measurement reports to identify potential "blind spots" where the secondary is unable to detect interference. Beyond simply shutting down or requesting guidance from the regulator database, A CCN can take a number of creative actions upon detecting such a blind spot. For example, the CCN could automatically send an SMS to a user requesting them to wander over to the blind spot area, perhaps incentivizing them with free network credit.

9.9 Conclusion

Rural areas are fundamentally hard for traditional telcos to serve profitably, leaving hundreds of millions of people beyond the reach of existing cellular phone networks for structural reasons. While community cellular networks appear to offer substantial advantages for providing sustainable rural service without subsidies or external support, their growth is stymied by a lack of rights to spectrum. However, exclusive spectrum licensing has created large areas of GSM whitespaces, areas in which GSM spectrum is allocated to a carrier but not actually used, as is the case in many rural areas worldwide.

This spectrum need not be wasted: we believe it represents an opportunity to enable community cellular networks to provide service in rural areas. In this chapter, we've proposed Nomadic GSM, a spectrum sharing technique for GSM whitespaces to leverage this opportunity. NGSM uses a combination of a spectrum database and a novel distributed spectrum scanning technique, leveraging the reporting capability of mobile phones, to ensure rapid detection of potential interference. Our proposal allows CCNs to safely share spectrum with incumbent carriers without their explicit cooperation, while mitigating key concerns that licenseholders might have with sharing their spectrum. By reporting spectrum measurements to a database, it enables regulators to verify what spectrum is actually in use in an area so that carriers can receive USO credit.

NGSM is compatible with existing, unmodified GSM phones: we've demonstrated its feasibility with both a prototype implementation in our lab as well as a real-world deployment on an existing community cellular network in Papua, Indonesia. We've demonstrated that with 70 daily active users, we are able to receive a measurement report at worst every 11.7 minutes while the BTS was on, with a 99.9th percentile interarrival of 56 seconds.

The implications of our system are important. Rural communities will build their own community cellular networks in increasing numbers, many of which will be operating in GSM whitespace. The situation is akin to that of WiFi in countries that had not yet adopted policies allowing unlicensed spectrum use—strong demand compels community cellular network operators to flout regulations and operate illegally, outside the control of regulators and at risk of interfering with the operation of licenseholders and other CCNs. Unlike WiFi, however, CCNs are still in their infancy, and an enlightened regulatory approach towards them can allow countries to maximize their benefits for providing rural service while mitigating impact on other users of the GSM bands.

Our proposal has attractive properties for achieving this, and, importantly, is deployable today, requiring no changes to existing mobile phones, network infrastructure, or operational practices of incumbent network operators. As such we feel it represents a strong first step towards a comprehensive policy for enabling legal coexistence of community cellular networks. We suggest that regulators take the following steps:

- Legalize use of GSM whitespaces with requirements that CCNs using them regularly (a) move between unused frequencies and (b) use NGSM (or similar) to monitor local GSM frequency and avoid causing interference.
- Facilitate creation of a GSM whitespace reporting database.
- Give carriers USO credit for CCNs operating in their spectrum allocations using the database reports as evidence for such claims.

Part VI

Methods

One of the major contributions of this thesis is its *quantitative* approach to studying spectrum whitespaces. Although small-scale models involving a few incumbents are invaluable, it is impossible for them to tell the whole story. Even if all small-scale scenarios were well-understood, translating this knowledge to a nationwide scale is critical to understanding the true impact of regulatory decisions.

Due to the nature of the problem, true quantitative approaches require the use of a computer in order to efficiently process all of the data at hand. For example, determining the amount of whitespace in the United States involves integrating data for thousands of TV stations. Each "pixel" in the whitespace availablility map must check to see if its whitespace availability is affected by each of these TV stations, a computation that would be completely impractical to do by hand for $200 \times 300 = 60,000$ pixels for each of the 49 channels.

Although these computations are not incredibly complex, it can take substantial effort to find the data, import it, and process it. I firmly believe that whenever possible, code should be shared among researchers in order to advance the field as quickly as possible. To that end, I have published my research code online since my first year of graduate school.

For the first four years of my research career, I built a code base in Matlab that grew to suit my evolving research needs. This code base underlies most of the results in this thesis and is described in Chapter 10.

Although intended to be general-purpose, the Matlab code base is in many ways purpose-built and somewhat inflexible. As a result, I developed Whitespace Evaluation SofTware (WEST), a Python code base designed to be more flexible than the Matlab code base. In addition, the choice of Python over Matlab means that the software will be more accessible to the average user[8]. The design of WEST and a few applications are described in Chapter 11. WEST was also critical in the studies done in Chapter 6.

WEST was used to generate the results in Chapters 6 and 11. All other results were generated using the Matlab code base.

---

[8]Matlab licenses often cost thousands of dollars, rendering them inaccessible to most outside the academic community and large corporations.

Chapter 10

MATLAB code

Some or all of the work in this chapter appeared in the author's Master's thesis, [38].

In order to understand the whitespaces on a nationwide scale, we developed a toolkit with the ability to let us explore using real-world data. We collected data from three major sources and created an interface for each. This library allowed us to build more complex models to answer some of the following questions:

- How many channels are available to secondaries?

- What data rates might be achievable in the TV whitespaces?

- How does the choice of communication range affect secondaries?

- Are whitespaces the right choice? How do they compare to traditional channel reassignment?

- Are TV receivers inside the protected region actually protected?

- How can we design the rules to be agnostic of secondary deployment?

This section covers the methodology behind the creation of the data presented in much of this thesis. Limitations of this data and our use of it are discussed in Section 10.7.

It is worth noting that there were two major challenges in developing this project: speed and organization. There are many ways to answer these questions, but many are computationally infeasible. It is important to find fast implementations in order to keep execution time down as much as possible. Having a rapid prototype process is key to developing ideas. Further, there is so much data generated in the process of making the figures featured herein–and even that is only a fraction of what is generated before the final version–so it is extremely important to keep everything organized so the results are accurately reported.

Further information as well as the toolkit itself can be found online at [6].

10.1 Basics

The code to generate all data figures is written and executed using a basic installation of Matlab (i.e. no toolboxes are required).

To create map-level data, we use a discretized version of the continental United States. In particular, we discretize a Mercator projection of the US, as shown for the state of Wisconsin in Figure 10.1. We often refer to these points as "pixels." The area represented by a pixel depends on the resolution of the map.
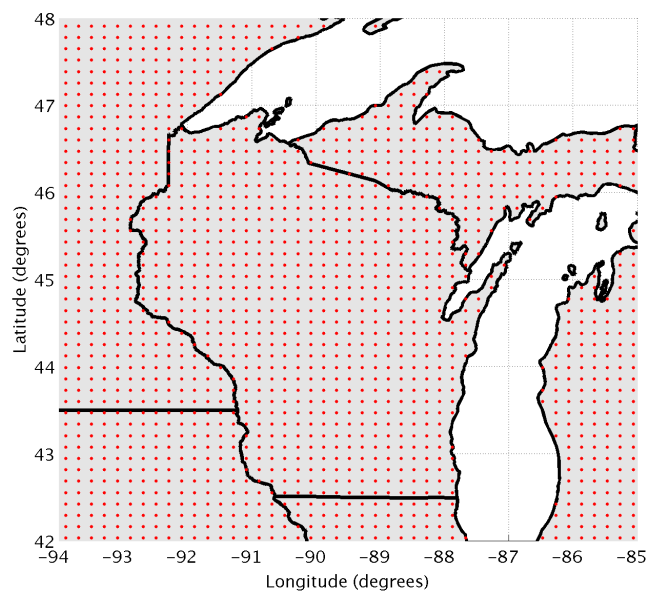
Figure 10.1: Zoom-in of discretized map

While this discretized version of the map provides valuable a good high-level insights, it fails to capture many local features. We will give one example of how to overcome this in Section 10.6.

10.2 External data

The simulations rely on three key pieces of external data, each of which will be discussed in turn:

- International Telecommunication Union (ITU) signal propagation model

- TV assignment data from the FCC's Consolidated Database System (CDBS)

- Population data from the US Census Bureau

Mishra, et al., were the first to develop a toolkit which incorporated these three elements [47, 48]. Our work builds on theirs and appears to have inspired similar work for Europe [31, 32]

10.2.1   ITU propagation model

The ITU propagation model uses transmitter height, power, and channel to determine the signal strength at any location[1]. This data is used for any plot which involves a protected radius, signal strength, or data rate. For example, all nationwide maps and CCDFs use this data. One can see instantiations of this propagation model in Figure 2.6 and online at [29]. The integration of the propagation model with Matlab code was done by Mubaraq Mishra.

10.2.2   TV assignment data

TV assignment data is used in the calculation of pollution levels, exclusion radii, number of TV channels potentially lost to aggregate interference, and safe power levels (as in Chapter 3.

The data used was downloaded from [28] on 16 September 2011. This website is a front-end for the FCC's CDBS, the resource used by TV Bands Database Administrators [131]. It includes geographic coordinates, transmit power, height above average terrain (HAAT), channel number, analog vs. digital, and much more for each TV assignment.

10.2.3   Population data

Population data is used to give per-person data rate estimates (e.g. maps) as well as to calculate per-person CCDFs.

Tract-level 2010 population data was obtained from the US Census Bureau [10]. The Census Bureau describes tracts as follows [11]:

> "Census tracts are small, relatively permanent statistical subdivisions of a county. Census tracts are delineated for most metropolitan areas (MA's) and other densely populated counties by local census statistical areas committees following Census Bureau guidelines (more than 3,000 census tracts have been established in 221 counties outside MA's). Six States (California, Connecticut, Delaware, Hawaii, New Jersey, and Rhode Island) and the District of Columbia are covered entirely by census tracts. Census tracts usually have between 2,500 and 8,000 persons and, when first delineated, are designed to be homogeneous with respect to population characteristics, economic status, and living conditions. Census tracts do not cross county boundaries. The spatial size of census tracts varies widely depending on the density of settlement."

Because of the way the census tracts are created, their size is directly correlated with their population density. For example, we can easily identify the urban areas of Madison and Milwaukee by the clusters of small tracts.

---

[1]Section 10.7 discusses why this is not strictly true.
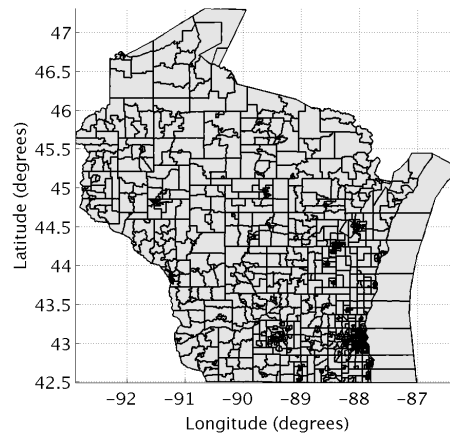
Figure 10.2: Census tracts in Wisconsin

Also note that some census tracts extend beyond the land border of the state and into the water (e.g. Lake Michigan).

In order to calculate the discretized version of the population, we first determine the population of the region represented by each pixel, shown in blue in Figure 10.3. Each intersecting census tract contributes a percentage of its population proportional to the percentage of its area included in this pixel. For example, a tract which has 25% of its area inside of a particular pixel will add only 25% of its population to the running total for the pixel. Since census tracts are designed to be roughly homogeneous [11], this is a reasonable assumption. Population density is calculated by dividing a pixel's population by the area it represents.
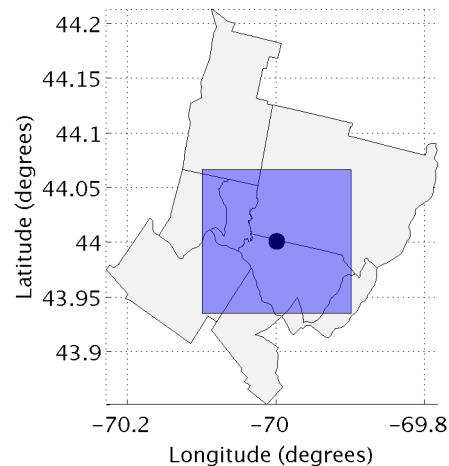
Figure 10.3: Per-pixel population computation illustration

Unfortunately, pixels are too large to sufficiently capture all of the variation in population. See Section 10.7.2 for further discussion of this limitation.

10.3 Calculating single-link rates

Figure 10.4 depicts the flow of information used to create a single-link rate map. Green blocks are input data, blue blocks represent intermediate pieces of data, and orange represents the final product.
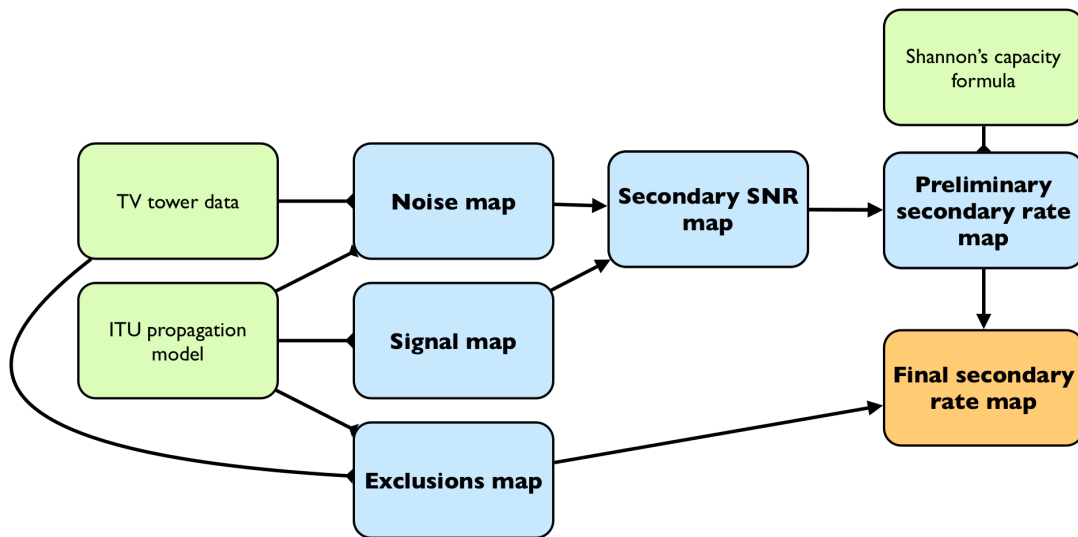
Figure 10.4: Flow of information used to create a rate map for the single-link case.

First, we compute the noise map, signal map, and exclusions map separately:

- A **noise map** represents the amount of TV-generated interference a whitespace device would experience. To create the noise maps, we use TV assignment data (location, transmit power, HAAT, and channel for each TV assignment) in conjunction with the propagation model to find the cumulative TV signal strength at each point on our map.

- The **exclusions map** is a binary map which contains information on where whitespace devices are allowed. We calculate the exclusions map, again using TV assignment data, propagation model, and the FCC's rules[2].

  - For digital TV assignments, we calculate using the F(50, 90) curves which means that the signal strength exceeds the threshold at 50% of locations 90% of the time.

---

[2]See Section 10.7 for caveats including HAAT calculation and propagation model for rule calculations.

– For analog TV assignments, we calculate using the F(50, 50) curves.

- The **secondary signal map** represents the received signal strength of a whitespace device at various locations. We create the signal map using the secondary's transmit power, HAAT, and channel to find its signal strength at its receiver. Note that in all but the $p = 2000$ case, the range is constant throughout the nation and thus the signal map is also constant. In the $p = 2000$ case, the variation in signal strength is due to the variation in ranges, which in turn are directly related to the local population density.

We divide the signal map by the noise map to yield a secondary SNR map, then use Shannon's capacity formula to calculate the rate map.

Finally, we apply exclusions to the capacity to get the final rate map. For (pixel, channel) pairs where whitespace operation is not allowed, the capacity is set to zero. Otherwise the original capacity value is retained. In this way, the exclusions map can be viewed as an exclusions *mask*. Note that we can apply the exclusions last since pixels do not interact in this example (which is true for all but the power-density model).

10.4   Calculating a cumulative distribution function (CDF)

The cumulative distribution functions (CDFs) we calculate are weighted CDFs, typically weighted by population but sometimes weighted by area. To calculate the CDF, we need two pieces of information: a weight map and a data map. The weight map assigns a weight to each corresponding point in the data map.

When the weight map is the population map, this can be viewed as giving each person a "vote." Once all of the votes are tallied, we create an empirical CDF of vote values. From this CDF we can easily calculate the empirical mean and median.

10.5   Computing MAC model rates

We use a medium access control (MAC) model to explore the effects of self-interference among secondaries. Our model allows secondaries to specify the size of the surrounding area which is free from other secondary transmissions, also called the *MAC exclusion area*. Outside of this exclusion area of radius r, we assume that the receiver is surrounded by neighbors as depicted in Figure 10.5. In our MAC scheme, any transmitter closer than r km (the *MAC exclusion radius*) from our receiver of interest must halt transmissions[3]. Thus we can consider a *capacity per area*, defined as the rate that this receiver achieves divided by its footprint, $\pi \cdot r^2$.

---

[3]One can imagine using some time-sharing protocol to ensure fairness among users, such as TDMA [18].

Finally, the value for r is chosen such that the capacity per area is maximized. Choosing a smaller r would reduce the size of the exclusion area but increase the interference levels, causing the capacity per area to shrink. A larger r would decrease the interference levels at the cost of enlarging the exclusion area, also causing the capacity per area to drop.
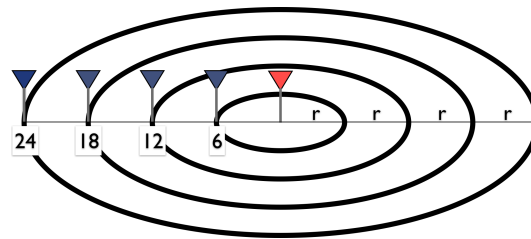
Figure 10.5: Secondary locations in the MAC model: the receiver of interest (in red) is surrounded by 6 neighbors at distance r, 12 neighbors at distance $2r$, and so on.

10.6 Computing power density levels and rate maps

This section describes the computations used in Chapter 3. As is shown in Figure 10.6, the steps can be broadly grouped into the categories "rule generation" and "rule evaluation." Because of this separation, we can calculate with one usage case in mind and evaluate with another.

We first calculate the allowed power density for each tower as if it were the sole TV tower[4]. We then conservatively combine these per-tower power densities by using the minimum allowed power at any location. Next, we assign powers to secondary cells by multiplying the local power density by the cell's footprint. We then calculate the self-interference from nearby secondaries and use the final SNR to calculate the fair rate.

These steps are described below; however, some implementation details have been omitted in favor of imparting a high-level understanding.

Per-tower power calculations

The per-tower power calculations can be broken down into the following steps:

- Assume that each TV tower is surrounded by a sea of secondaries which extends infinitely[5] far outward from r_p. We approximate the sea as a series of secondaries located on rings centered around the TV tower. Symmetry implies that secondaries on the same ring are equivalent, thus drastically reducing the number of required computations by considering a "representative" secondary.

---

[4]We also neglect the effects of the borders and coastlines. For example, even a TV tower in San Francisco is assumed to be completely surrounded by secondaries.

[5]Practically, we do not calculate beyond 750 km. We found that results vary little with a "world" of this size.
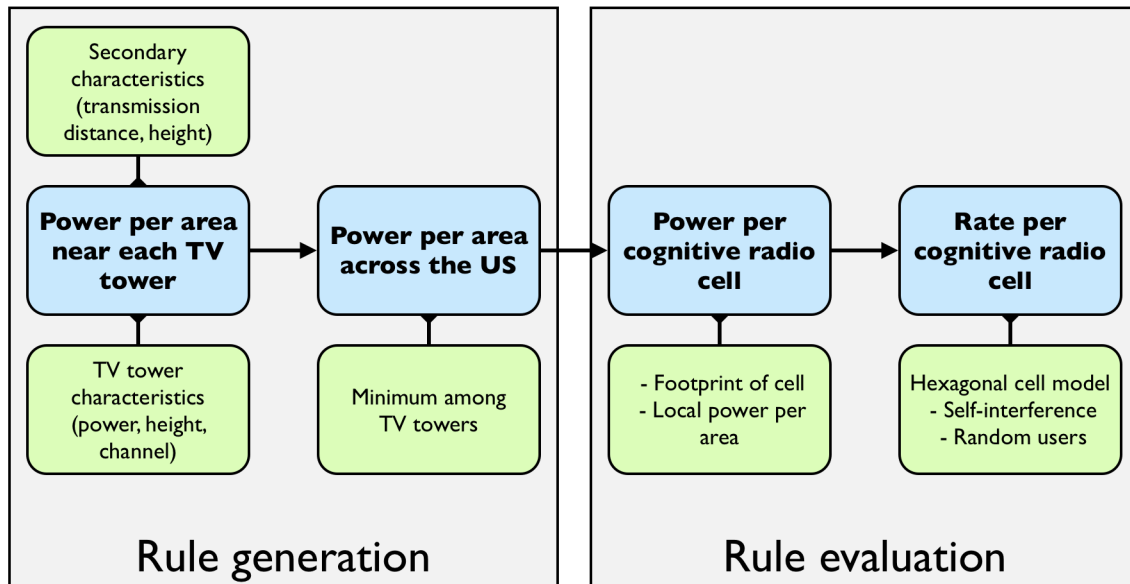
Figure 10.6: Flow of information used to create power density maps and rate maps.

- The representative secondary from each ring calculates its "dream" power and rate. Recall that the "dream" power for ring n finds the safe uniform power density assuming that only secondaries at ring n or further may transmit.

- Secondaries whose "dream" rates are too low (below the threshold[6] β) are no longer allowed to transmit. This threshold effectively defines the value of $r_n - r_p$.

- We then find the maximum γ such that the total interference to a TV receiver at r_p is less than thermal noise[7].

Now we know the maximum safe power as a function of distance to r_p for towers individually. The next step synthesizes this data into nation-wide safe power levels.

Map-level power calculations

For each (channel, pixel) pair, we calculate the minimum power density allowed given the range from each of the TV transmitters on this channel. We call this the cochannel allowed power density for channel c, P_c.

We assume that TV receivers are able to attenuate adjacent-channel signals by 50 dB. This gives rise to a new adjacent-channel "exclusions" rule. Each (channel, pixel) pair is only allowed to use power density $\min\{P_c, P_{c+1} + 50\text{dB}, P_{c-1} + 50\text{dB}\}$.

---

[6]This is formally defined in Section 3.2.3.

[7]Under our assumptions, this is the maximum tolerable interference.

For each location on our map we now have a set of (channel, safe power density) pairs. This completes the rule generation and our next step is to evaluate the rule using the data-rate metric.

Map-level rate calculations

As in the cellular model, the size of the cell depends on the local population density and the value of p. The total allowed transmit power is determined by multiplying the cell area by the local power density.

The calculations of desired signal strength and pollution level are identical to those in the cellular model. However, the self-interference calculations are more involved. We divide our interfering neighbors into two categories, distant and nearby, and add together the interference from each.

- *Distant transmitters*. Since the distance is so great, we can approximate the noise from far-away transmitters by calculating the interference from the center of their pixel to the center of our own. Since pathloss computations are time-consuming, this greatly reduces the run-time.

- *Nearby transmitters*. For nearby interferers the pixel-to-pixel approximation breaks down because the relative error is much higher. Instead, we use nearby power densities in our cellular model which is much more accurate at this smaller scale.

Now that we have the necessary pieces of information, we can calculate the SNR and therefore the rate of each secondary. As in the cellular model, we use the harmonic mean of all users in the cell to represent the fair rate per person of the cell. We now know the achievable data rate available to cellular-style deployments under the power-scaling rule.

10.7 Limitations

Here we detail the assumptions and limitations of our toolkit.

10.7.1 Propagation model

- For simplicity and time, we completely ignore the effects of terrain and horizon. Mähönen, et al., showed that this will likely change local details but not nationwide trends [31].

- We use the ITU propagation model [29] rather than the Section 73.699 propagation model [132] for which the FCC regulations [9] are intended. WEST, as described in Chapter 11, does use the official propagation model.

- The ITU model cannot calculate the pathloss for distances greater than 1000 km nor for heights below 10 meters (roughly rooftop-height).

10.7.2 Population

- We assume that the population of a pixel is roughly uniform. However, the population of the United States varies too quickly for this to actually be true. We show the magnitude of this discrepancy in Figure 10.7 where we plot the CCDF by person of the population density itself. This shows that the statistics of our sampled population distribution do not match those of the true distribution[8]. Even with a 16-fold increase in sampling frequency we do not match the true distribution. It is infeasible to compute such high-resolution maps, so we leave it as future work to show the magnitude of this error after it propagates through our various models.
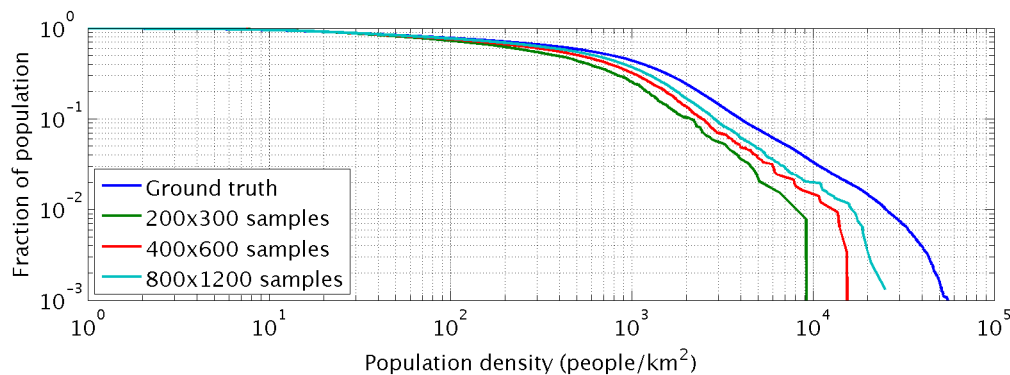
Figure 10.7: Population density sampling results

- Two major regions are missing population data. The first is in southern Nevada and the second is in southwest Arizona. These are shown as black regions on the population map in Figure 3.4. We believe that these are the locations of military bases.

10.7.3 FCC rule interpretations, towers

- The FCC regulations specify calculating the protected contour using eight radial arms extending from the TV transmitter. For each, the radial height above average terrain (HAAT) is calculated, rounded up to 30 meters if necessary, and then used in the protected contour calculation. We use a single HAAT which we round to 10 meters (the smallest possible in the ITU model) if necessary. Because we use a single HAAT, our protected region is circular rather than irregular.

- We ignore all radio astronomy exclusions [9, §15.712(h)]. The locations are shown in Figure 10.8. For visibility the circles shown have a radius of 24 km, ten times that of the actual exclusions. It is unclear how to account for these excluded areas in our power-density model.

---

[8]This was calculated using census-tract-level data which we assume to be the ground truth.
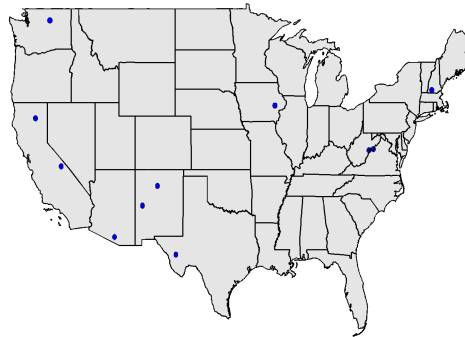
Figure 10.8: Map of the United States showing the locations of the protected radio astronomy sites. For visibility, the protected regions are not shown to scale: the actual excluded regions are 1/100th the area shown. No secondaries may transmit in these regions on any channel.

- We ignore the PLMRS/CMRS exclusions in the 13 major metropolitan areas [9, §15.712(d)]. The number of reserved channels is shown in Figure 10.9(a). When we consider that some channels are already "lost" to TV tower exclusions, we see that the effect of including the PLMRS/CMRS exclusions is not so drastic; this effect is shown in Figure 10.9(b). WEST, as described in Chapter 11, includes these protected entities.
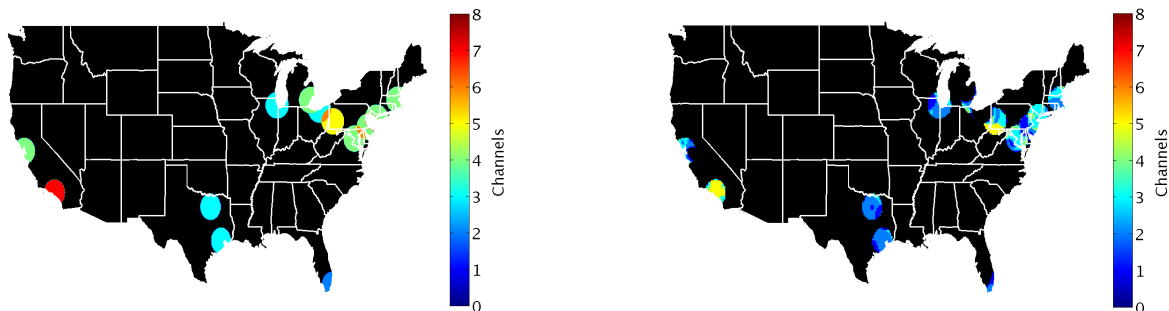
(a) Reserved PLMRS/CMRS channels in the 13 major metropolitan areas.

(b) Number of secondary channels excluded *only* because of PLMRS/CMRS exclusions.

Figure 10.9: Maps of the United States showing the effect of the PLMRS/CMRS metropolitan area exclusions.

- We ignore border protection rules. The FCC rules state that foreign protected contours need only be protected in their country of origin [9, §15.712(d)].

- We assume reception threshold is 15dB, even for analog stations. In fact, except for calculating r_p, we assume all stations are digital.

10.7.4 Power density model

- In order to calculate the "dream" rate and the "new" rate for each tower, we need to assume a transmission range. In the p-people-per-tower models, we assume that the local population density is the average of the population densities inside the TV station's protected contour. In fact, this is exactly the population that will not be served by the secondary devices for which we are calculating the power density. Furthermore, TV stations are located near population centers which means that the population is changing rapidly. We have preliminary results suggesting that this could cause improper power-density scaling but we have not yet remedied the problem.

- We assume that the TV decodability threshold is 15 dB.

Chapter 11

WEST: Whitespace Evaluation SofTware

The work in this chapter has been submitted to IEEE DySpAN 2015 and is pending review. This work was done in collaboration with Vidya Muthukumar[1] and Anant Sahai.

Spectrum whitespaces and dynamic spectrum sharing have become important and interesting topics in recent years. The USA authorized the use of TV whitespaces in 2008 and the UK and Canada followed suit in early 2015. In light of the PCAST report of 2012, additional bands are being evaluated for spectrum sharing in the USA and abroad.

With the increasing momentum of spectrum whitespaces, it is more important than ever to understand the consequences of regulatory decisions. For example, what is the effect of increasing the separation distance from 10km to 15km? Regulators need the ability to understand tradeoffs like this so that they can make informed decisions based on *actual*, not hypothetical or supposed, impact.

Despite the clear need, data-driven analyses appear to be quite rare among regulators, industry members, and researchers alike. Although the data is often freely available, employing it can be an onerous task. In order to reduce this barrier, we have created an open-source software package, WEST, that quickly allows a user to estimate the amount of whitespace in a given region.

For example, after collecting the requisite data, we produced estimates of the amount of whitespace in Canada in under an hour. To demonstrate the power of our software, we present novel results on whitespace availability in Canada and Australia. However, the true potential of WEST lies in the ability to configure it to use existing or hypothetical rulesets. We thus use WEST to compare the FCC and Industry Canada (IC) rulesets, showing that each citizen loses approximately one whitespace channel, due only to the increased size of IC's separation distances as compared to the FCC's. We also showed that although the effect of taboo channels (a notion introduced in the IC ruleset) is small in Canada, it would be much larger if applied to the USA. The identification of

---

[1]A true testament to WEST's ease of use, Vidya produced 100% of the plots within this chapter using the current version of WEST and very little guidance.

the real-world effects of these regulatory decisions was made possible by WEST's ability to create "chimera rulesets," i.e. mosaics of the IC and FCC rules, so that we could examine each variable in isolation.

Finally, we describe the high-level design of WEST. The modular design makes it easy for users to combine, replace, modify, or remove various components to achieve the desired effect. We sincerely hope that the community will use and contribute to WEST, turning it into an even more powerful tool than it is today.

Figure 11.1: Map showing the amount of whitespace (in MHz) available to a fixed device in North America under the FCC ruleset.

11.1 Introduction

Although it may appear from recent spectrum auctions[2] as though all useful spectrum is fully utilized, studies have shown again and again that this is incorrect. For example, [21] and [22] have shown that actually a large fraction of allocated spectrum lays fallow. Since it is not practical to make sweeping changes to existing allocations and deployments, dynamic spectrum access (DSA) is critical for harnessing this spectrum that is allocated yet unused [54].

The incarnation du jour of DSA is as TV whitespaces, the interstices between over-the-air TV stations. The Federal Communications Commission (FCC) in the USA made use of the whites-

---

[2] The 2015 AWS-3 auction sold 65 MHz of spectrum for almost $45 billion USD [19].

paces legal in 2008 [1] (with updates in 2010 [9] and 2012 [23]) and Singapore followed suit in 2014 [133]. Ofcom in the UK [25] and Industry Canada [134] did the same in early 2015. We fully expect that it is simply a matter of time before TV whitespaces around the world are legal to use.

At the same time, other bands are under consideration for spectrum sharing. In the USA, the 3550-3650 MHz and 5350-5470 MHz bands are undergoing sharing investigations [2, 3]. As discussed in Chapter 9, GSM whitespaces are already being used in other countries (albeit illegally).

One of the most common questions asked when considering opening up a new band or region for whitespace use is: what is the size of the opportunity? That is, "how much whitespace is there?" Early papers [48] on TV whitespace in the USA quantified this in terms of MHz available across the USA, and later [30] (also Chapter 2) in terms of potential data rates under a given deployment model. These papers helped inspire future work quantifying the amount of whitespace in Europe [31, 32]. Other papers have recognized the importance of these studies and called for them to be done in other regions [36]. Some regulators have even started using them to highlight the effects of their regulatory decisions [37].

Despite the universally-recognized importance of these studies, there is no easy way for an interested party (regulator, researcher, or industry member) to carry one out themselves without a lot of preliminary work. Existing solutions include static images on websites [135–138] and closed-source software tools [139], all of which work exclusively for TV whitespaces and most only for a single region.

Much of our prior work has relied heavily on our own Matlab code base [6]. While it has been open-source for several years and has several users, it is focused on TV whitespaces in the United States. Because it grew organically throughout the course of our research, it was not designed with flexibility in mind. Finally, it was written in Matlab which is not freely available to all who may wish to use it, nor is it easy to integrate with other tools (e.g. Amazon's AWS).

After years of experience in the field of policy for dynamic spectrum access, it was clear that there was an unfulfilled need for open-source software which was flexible enough to study a wide variety of bands in any region of the world. As a result, we decided to build WEST [7], a framework for evaluating the whitespace opportunity in any region and with any incumbents. WEST is written in Python and features a modular design which makes it easy to combine, replace, modify, or remove various components to achieve the desired effect. WEST has already been used to study spectrum reallocation scenarios related to the FCC's upcoming incentive auctions (see Chapter 6) and we hope it will be used for many more papers in the future.

The key idea behind WEST is that most meaningful results stem from a single structured collection of data: a whitespace availability map. WEST provides the tools necessary to produce these maps from a wide variety of data sources. Section 11.3 elaborates on this idea that whitespace availability maps are the "thin waist" of whitespace explorations and gives examples of derived data.

This chapter is organized as follows. First, we further motivate the design of WEST by presenting novel results on the amount of whitespace in Canada and Australia. We also compare the FCC and Industry Canada rulesets and quantify the real-world impact of their differences. We then briefly show results on the amount of *contiguous* whitespace spectrum available in the United States.

After highlighting some of the key capabilities of WEST, we then describe its design and how it can be extended for use with new regions. Bonus Section 11.8 provides detailed descriptions of the components of WEST while Section 11.9 gives a short code sample.

11.2 Applications of WEST

Since WEST was designed to study whitespaces with different rulesets and in different regions, we put it to the test. We show the results of the first public study of the whitespace opportunity in Canada and Australia — under both FCC and Industry Canada regulations. We then analyze the differences between the FCC and Industry Canada rulesets, highlighting not only the regulatory differences between the two but also the *impact* of these differences.

11.2.1 Canadian whitespaces under FCC regulations

As a test of WEST's extensibility, we carried out an exercise to quantify the amount of whitespace available in Canada *under the FCC's rules*. This required two pieces of external data:

- A geographic boundary file describing Canada's borders[3]

- A listing of the Canadian TV stations[4]

The main difficulty lay in finding the data.

Geographic boundary files are not always accurate or without syntax errors. However, once a suitable boundary shapefile[5] was found, it took only 5 lines of code to import the data into WEST. This is in large part because WEST includes support for reading shapefiles since they are a very common geospatial data format.

Industry Canada freely provides their TV station database to interested parties[6] but it is in an ancient format that only Windows computers can read. Google's spectrum database data download page[7] includes data for Canadian TV stations, but only those near the USA border. However, the

---

[3]See http://www.nws.noaa.gov/geodata/catalog/national/data/province.zip.
[4]See http://data.fcc.gov/download/incentive-auctions/Constraint_Files/Canadian_Allotment_List_2014May20.xlsx.
[5] A shapefile is a common geospatial data format for Geographic Information System (GIS) software.
[6]See http://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/h_sf09484.html.
[7]See https://www.google.com/get/spectrumdatabase/data/.

FCC has released the full Canadian TV station set[8] as part of the upcoming incentive auction dataset. Once the correct data was found, existing code for USA TV stations was readily adaptable to the Canadian TV dataset.

It took about 2 hours and 200 lines of code to find and import the data described above. It only took 7 lines of code to calculate the number of available whitespace channels and another 7 lines to plot them, as seen in the Canadian part of Figure 11.1. The source file for this calculation can be found at `https://github.com/kate-harrison/west/blob/master/examples/canada_tvws.py`.

To our knowledge, this is the first study that has been conducted on nationwide whitespace availability in Canada. Although Spectrum Bridge offers a whitespace finder[9], it is limited to providing a list of available channels at a user-specified location, making nationwide trends impossible to discern.

11.2.2 Australian whitespaces under FCC regulations

Because it was so quick and easy to do, we gathered data for Australia [140–143] and computed the amount of available whitespace across Australia, despite the fact that there are no candidate whitespace rules for this region. The results are shown in Figure 11.2.

As expected, a great portion of Australia's area has copious amounts of whitespace due to a lack of inhabitants. The populous coastal regions, however, are doing somewhat better than expected: after the digital transition, about five channels allocated for TV use were not actually used. Also, a quick look at the locations of TV stations in Australia (on Google Earth) tells us that the stations are much more spread out along the coastal regions than in the United States or Canada. This combined effect leaves Australians with a *minimum* of 140 MHz of whitespace spectrum everywhere. This is confirmed by the CCDF in Figure 11.4(a). Unlike the United States and Canada, Australia has a wealth of whitespace spectrum waiting to be unleashed by regulators.

Although this work was suggested in 2011 by [36], it has been sufficiently difficult to complete that no one has done it — until now. Once the Australian TV allocation data was in hand, it took about an hour to integrate it into WEST and produce the map below. Some of the adaptations that were unique to Australia were its 7 MHz channel bandwidth and its slightly different frequency allocations[10], but WEST was able to handle both of these easily.

---

[8] See `http://data.fcc.gov/download/incentive-auctions/Constraint_Files/`.

[9] See `http://whitespaces-canada.spectrumbridge.com/`.

[10] The TV bands in Australia are not precisely aligned with those in North America. For example, Australia's "VHF Band III" covers 137-230 MHz while the F-curves propagation model is defined for only a subset of these frequencies. We were able to produce reasonable estimates for Australian whitespace availability by approximating the service areas of TV stations operating on these out-of-range frequencies.
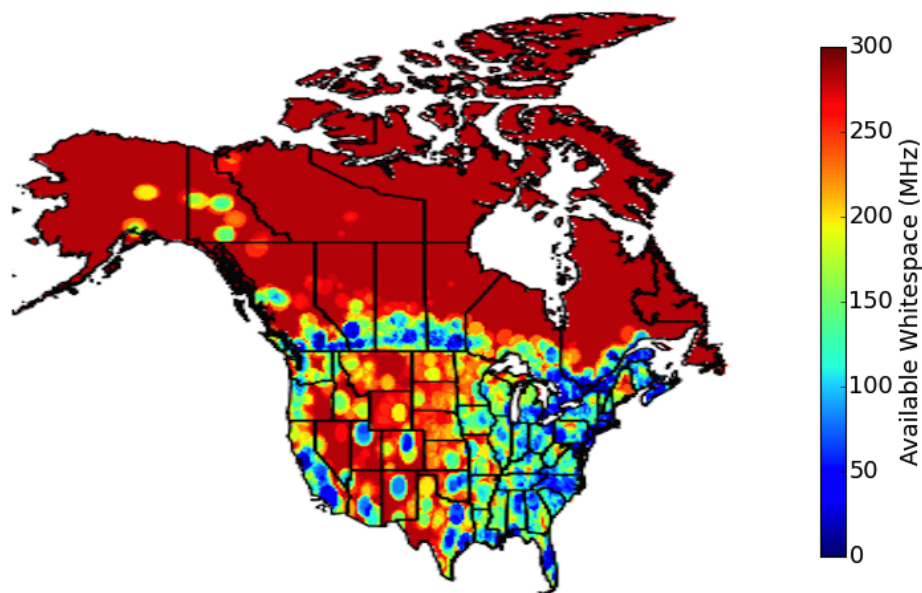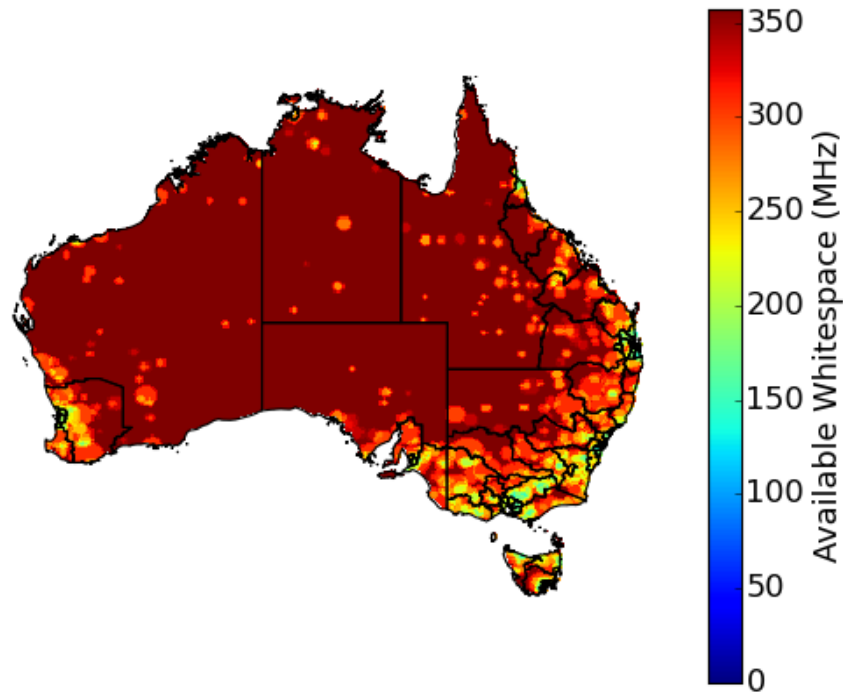
Figure 11.2: Map showing the amount of whitespace (in MHz) available to a fixed device in Australia under the FCC ruleset.

11.2.3 Studying the Industry Canada whitespace regulations

Industry Canada released their regulations earlier this year, providing an excellent opportunity to test the flexibility of WEST's regulatory module[11]. Although the IC ruleset is very similar to that of the FCC, there are a few key differences:

1. Definition of protected contour. The target field strength for UHF stations is fixed in the FCC regulations whereas it changes with frequency in the IC regulations (achieving equality with the FCC at 615 MHz, the middle of the UHF band).

2. Minimum separation distances[12] are typically larger under the IC regulations.

---

[11] Ofcom's rules are unfortunately unsuitable for use outside the UK because they require proprietary information that only exists for the UK and is only available to specific entities. Singapore's rules are similarly very purpose-built and hard to generalize outside of Singapore.

[12] A device may only utilize co-channel whitespaces if it is at least this far from the nearest protected TV contour. Different minimum distances are defined for adjacent- and taboo-channel operation. The distances may also depend on the characteristics of the whitespace device and the TV station.

3. The IC regulations also impose minimum far-side separation distances[13].

4. Taboo channels. The IC regulations define "taboo channels" to protect analog stations. This has the effect of banning whitespace devices from not only transmission on the same and adjacent channels, but also on several other channels within the service contour of the analog TV station.

Both rulesets utilize the same propagation curves [132] for calculating the service areas of TV stations and define identical device classes. This overview covers the most important differences between the two rulesets; readers interested in the full set of differences are referred to the source material [23, 24].
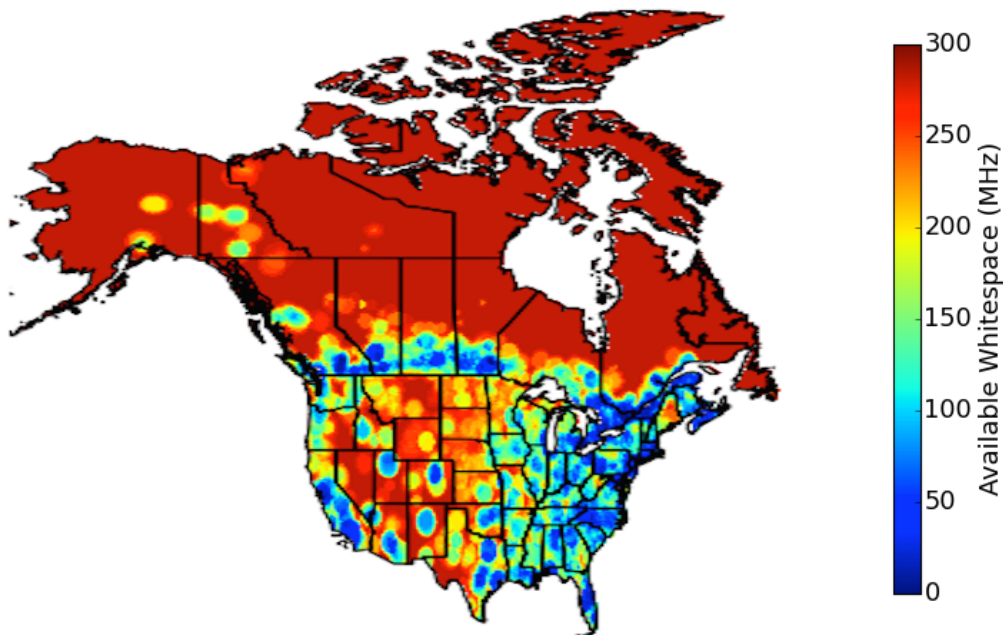
Figure 11.3: Map showing the amount of whitespace (in MHz) available to a fixed device in North America under the Industry Canada ruleset.

Figures 11.1 and 11.3 show the amount of available whitespace in North America under the FCC and Industry Canada rulesets for fixed devices. We restrict our discussion to the United States and

---

[13] The far-side separation distance defines the minimum separation distance from the *far* side of a TV station's protected contour (in contrast, all other separation distances are defined with respect to the nearest point on the contour). See Section 4 of http://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/sf10928.html for additional details.

(a) CCDF by population

(b) CCDF by area
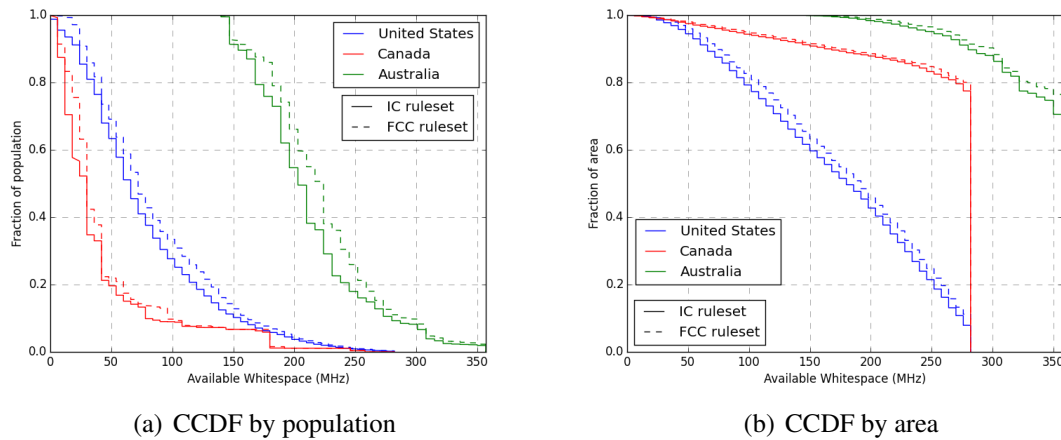
Figure 11.4: Complementary cumulative distribution functions (CCDFs) of available whitespace by population and area in the United States, Canada, and Australia for fixed devices.

Canada to keep it brief; there is no reason the same analysis could not be done for Australia. The differences are difficult to discern from these plots alone but are more obvious in Figure 11.4. We make two key observations:

- **People have more whitespace in the USA than in Canada.** Regardless of the ruleset used, Americans appear to have a greater number of whitespaces than Canadians. The median Canadian has access to about 5 fewer whitespace channels (30 MHz) than the median American. Although much of Canada's *area* experiences an overabundance of whitespaces, its population is more concentrated in urban areas which are typically less rich in whitespace.

- **The FCC ruleset appears to be more permissive in both the USA and Canada.** In the USA, each person has approximately one additional whitespace channel under the FCC ruleset. The difference is less pronounced but still present in Canada.

Fully understanding the differences between the two rulesets and how they are manifested in the real world takes a bit of work. We made use of WEST's modular design to create "chimera rulesets" which blend together the two rulesets. For example, one such chimera ruleset may be identical to the FCC ruleset *except* that it uses the separation distances specified by the IC ruleset.

Figure 11.5: Block diagram showing how the chimera rulesets were created from the FCC and IC rulesets.

Figure 11.5 shows how we created these hypothetical chimera rulesets from the FCC ruleset and IC ruleset. In this way we isolate each effect in turn and are thus able to quantify the effect of each decision.

In Canada it is clear that the main real-world difference between the rulesets shows up in the size of the separation distances and, secondarily, in the definition of a TV's service area. Other ruleset differences have negligible effect on whitespace availability.

However, we see that the United States is significantly impacted by the effect of taboo channels. These restrictions affect only areas serviced by analog TV stations, which in the USA are both more numerous and more likely to appear in urban areas. From this we learn that a ruleset is not necessarily objectively "good" or "bad": the context in which it is applied matters a great deal.

(a) Canada   (b) United States
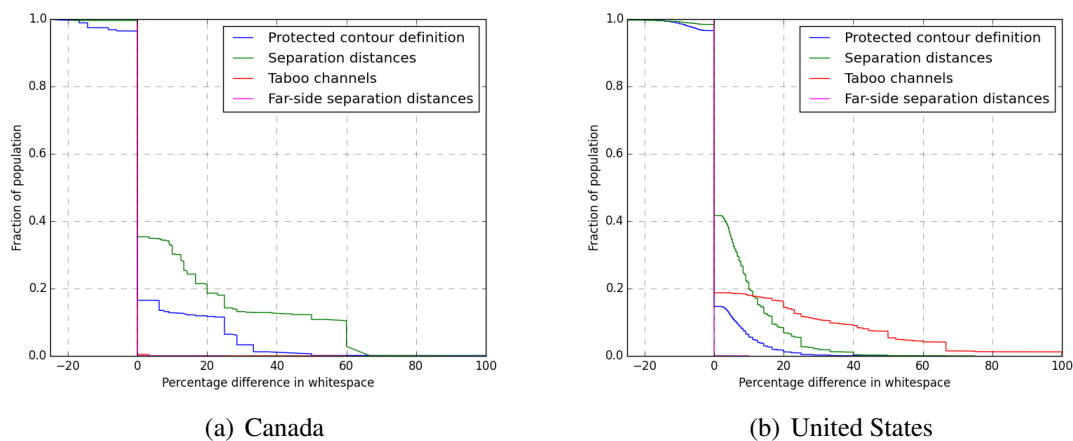
Figure 11.6: Complementary cumulative distribution functions (CCDFs) of the percentage change in whitespace by population due to each effect being applied incrementally in the United States and Canada. For example, using the IC-mandated separation distances instead off the FCC-mandated separation distances costs about 20% of Canadians about 20% of their whitespace.

We now describe in greater detail how each of the ruleset differences impacts whitespace availability:

1. *Separation distances:* For almost all TV stations, the IC ruleset mandates greater separation distances from stations' protected contours than the FCC ruleset. Under the IC ruleset, separation distances are greater by anywhere from 1 kilometer to 30 kilometers. Consequently, whitespace devices need to maintain a much greater distance from TV stations to be allowed to operate than under the FCC ruleset. This results in a decrease in available whitespace across both Canada and the United States.

2. *Target field strength:* The target field strength of a TV station, and therefore its defined service contour, differs by at most a kilometer between the two rulesets. This has the small

effect of slightly increasing or decreasing the amount of available whitespace for a relatively small number of people in both the United States and Canada.

3. *Taboo channels:* The introduction of taboo channel exclusions in the IC ruleset affects whitespace availability only in locations that are served by analog TV stations. In Canada, only 2% of TV stations are analog, and they serve only 2% of the population. In contrast, analog TV stations serve 27% of the USA population. So while taboo channel exclusions have a negligible effect on available whitespace in Canada, they would cause a substantial decrease in available whitespace in the United States.

4. *Far side separation distances:* Far side separation distances cause additional loss in whitespace only around stations with very small protected contours. This condition affects the protected contours of only 2-3 stations in both the United States and Canada, which translates to 0.3% of studied locations in Canada and 0.1% of studied locations in the United States, Therefore, the far side separation distance condition would affect whitespace availability in a negligible way.

The power of WEST is clear from this exercise. First, it allows the quick exploration of whitespace availability in another region as discussed in Section 11.2.1. Second, it allowed us to easily create "chimera rulesets" in order to explore the real-world effect of regulatory decisions. Note that any ruleset can be modified arbitrarily and need not be based on a real-world ruleset. For example, we could just as easily quantify the effect of adjacent-channel exclusions by modifying our implementation of the FCC ruleset to remove these exclusions. Other publicly-available tools simply do not support these kinds of operations and explorations.

11.2.4 Contiguous whitespace channels

One of the major barriers to whitespace utilization is the fact that whitespaces represent discontiguous spectrum. Since whitespaces are only found in the interstices of the TV spectrum, there is no guarantee that whitespace channels will be adjacent in frequency. This presents challenges in the design of whitespace devices.

With WEST, it is simple to calculate the number of contiguous whitespace channels: it can be done in less than 50 lines of code[14]. Figures 11.7 and 11.8 show the result. We see that most places (Dallas, TX, and New York City being the most notable exceptions) have at least two contiguous whitespace channels available. We also see that the median United States citizen has about 5 contiguous whitespace channels, translating to 30 MHz of contiguous whitespace spectrum. The results are similar for portable devices with the notable exception that their maximum number of contiguous channels is limited by the number of whitespace channels which are open for use by

---

[14] See https://github.com/kate-harrison/west/blob/master/examples/contiguous_channel_count.py.

personal/portable devices[15]. This high availability suggests that systems should be designed to take advantage of contiguous spectrum when it's available, with the ability to fall back to a single whitespace channel if necessary.
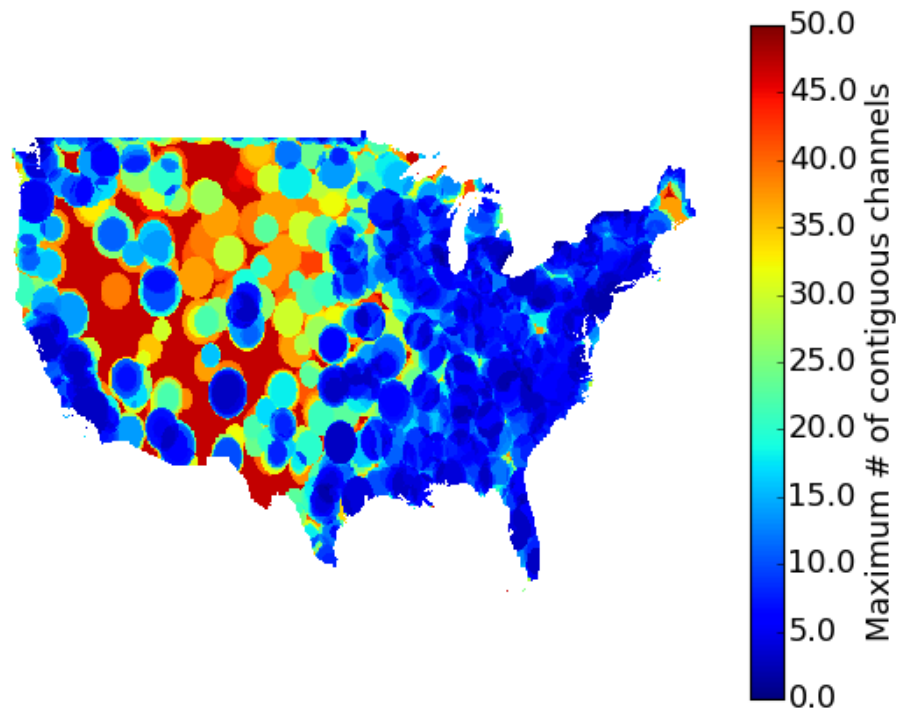
Figure 11.7: Map showing the maximum number of contiguous whitespace channels available to a fixed device in the United States.

11.3   Using WEST

Creating whitespace availability maps—like those shown throughout this chapter—is remarkably easy with WEST. In fact, these maps are merely the most basic output: as Figure 11.9 shows, there are many ways of exploring and presenting the data.

The user begins by specifying the device (e.g. fixed vs. personal/portable) and Region (e.g. United States) of interest. The Region object contains information about the protected entities in that region, as described in Section 11.8. Since WEST's calculations are done on "pixelized" maps,

---

[15]Currently portable devices are only allowed to use TV channels 21-51 while fixed devices have access to up to 17 additional channels.
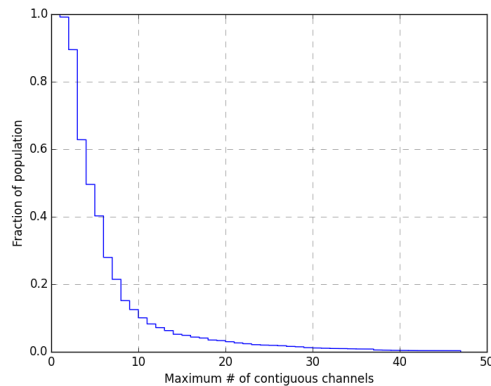
Figure 11.8: Complementary cumulative distribution function (CCDF) of number of contiguous whitespace channels available by population in the United States for fixed devices.

the user must also specify the resolution for the output data in the form of a DataMap2D, which is essentially a 2-D matrix with geographic metadata.

This data is used as input to a Ruleset object (e.g. one implementing the FCC 2012 ruleset) along with a channel number. The Ruleset object then calculates which locations are considered whitespace on that channel and returns a binary DataMap2D holding this information.

From here, the possibilities are numerous:

- One of the most common things — as shown in Figures 11.1 and 11.2 — is to perform this operation for all whitespace-enabled channels in the region and create a map of the **whitespace channel count**.

- By repeating this process for a second ruleset and then subtracting the results, one can also create a **whitespace delta map** which highlights the differences between two rulesets. This can also be done to compare e.g. fixed vs. portable whitespace availability.

- Seeing a **CCDF by area or population** of the same data can often lead to additional insights and is a natural way to quantify the effect, as seen in Figures 11.4 and 11.6.

- Other work of ours (see Chapter 6) utilized the same machinery to compare band reallocation with the whitespace approach for recovering spectrum via **plots of the Pareto curves**.

- The same work also used **2D histograms** to compare whitespaces under two different scenarios. Details on how WEST was used to generate these figures can be found in Appendix B.

- With the additional assumption of a self-interference or deployment model, one could also calculate **data rate maps** like those in [30] and Chapter 2.

The key is that all of these results come from further processing of a basic data type — the whitespace availability map — which is what WEST is designed to compute.
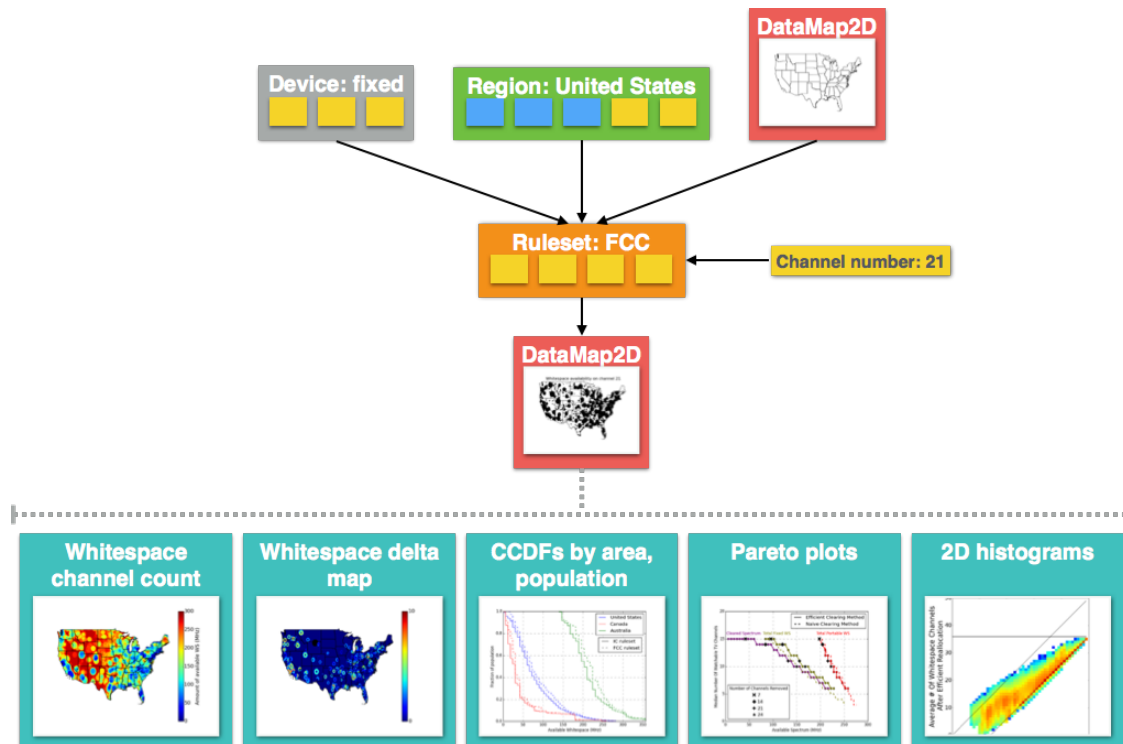
Figure 11.9: Flow diagram for generating a whitespace availability map and derived output.

11.4 Extending to other regions

Extending WEST to support a new region, as was done in Section 11.2, is quite easy once the correct source data is in hand. Using Canada as an example, we took the following steps:

1. Create a new BoundaryCanada class. This class represents the geographic boundary of the region and is primarily used for display purposes (e.g. drawing provincial borders). Since there is built-in support for shapefiles in the Boundary class, the new class only needed to specify the filename for the source data once an appropriate shapefile was found.

2. Create a ProtectedEntities class to read in and hold data about USA and Canadian TV transmitters. Although the code is specific to the format of the source data file, much of the code from the USA-only data files could be reused.

3. Create a RegionCanada class to specify e.g. which channels are available as potential whitespace and which entities should be protected (i.e. the class from Step 2).

4. (Optional) Create a DataMap2DCanada class which will, by default, have latitude and longitude bounds appropriate for Canada. This class specifies how the region's area will be pixelized. Creating a specific DataMap2D subclass makes it both easier to reference as well as ensures compatibility between generated data.

These classes are described in detail in Section 11.8. The code for all of these steps can be found at `https://github.com/kate-harrison/west/blob/master/examples/canada_tvws.py`. It uses USA and Canadian TV station data with the FCC ruleset to compute a map estimating the whitespace availability in Canada.

11.5 Extending to other rulesets

The analyses we carried out in Section 11.2 were made possible primarily because the architecture of WEST enables us to easily create extended or modified rulesets, whether based on real, candidate, or hypothetical rulesets.

The first ruleset we created to compare to the FCC ruleset was a reasonably faithful implementation of Industry Canada's 2015 ruleset. Implementing the Industry Canada ruleset required creating a new subclass of the Ruleset class (see Section 11.8 for more details about this class). Since there are many similarities with the FCC ruleset, much of the code could be borrowed. By obeying the interface defined by the Ruleset class, we ensured that the FCC and IC rulesets could be used interchangeably.

We next created "chimera rulesets" as described in Section 11.2. Due to the carefully-planned structure of the code, we were able to easily add or remove ruleset components with only a few lines of code. Exploiting the structure of the Ruleset class enabled us to create all three chimera rulesets in less than 20 *total* lines of code.

What makes WEST so powerful is its ability to support not just implementations of existing rulesets, but also the creation of hypothetical rulesets. The advantage of these hypothetical rulesets is twofold. First, regulators can test, ahead of time, the empirical effects of their candidate rulesets on available whitespace, enabling them to come up with better designs. Hypothetical rulesets are also useful to researchers, as they can facilitate more detailed whitespace analyses across different regions. Second, we can quantify the whitespace availability in countries that have not yet enabled the use of whitespace spectrum by applying other countries' existing rulesets. This is exactly what we did in Australia since ACMA has not developed a candidate ruleset yet; this allowed us to generate the first public maps estimating whitespace availability in the region.

11.6 Conclusions

This chapter began by presenting novel results on the amount and location of whitespaces in both Canada and Australia. It continued with a comparison between the FCC and Industry Canada whitespace rulesets as applied to the United States and Canada. We showed that the FCC ruleset is more permissive than the IC ruleset, particularly in the USA. Via explorations with WEST, we discovered that this was primarily due to the effect of increased separation distances from stations' protected contours, and taboo channel exclusions.

The results above are only made possible by WEST, an open-source software package we developed to help researchers, industry members, and regulators alike. Our goal with WEST is to enable data-driven analysis of both the current and the hypothetical whitespace opportunities. It is designed to be flexible enough to work with a variety of regions, protected entities, rulesets, and propagation models. This chapter has demonstrated some of the power of WEST but has not exhausted its capabilities. We look forward to future papers in the community that will use WEST in new and interesting ways.

We strongly believe that regulations in particular should be data-driven, giving regulators full understanding of the tradeoffs they are making between interested parties. This not only increases transparency but also the ability of regulators to strike the tradeoff points they desire. Furthermore, it gives them the tools they need to independently analyze these tradeoffs rather than relying on industry members and lobbyists for these results.

11.7 Future work

The most important future work will be done by members of the spectrum sharing community, each contributing their own data-driven perspective on whitespaces. We strongly encourage interested parties to build on WEST and share their extensions with the community if and when possible.

In particular, users of WEST would benefit from the following additions:

- Terrain data (we currently assume the world is flat which yields circular contours)

- More regions (currently only USA, Canada, and Australia are supported)

- More protected entities (WEST is currently focused on TV but also includes support for radioastronomy sites and PLMRS/CMRS; one could add support for MVPD sites, BAS links, etc.)

- Test on another band (true test of flexibility will be application to e.g. 3.6 GHz)

- More rulesets (e.g. DSA model rules [144], any European rules that may come out)

- More propagation models (e.g. Hata)

As with most open-source projects, WEST relies on the community for growth. Anyone wishing to contribute to WEST is encouraged to send an email to the authors.

But the goal is not just to add support for existing concepts like those listed above: we also want to see WEST used in new and creative ways and by a variety of users. The community as a whole will benefit greatly from seeing perspectives more diverse than those of a single research group. To that end, we encourage the reader to pursue data-driven explorations – whether they use WEST or not.

11.8   Modules

In this section, we describe the design of WEST. WEST contains the following high-level modules:

- **DataMap** (DataMap2D or DataMap3D): the standard format for data is a 2-D matrix with geographical metadata. A plotted DataMap2D yields figures like Figure 11.2. A DataMap3D allows for logical aggregation of DataMap2D objects.

- **Population**: reads in population data and creates a population DataMap2D.

- **Region**: specifies various parameters about a region such as which channels are available for whitespace use, channel bandwidth, and the set of protected entities.

- **Boundary**: specifies the boundary of the given region; mostly used for plotting purposes.

- **ProtectedEntity**: specifies an entity (e.g. TV station) that may be eligible for protection.

- **ProtectedEntities**: a collection of ProtectedEntity objects (e.g. the set of all TV stations in the United States).

- **Ruleset**: describes how to protect various entities.

- **Specification**: describes an experiment in a parametrized way; used to quickly recall or generate data.

Each of these will be discussed in turn in the following subsections.

11.8.1   ProtectedEntity

A ProtectedEntity object describes an entity which may be eligible for protection from whitespace devices. Figure 11.10 shows an example of a specific protected entity: a TV station. The object contains the essential information about the protected entity (e.g. location, height, transmission power, frequency, transmitter type) but notably does not contain any information about the nature of the protection. The nature of the protection is inherently a property of the Ruleset, not the protected entity itself.

Other protected entities may include radioastronomy sites or PLMRS sites, as in the case of the TV whitespaces in the United States.
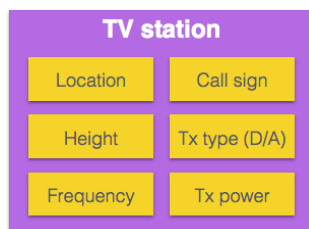
Figure 11.10: Visual representation of a ProtectedEntityTvStation object.

11.8.2   ProtectedEntities

A ProtectedEntities object is a collection of ProtectedEntity objects of a single type. For example, Figure 11.11 shows a collection of TV stations. The ProtectedEntities class is responsible for specifying the data source and parsing it to create individual ProtectedEntity objects.

While some ProtectedEntities objects will differ in the type of ProtectedEntity they contain, others will differ based on the source of the data. For example, a user may create a ProtectedEntitiesTvStationsFromGoogle[DATE] class which ingests a file downloaded from Google's spectrum database data download page [145] and another ProtectedEntitiesTvStationsFromTvQuery[DATE] class which ingests data from the FCC's TV query website [28]. In this way, the user can easily switch between the two datasets by simply specifying the class s/he wishes to use.
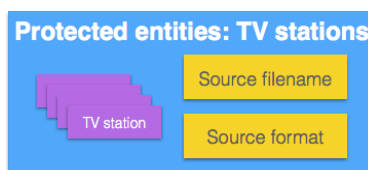
Figure 11.11: Visual representation of a ProtectedEntitiesTvStations object.

11.8.3   Region

A Region object contains a collection of ProtectedEntities which may be afforded protection within the region. It also contains geographical information (in the form of a Boundary object) which describes the physical boundary of the region.

Finally, it contains information on the channels (and their corresponding frequency representation) which *may* be available for whitespace use, subject to the protection of the protected entities. A distinction is made between channels which may be available for portable devices vs. fixed devices. While this information could also be contained in the Ruleset class, it fits more naturally in the Region class.
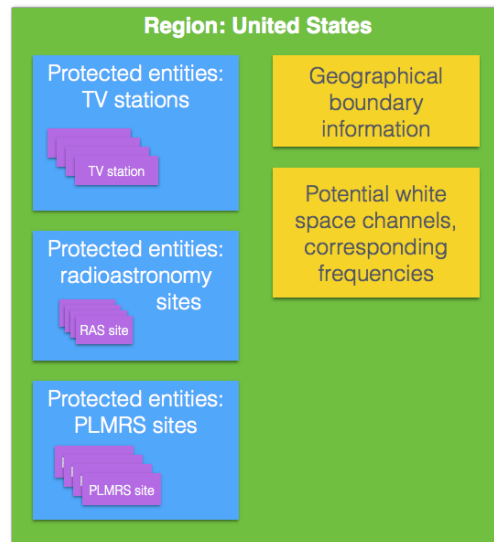
Figure 11.12: Visual representation of a Region object

11.8.4 DataMaps

Information is collected in objects called DataMaps. They come in two varieties: 2D and 3D. Typically, a DataMap2D will be used to describe a single piece of geographically-dependent information, e.g. the availability of a single channel for whitespace use. DataMap3D objects, on the other hand, contain multiple pieces of geographically-dependent information, e.g. the varied availabilities of a collection of channels for whitespace use.

A DataMap2D is essentially a two-dimensional matrix with metadata and helper functions. It is specified by its geographical bounding box and the number of latitude and longitude divisions (i.e. resolution). DataMap2D objects typically but not necessarily describe points within a Region, as suggested by the image in Figure 11.13. Two DataMap2D objects are considered comparable if they have the same geographic boundary and resolution.
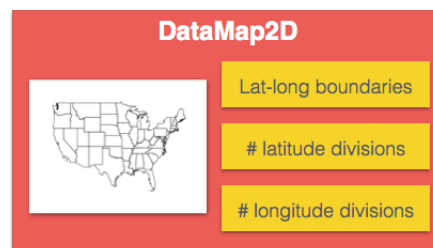
Figure 11.13: Visual representation of a DataMap2D object.

A DataMap3D is little more than a collection of DataMap2D objects and associated helper functions to facilitate the combination of the DataMap2D objects (e.g. pointwise summation). Each

constituent DataMap2D is called a layer and has a unique label. Layers within a DataMap3D must be comparable and this is strictly enforced.
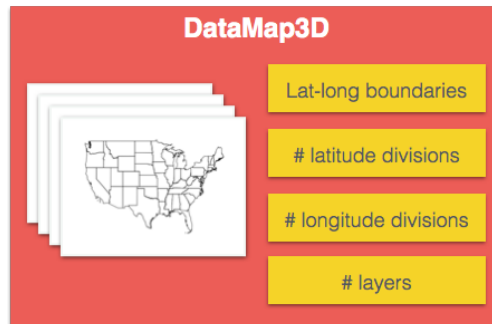
Figure 11.14: Visual representation of a DataMap3D object.

11.8.5 Submaps

DataMap2D objects have the ability to extract a submatrix of themselves and its associated meta-data, creating another DataMap2D. This allows for more regional processing of data. Submaps are themselves DataMap2D objects and can therefore be used in the same way as the parent DataMap2D.

A submap is created by specifying its geographic bounding box to the original DataMap2D. Steps 1 and 2 of Figure 11.15 shows a visual representation of this process. The submap is populated with a *copy* of the data in the original DataMap2D. To use the values from the submap in the original DataMap2D, the submap must be "reintegrated."

The submap infrastructure was used heavily in Chapter 6. Because we were computing the amount of available whitespace under tens of thousands of different TV station allocation scenarios[16], it was necessary to cache the protected region of each TV station. A cached protected region was a binary-valued submap of the continental USA DataMap2D. To determine the union of protected regions, we simply reintegrated the submap corresponding to each TV station, combining the parent and submap's values using logical OR. Figure 11.15 shows this process for a single TV station.

11.8.6 Ruleset

One of the more complicated objects is the Ruleset object. As shown in Figure 11.16, it contains the protection criteria for all applicable protected entities and the logic for combining various protection rules. The design of each Ruleset class will depend heavily on the style of the regulations.

WEST currently supports a version of the FCC's TV whitespace rules in the class RulesetFcc2012. This class was designed to be extensible, with most functions performing a very specific task so

---

[16]While the channel assignment of a given station changed between assignments, its service area was preserved.
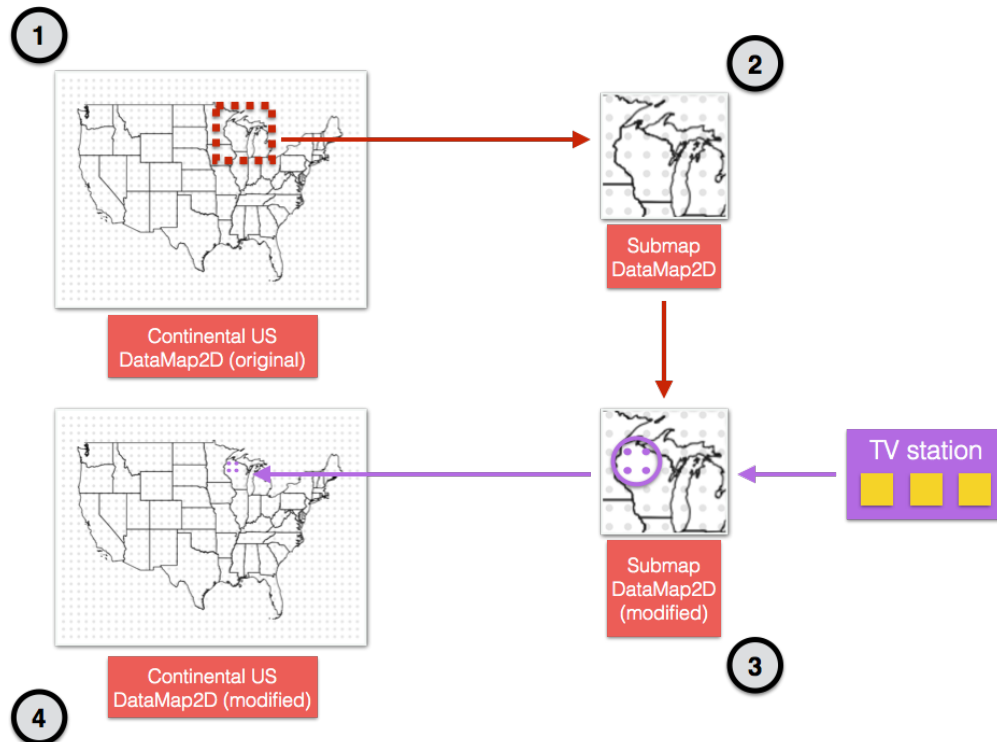
Figure 11.15: Submap reintegration process. Step 1 shows the original DataMap2D which contains data for the entire continental USA. Step 2 shows an extracted submap containing data for a portion of the midwest. Step 3 computes which pixels of the submap are not available for whitespace use due to the need to protect a particular TV station. Step 4 reintegrates the submap into the original DataMap2D. Note that the purple pixels in the modified submap are also purple in the modified continental USA DataMap2D.

that they can easily be replaced. For example, to evaluate the amount of whitespace that would be available using the FCC's rules but different separation distances, one would simply need to subclass RulesetFcc2012 and overwrite the two functions which return the co- and adjacent-channel separation distance, respectively[17].

11.8.7 Specification

Data-driven investigations often cache data in order to trade disk space for computational time. We have found this to be especially necessary when working with whitespace-related data. For example, caching the protected regions as discussed in Section 11.8.5 allowed us to perform tens of thousands of computations that otherwise might have taken up to 15 minutes each. In order to facilitate the creation, storage, and retrieval of data, we created the Specification class.

---

[17]These functions are called get_tv_cochannel_separation_distance_km() and get_tv_adjacent_channel_separation_distance_km().
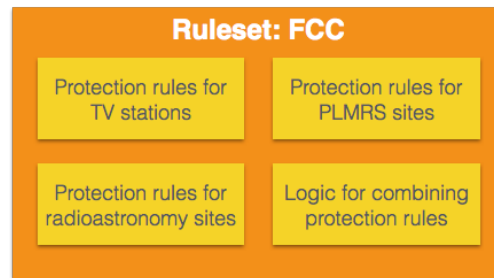
Figure 11.16: Visual representation of a Ruleset object.

A Specification subclass describes how to create a particular type of data and how to programmatically generate a [hopefully unique] filename for the resulting data. A Specification object is instantiated with the parameters needed to uniquely describe the data. When the user calls the fetch() method, the following happens:

- If a file with the corresponding filename *does not* exist, generate the data and save it to disk. Then return the data to the user.

- If the file does exist, load the data and return it to the user.

In this way the user does not need to worry about whether or not the data already exists when fetching the data. This prevents a lot of boilerplate code and streamlines the entire process. See Section 11.9 for an example.

WEST itself defines several Specification subclasses:

- **SpecificationDataMap** describes a DataMap2D or DataMap3D. It is primarily used as input to other Specifications.

- **SpecificationRegionMap** describes a DataMap2D whose values will be 1 inside the Region's Boundary and 0 outside.

- **SpecificationWhitespaceMap** describes a DataMap3D whose values will be 1 at locations which are available for whitespace use and 0 otherwise.

- **SpecificationPopulationMap** describes a DataMap2D where the value of each "pixel" is the number of people living inside the area covered by the pixel.

Note that the Specification class *does not* tackle the problem of cache invalidation. If the user edits the procedure for data generation or any prerequisite data, s/he is responsible for recognizing which files are affected and deleting them.

The Specification class also *does not* make any strong guarantees on unique naming. The Specification subclasses defined in WEST itself will generate unique names; however, no such guarantee can be made about user-defined subclasses and hence the user should take care when defining a new Specification.

11.8.8 Other modules

There are several remaining modules which are just as essential yet are more straightforward than the previous modules. We briefly describe them in this section. We also note that WEST fully supports the use of Sphinx[18] which generates HTML documentation for all modules. We refer interested parties to the documentation itself, as that is both more complete and more current than this chapter.

11.8.8.1 Boundary

This class takes in a geographic boundary of a region (e.g. the outline of the continental United States or the outlines of individual states). Although it is meant to work with generic geographic data, it has built-in support for the common shapefile format. This class is used when creating the logical DataMap2D that describes which pixels are inside a Region, used both to speed up computations as well as to make the map "background" white. It is also used to draw Region outlines when plotting DataMap2D (e.g. the state outlines in Figure 11.1). Note that the Boundary objects used for these two functions may be distinct.

11.8.8.2 Population

Figure 11.4 above used data about each Region's population to calculate CCDFs *by population*. Our USA population data was obtained from [10, 11], our Canadian population data from [146, 147], and our Australian population data from [142, 143]. The Population class in WEST makes relatively it easy to read in the necessary geographic and census data, automatically linking it together and creating a DataMap2D with pixelized population data. For example, the code to create the USA population map is about 75 lines long.

11.8.8.3 Plotting

Plotting functionality is naturally included within WEST. The most notable features include the ability to overlay arbitrary Boundary outlines (of any color) and the ability to use a white background for the image.

11.8.8.4 PropagationModel

While WEST currently implements only the FCC's F-curves, it is possible to add support for arbitrary propagation models. The interface to a PropagationModel object is well-defined and thus PropagationModel objects can be used interchangeably. For example, one could easily specify a new FCC-inspired ruleset which uses a different propagation model to calculate the TV service contours.

---

[18]See http://sphinx-doc.org/.

11.8.8.5 Data manipulation

Since CCDFs form a core part of our explorations, we provide functions which will calculate the CDF from a given DataMap2D using arbitrary weights. We commonly use the population data as weights to produce CCDFs by population.

11.9 Code sample: calculating the amount of available whitespace

```
# This code sample demonstrates how to generate and plot data representing TV
# white space availability in the United States.

from west.data_management import *
from west.data_map import *
from west.boundary import BoundaryContinentalUnitedStates, \
    BoundaryContinentalUnitedStatesWithStateBoundaries
from west.region_united_states import RegionUnitedStates
from west.ruleset_fcc2012 import RulesetFcc2012
from west.device import Device

# Specify the type of device
test_device = Device(is_portable=False, haat_meters=30)

# The data will be on a 200x300 lat-long grid. The first argument specifies
# the lat-long boundaries of this grid. In this case, those boundaries are:
#    [24.5, 49.38] degrees latitude
#    [-124.77, -66] degrees longitude
# These values can be see in the definition of
# DataMap2DContinentalUnitedStates in west.data_map and can be altered by
# subclassing DataMap2DWithFixedBoundingBox in that module.
datamap_spec = SpecificationDataMap(DataMap2DContinentalUnitedStates, 200, 300)

# Data will only be computed for locations inside the continental United States.
region_map_spec = SpecificationRegionMap(BoundaryContinentalUnitedStates,
                                         datamap_spec)

# Whitespace will be computed using the values above, protected entities
# defined in RegionUnitedStates, and the ruleset defined in RulesetFcc2012.
is_whitespace_map_spec = SpecificationWhitespaceMap(region_map_spec,
                                                    RegionUnitedStates,
                                                    RulesetFcc2012, test_device)

# Generate the TVWS availability data if it does not already exist;
# otherwise, the data is loaded from disk. The data is in the form of a
# DataMap3D with each layer representing the whitespace availability on a
# particular TV channel. A value of '1' indicates that whitespace is
# available at that location; '0' means WS is not available.
is_whitespace_map = is_whitespace_map_spec.fetch_data()

# Turn the DataMap3D into a DataMap2D by summing all layers at each location.
total_whitespace_channels = is_whitespace_map.sum_all_layers()

# Fetch (generate or load) a region map which will be used to mask the
# DataMap2D in order to create a white background in the image and the
# appearance of a "floating" map.
is_in_region_map = region_map_spec.fetch_data()

# Plot the data
plot = total_whitespace_channels.make_map(is_in_region_map=is_in_region_map)

# Add outlines of the individual states and customize their appearance.
plot.add_boundary_outlines(
    boundary=BoundaryContinentalUnitedStatesWithStateBoundaries())
plot.set_boundary_color('k')
plot.set_boundary_linewidth('1')

# Add a title and colorbar
plot.set_title("Number of available TVWS channels")
plot.add_colorbar(vmin=0, vmax=50, label="Channels")

# Save the plot
plot.save("Number of TVWS channels in the United States.png")
```

Part VII

Conclusions

Chapter 12

Conclusions

This thesis has taken a quantitative perspective on whitespace regulation. It began with what was at the time a novel concept: quantifying the amount of whitespace available across the nation. We showed that urban areas tend to have fewer whitespaces than rural areas and that the urban areas face the additional challenge of an increased noise floor due to TV signals which propagate beyond the official TV service areas. To study the magnitude of this effect, we used several models for secondary device deployment and calculated the theoretical maximum achievable data rate taking into account variations in spectrum availability, the raised noise floor, and interference from other secondary devices.

In Chapter 3, we showed that the current FCC regulations do not adequately account for aggregate interference from secondary devices to incumbent systems. In response, we proposed a candidate rule for safely adjusting the power limit of secondary devices outside of a station's protected region. The primary goal for any power scaling rule is and should be to protect the incumbent services; the secondary goal is increasing utility for whitespace users. The candidate rule we proposed, while only one of many possibilities, achieves both of these goals and approximately equalizes potential data rates across the country despite the extreme variation in spectrum availability and noise levels.

Chapter 4 takes this goal of equality one step further and explicitly seeks to dynamically set power levels in order to provide uniform quality of service to secondary devices. We argued that rules that aim to be both fair and spectrally efficient need to be both frequency- and spatially-aware. We showed that either of these qualities alone is insufficient but that together they are quite potent. These results have implications for database, regulation, and spectrum market design.

Chapter 5 tackles the larger question of efficient allocation of spectrum. In particular, it compares reallocating TV channels for unlicensed use with the legalization of whitespace spectrum use (with variable incumbent protection criteria). In all secondary deployment scenarios that we studied, whitespaces represented the best tradeoff if the incumbent service is highly valued. If secondary services are valued more highly than the incumbents', reallocation is the better strategy since much higher data rates are possible without the raised noise floor.

Chapter 6 took this analysis a step further and opportunistically used data released in connection with the upcoming incentive auctions to examine the tradeoff between clearing spectrum and enabling the use of whitespaces. This chapter compares a naive partial reallocation method (nearly identical to the one used in Chapter 5) and an efficient partial reallocation method ("repacking"). While naive reallocation affects incumbent services in many geographic areas, efficient reallocation confines the effect on incumbent services to the "pinch points" which are typically major metropolitan areas. We also showed that the total amount of spectrum available for new uses is relatively insensitive to how incumbents are removed and that efficient repackings basically trade whitespace spectrum for cleared spectrum. However, even the most efficient repackings leave plenty of whitespace spectrum.

In Chapter 7, we switched gears slightly and leveraged our experience in the field to rethink some aspects of the whitespace ecosystem that many take for granted. While most assume that the abilities to geolocate and to contact a whitespace database will reside within whitespace devices themselves, we challenge this paradigm and show that doing so can lead to significant gains. In particular, we show how a small paradigm shift can help solve the problem of enabling whitespace access for indoor devices, currently an open problem since the de facto localization technology, GPS, notoriously does not work indoors. While our proposed architecture allows for much greater flexibility than is seen in regulations today, it is also backwards compatible and supports today's architectures. Furthermore, it provides several advantages in terms of simplifying certification procedures and developing new regulations.

Chapter 8 further tackles the problem of weakly-localized devices, i.e. devices that do not possess their own localization technology or whose technology is nonstandard (i.e. not GPS). We showed that while the FCC's approach to supporting these devices is too lax, Ofcom's is unnecessarily restrictive. We proposed an enhancement to Ofcom's approach which leverages existing capabilities of secondary devices to safely allow them to access a significant fraction of the actually-available whitespace despite not knowing their location. Although we believe that our proposed enhancement is promising, its true purpose is as a proof of concept. We strongly believe that "localization" services can take many forms and constraining localization to a narrow definition, as is done in the US and the UK today, unnecessarily limits the potential for innovation.

While the TV whitespaces are exceedingly interesting and present many challenges, the future of spectrum sharing relies on promoting dynamic spectrum access in other bands as well. Chapter 9 explored the technical and political hurdles involved in sharing carrier-held GSM spectrum with operators of community cellular networks (CCNs). We proposed Nomadic GSM, a spectrum sharing technology that leverages existing capabilities of GSM phones and the GSM protocol to safely share GSM spectrum with incumbents. We also argued that incumbent carriers should receive universal service obligation (USO) credits when their spectrum is used by a CCN, thus providing incentives to share. Not only is our proposal attractive to all involved parties, but it is also deployable today and indeed has already been deployed in rural Indonesia.

Computation has played an invaluable role in the results presented in this thesis. As such, we feel it is a key part of responsible research to publish our entire code base online for others to freely

view and use. Two distinct but similar code bases were described in Chapters 10 and 11. The first grew organically while the other was carefully designed to maximize flexibility for use in future research projects. The first, written in Matlab, was used to generate most of the results in this thesis. The second, in Python, was used in Chapter 6. In Chapter 11 we described the design of this second code base and used it to quantify the whitespace opportunity in Canada and Australia using both FCC and Industry Canada whitespace regulations. We showed that, with the necessary data in hand, it can take less than an hour to evaluate the whitespace opportunity (using an existing ruleset) in a new jurisdiction.

This thesis presented many novel results in the field of dynamic spectrum sharing, including two which were presented in award-winning papers at the premier conference in the field. Throughout we took the approach of quantifying the effects of regulatory decisions and backed up our claims with analysis of real-world data. We firmly believe that this is an extremely useful technique that enhances our collective understanding and have therefore released all of our code online so that future research can build upon it.

Chapter 13

Future work

Spectrum management is a field that is undergoing enormous change. The first US spectrum auction occurred in 1994 and only 14 years later we were already so pressed for spectrum that the FCC opened up the television whitespaces for unlicensed use. In some sense, whitespaces are in their infancy: current regulations are very primitive and their adoption is nowhere near worldwide. However, they are gaining incredible momentum with each passing year: 2015 alone added several countries to the ranks of those that have opened up their whitespaces.

13.0.1 Evolving regulations

At the same time, the United States, with the most established whitespace rules of all, is working on forward-looking revisions to their regulations such as more flexible notions of localization and variable power limits [40]. These changes are encouraging and are likely due at least in part to the recent UK regulations which already incorporate some of these changes. Despite this progress, there is still a long way to go, in the US and the rest of the world.

For example, although the FCC is proposing in [40] to support a wider variety of device power limits, the implementation remains primitive. Rather than give a formula by which one can compute the maximum permitted power as a function of the distance to the nearest protected contour, the FCC proposes to provide a lookup table with discrete values. The approach is similar in other regulatory domains. Although this may reduce the burden on whitespace databases, it does not support a flexible approach that would e.g. allow a device to substantially reduce its power in exchange for greater spectrum availability. Similar to the work in Chapter 4, WEST could be used to quantify the impact of one approach over another.

Existing regulations also do not realistically address the problem of aggregate interference. As shown in Chapter 3, aggregate interference can pose a significant problem for incumbents, especially if whitespace devices become popular. In order to address this problem, whitespace databases would need to have an approximate record of which devices are currently using which spectrum and adjust responses to future spectrum access requests accordingly.

Similarly, existing regulations do not attempt to balance the needs of rural vs. urban users. As discussed in Chapter 4, rural areas desperately need higher emissions limits while urban areas need more spectrum. A one-size-fits-all approach to emissions limits cannot adequately provide for the needs of both types of users, and instead the emissions limits and separation distances should be specified on a per-channel basis. Finding the correct balance between both types of users, as well as providing enough certainty and consistence for device manufacturers, is an area which requires extensive work. WEST is particularly well-suited to this kind of research.

Whitespace databases today are expected to return a single "correct" answer[1]. In extreme cases the correctness of one database's response may be checked against the response of another database. Although this will guarantee that the databases return identical (and presumably correct) responses, it leaves no room for innovation and differentiation between databases. For example, one database may choose to return identical responses to all devices querying from the same location. Another database may attempt to provide a superior experience to its users by attempting to "steer" them to channels which are expected to have lower levels of interference (e.g. from incumbents or other secondary devices) via returning only a *subset* of the truly available whitespace channels [148]. Although this is clearly safe from the incumbent's perspective, it would be difficult for such a database to pass today's certification procedures.

In a similar vein, whitespace regulations today conflate two distinct problems: localization of a device and determining which whitespace channels are available for it to safely use. This was discussed extensively in Chapter 8 but the concept has no regulatory support. I strongly believe that this approach has a lot to offer the whitespace ecosystem but the concepts still need to be socialized among regulators since certification procedures would need to be adapted. WEST can assist in this process by allowing stakeholders to quantify the benefits of permitting additional localization methods, similar to what was done in Chapter 8.

In fact, the problem of granting whitespace access to a non-geolocated device is quite difficult in the current regulations. For example, Ofcom in the UK allows a geolocated device to broadcast "generic operating parameters" which are meant to be safe for all devices operating within its service area. The permissiveness of these parameters depends heavily on the size of the service area and in some cases they are so restrictive that slaves which are in the master's service area are unable to communicate back to the master, even if there is actually usable whitespace available at the slave. Confronted with this situation, the master device may naturally choose to reduce its service area until the corresponding generic operating parameters are permissive enough to allow any slave within its service area to reply. This artificial reduction in service area is due in large part to the uncertainty of the slave device's location. Again, the magnitude of problems such as these can be quantified using WEST.

Finally, although innovation within the regulations is critical, it is just as important that the regulations themselves not become too fragmented. Wildly diverse regulations mean that it may

---

[1]See e.g. http://transition.fcc.gov/oet/whitespace/guides/TVWS_Database_System_Approval_Process3.pdf.

be difficult to build devices which can operate in more than one market, substantially increasing the price of devices and/or reducing the proliferation of whitespace devices. Figuring out how to best strike the delicate balance between diverse regulations and regulatory harmonization is very practical and important work that will itself require a diversity of backgrounds: researchers, industry members, standards body members, and regulators.

13.0.2 Sharing in other bands

Perhaps even more exciting than the proliferation of TV whitespaces to new jurisdictions is the fact that some of them are looking into opening up whitespaces in other bands as well. For example, the United States recently opened the 3.5 GHz band [27] and is considering opening the 5 GHz band [2]. Spectrum sharing in these bands presents more of a challenge than in the TV whitespaces due to the nature of the incumbents. These federally-held bands are used for ship- and air-borne radar systems which are critical to the operation of the US's Department of Defense both at home and abroad. Unlike TV stations, radar systems are mobile and have highly-variable usage patterns.

Although there are ongoing discussions about sharing military radar bands, there is a plethora of unanswered questions that remain. Many of these questions were raised at the Wireless Spectrum Research and Development (WSRD) workshop in October 2014[2].

- How will secondary devices recognize that a particular (frequency, time, location) triplet is protected? There are two leading options for this, each with their own set of follow-on questions:

  - **Database**: under this model, the entity responsible for incumbent operations would provide information on current and future incumbent locations to a database provider. The database may be hosted by the responsible (federal) entity or by a third party, as is done in the TV whitespaces.

    * What information can be stored in the database? For example, federal security procedures may prevent sharing information outside of federal systems.
    * What information needs to be stored in the database? Is there a way to trade off spectrum availability with the amount of data that is required? What does this tradeoff look like?
    * Should information stored in the database be obfuscated and/or secured in some way? If so, how?
    * How can we protect against repeated-query attacks that attempt to learn the contents of the database through legitimate-looking queries?

---

[2]Further information about this workshop can be found at `https://www.nitrd.gov/nitrdgroups/index.php?title=Federal_-_Commercial_Spectrum_Data`. The workshop report was authored in large part by myself and Prof. Anant Sahai.

– **Sensing**: devices operating in other parts of the 5 GHz band today employ DFS (dynamic frequency selection) in which the device senses for the incumbent before commencing operation. While some argue that this is sufficient for these bands as well, others propose the introduction of dedicated sensing devices which report any detection of incumbent activity to nearby whitespace devices.

* What are the benefits and drawbacks of on-device vs. dedicated sensing? Is either (or both) sufficient to protect the incumbent? What set of permissions correspond to each proposed solution? What impact does each solution have on the type of device that can be deployed in the band?

* Should the sensing be threshold- or pattern-based? What is the effect of each on the false-positive rate? How does this false-positive rate affect device operation?

* What should the sensing threshold be?

* Can radar signatures for classified military systems be released for this limited purpose?

* How can future radar systems be protected? How would on-board or dedicated sensing equipment be updated when new radar systems are deployed?

• What is a reasonable channel clearing time[3]?

• What harm do secondary devices realistically pose to radar systems? What harm do radar systems realistically pose to secondary devices?

• How can secondary use be audited and enforced?

• How much whitespace could be made available with aggressive regulations? What is the tradeoff between recoverable spectrum and assurance of primary protection?

• What kinds of devices are well-suited to sharing with radar systems?

These questions and many more will be tackled in the coming years by researchers, industry members, and regulators alike. Proponents of whitespaces in these bands hope to secure a study item in the World Radiocommunication Conference 2015 (WRC-15) so that sharing will be studied at an international level, with the goal of an international sharing decision at WRC-19.

I believe that we are moving toward a world where all spectrum is dynamically shared. Opening up the TV whitespaces and considering doing the same for federally-held bands is a great step along that path. But we must continually ask ourselves: what other bands are good candidates for spectrum sharing? What unique challenges do the various incumbents pose for protection? How can regulators, industry members, and researchers work together to promote dynamic spectrum access?

---

[3]The channel clearing time is the time between when a radar pulse is detected and the affected device(s) cease transmitting.

13.0.3 Sharing with other devices

The rapid increase of available unlicensed spectrum poses other interesting problems. For example, a variety of companies are developing a new cellular phone standard called LTE-U, a variant of LTE meant to augment licensed spectrum with unlicensed bands. With the development LTE-U come many questions, such as:

- Is it "fair" for LAA (licensed-assisted access) devices to operate in the unlicensed bands?

- What does fair sharing in unlicensed bands mean or look like?

- What impact would LTE-U devices have on existing applications such as WiFi?

- What if they crowd out these existing applications? Is this an acceptable situation? If not, how can it be avoided?

Companies such as Qualcomm have done studies[4], but not everyone in the community is convinced. Answering the questions above will be key to the future of unlicensed spectrum.

13.0.4 Future work on WEST

So far we have concentrated on the aspects of whitespace research that can be helped by WEST. However, WEST itself has room to grow. Some of these opportunities were identified in Section 11.7. We discuss a few higher-level points here.

First and foremost, WEST need users. Each user brings with her a unique perspective on the problems in spectrum sharing and will exercise the code in different ways. Although we have used WEST extensively ourselves, we can only design for the types of problems that we encounter. Having more users means that the existing code in WEST will be vetted and, hopefully, that new code will be added. Support for additional rulesets, regions, propagation models, etc. will greatly increase WEST's versatility and utility.

WEST may not be a one-size-fits-all solution to every research problem—and it's not meant to be. One of the key ideas behind WEST is to introduce researchers to data-driven work in a low-cost-high-reward manner. We hope that once the value of this data-driven approach is better understood, others will create similar packages that solve different sets of problems or solve the same problems in different ways. We look forward to a day when WEST is one of many tools freely available to stakeholders.

---

[4]https://www.qualcomm.com/invention/technologies/1000x/spectrum/unlicensed
supposedly showing that LTE-U "is a better neighbor to WiFi than WiFi itself"

Bibliography

[1] Federal Communications Commission, "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Second Report and Order and Memorandum Opinion and Order," Tech. Rep. 08-260, Nov. 2008. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-260A1.pdf

[2] ——, "Revision of Part 15 of the Commission's Rules to Permit Unlicensed National Information Infrastructure (U-NII) Devices in the 5 GHz Band: Notice of Proposed Rulemaking," Feb. 2013. [Online]. Available: http://www.fcc.gov/document/5-ghz-unlicensed-spectrum-unii

[3] ——, "In the Matter of Amendment of the Commission's Rules with Regard to Commercial Operations in the 3550-3650 MHz Band: Further Notice of Proposed Rulemaking," Apr. 2014. [Online]. Available: http://www.fcc.gov/document/proposes-creation-new-citizens-broadband-radio-service-35-ghz

[4] K. Heimerl, S. Hasan, K. Ali, P. Bloom, J. Blumenstock, E. Brewer, and T. Parikh, "Designing Community Cellular Networks," in *In Submission*.

[5] K. Heimerl, S. Hasan, K. Ali, E. Brewer, and T. Parikh, "Local, Sustainable, Small-Scale Cellular Networks," in *Proceedings of the Sixth International Conference on Information and Communication Technologies and Development*, ser. ICTD '13. Cape Town, South Africa: ACM, 2013.

[6] K. Harrison, "Whitespace evaluation code." [Online]. Available: https://github.com/kate-harrison/whitespace-eval

[7] ——, "Whitespace Evaluation SofTware (WEST)." [Online]. Available: https://github.com/kate-harrison/west

[8] [Online]. Available: http://www.ntia.doc.gov/osmhome/allochrt.html

[9] Federal Communications Commission, "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Second Memorandum Opinion and Order," Tech. Rep. 10-174A1, Sep. 2010. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-174A1.pdf

[10] [Online]. Available: http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml

[11] [Online]. Available: http://www.census.gov/geo/www/cen_tract.html

[12] K. Harrison and A. Sahai, "Potential collapse of whitespaces and the prospect for a universal power rule," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 316–327.

[13] Federal Communications Commission, "In the Matter of Expanding the Economic and Innovation Opportunities of Spectrum Through Incentive Auctions: Report and Order," Jun. 2014. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/FCC-14-50A1.pdf

[14] D. Setiawan, "Overview of Spectrum Issues in Indonesia," http://denysetia.files.wordpress.com/2010/09/290610-broadband-wireless-gsma-london-denny-ins.ppt, June 2010.

[15] A. De Vany, R. Eckert, C. Meyers, and D. O'Hara, "Property system for market allocation of the electromagnetic spectrum: A legal-economic-engineering study, a," *Stan. L. Rev.*, vol. 21, p. 1499, 1968.

[16] Barnouw, Erik: "A Tower in Babel", New York, Oxford University Press, 1966.

[17] U.S.C. Section 151.

[18] A. Goldsmith, *Wireless communications*. Cambridge Univ Pr, 2005.

[19] "Advanced Wireless Services (AWS-3) Auction Summary," 2015. [Online]. Available: http://wireless.fcc.gov/auctions/default.htm?job=auction_summary&id=97

[20] Federal Communications Commission, "National Broadband Plan," 2010. [Online]. Available: http://www.fcc.gov/national-broadband-plan

[21] T. Taher, R. Bacchus, K. Zdunek, and D. Roberson, "Long-term spectral occupancy findings in Chicago," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 100–107.

[22] [Online]. Available: http://bwrc.eecs.berkeley.edu/Research/MCMA/CR_White_paper_final1.pdf

[23] Federal Communications Commission, "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Third Memorandum Opinion and Order," Tech. Rep. 12-36, Apr. 2012. [Online]. Available: http://www.fcc.gov/document/tv-white-spaces-rule-changes

[24] "White Space Database Specifications," Feb. 2015. [Online]. Available: http://www.ic.gc.ca/eic/site/smt-gst.nsf/eng/sf10928.html

[25] "Implementing TV White Spaces," Feb. 2015. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/white-space-coexistence/statement/tvws-statement.pdf

[26] Federal Communications Commission, "Revision of Part 15 of the Commission's Rules to Permit Unlicensed National Information Infrastructure (U-NII) Devices in the 5 GHz Band: First Report and Order," Mar. 2014. [Online]. Available: http://www.fcc.gov/document/5-ghz-u-nii-ro

[27] ——, "In the Matter of Amendment of the Commission's Rules with Regard to Commercial Operations in the 3550-3650 MHz Band: Report and Order and Second Further Notice of Proposed Rulemaking," Apr. 2015. [Online]. Available: http://www.fcc.gov/document/citizens-broadband-radio-service-ro

[28] [Online]. Available: transition.fcc.gov/mb/video/tvq.html

[29] "Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz," International Telecommunications Commission (ITU), Recommendation ITU-R P.1546-3, 2007.

[30] K. Harrison, S. Mishra, and A. Sahai, "How much white-space capacity is there?" in *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on*. IEEE, 2010, pp. 1–10.

[31] J. van de Beek, J. Riihijarvi, A. Achtzehn, and P. Mahonen, "UHF white space in Europe - a quantitative study into the potential of the 470–790 MHz band," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 1–9.

[32] R. Jantti, J. Kerttula, K. Koufos, and K. Ruttik, "Aggregate interference with FCC and ECC white space usage rules: Case study in Finland," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 599–602.

[33] M. Barbiroli, C. Carciofi, D. Guiducci, and V. Petrini, "White spaces potentially available in italian scenarios based on the geo-location database approach," in *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 416–421.

[34] D. Makris, G. Gardikis, and A. Kourtis, "Quantifying tv white space capacity: quantifying tv white space capacity," *Communications Magazine, IEEE*, vol. 50, no. 9, pp. 145–152, 2012.

[35] A. Achtzehn, M. Petrova, and P. Mähönen, "Deployment of a cellular network in the tvws: A case study in a challenging environment," in *Proceedings of the 3rd ACM workshop on Cognitive radio networks*. ACM, 2011, pp. 7–12.

[36] B. Freyens and M. Loney, "Opportunities for white space usage in Australia," in *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology (Wireless VITAE), 2011 2nd International Conference on*. IEEE, 2011, pp. 1–5.

[37] "Geographic maps showing TVWS availability," Sep. 2013. [Online]. Available: http://stakeholders.ofcom.org.uk/consultations/white-space-coexistence/maps/

[38] K. Harrison, "Cognitive radios in the TV whitespaces: challenges and opportunities," 2011.

[39] S. Hasan, K. Heimerl, K. Harrison, K. Ali, S. Roberts, A. Sahai, and E. Brewer, "GSM whitespaces: An opportunity for rural cellular service," in *Dynamic Spectrum Access Networks (DySpAN), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 271–282.

[40] Federal Communications Commission, "Amendment of Part 15 of the Commission's Rules for Unlicensed Operations in the Television Bands, Repurposed 600 MHz Band, 600 MHz Guard Bands and Duplex Gap, and Channel 37: Notice of Proposed Rulemaking," Sep. 2014. [Online]. Available: http://www.fcc.gov/document/part-15-nprm

[41] [Online]. Available: http://www.fcc.gov/encyclopedia/tv-service-contour-data-points

[42] C. Shannon and W. Weaver, *The mathematical theory of communication*. University of Illinois Press Urbana, 1962, vol. 19, no. 9.

[43] M. Costa, "Writing on dirty paper (corresp.)," *Information Theory, IEEE Transactions on*, vol. 29, no. 3, pp. 439–441, 1983.

[44] H. Wu and Y. Pan, *Medium Access Control in Wireless Networks*. Nova Science Pub Inc, 2008.

[45] Y. Wu, E. Pliszka, B. Caron, P. Bouchard, and G. Chouinard, "Comparison of terrestrial DTV transmission systems: the ATSC 8-VSB, the DVB-t COFDM, and the ISDB-t BST-OFDM," *Broadcasting, IEEE Transactions on*, vol. 46, no. 2, pp. 101–113, 2000.

[46] "Technical and operational requirements for the possible operation of cognitive radio systems in the white spaces of the frequency band 470-790 MHz, ECC Report 159." [Online]. Available: http://www.erodocdb.dk/docs/doc98/official/Pdf/ECCRep159.pdf

[47] M. Mishra and A. Sahai, "How much white space is there?" *EECS Department, University of California, Berkeley, Tech. Rep. UCB/EECS-2009-3, Jan*, 2009.

[48] S. Mishra and A. Sahai, "How much white space has the FCC opened up?" *IEEE Communication Letters*, 2010.

[49] H. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 443–454.

[50] N. Hoven, "On the feasibility of cognitive radio," *Master's thesis, University of California, Berkeley*, 2005.

[51] K. Harrison and A. Sahai, "Seeing the bigger picture: Context-aware regulations," in *Dynamic Spectrum Access Networks (DySpAN), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 21–32.

[52] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner, and J. Grosspietsch, "Geo-location database techniques for incumbent protection in the tv white space," in *New Frontiers in Dynamic Spectrum Access Networks, 2008. DySPAN 2008. 3rd IEEE Symposium on*. IEEE, 2008, pp. 1–9.

[53] M. Kokar and L. Lechowicz, "Language issues for cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 689 –707, april 2009.

[54] President's Council of Advisors on Science and Technology, "Realizing the full potential of government-held spectrum to spur economic growth," Tech. Rep., 2012. [Online]. Available: http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast_spectrum_report_final_july_20_2012.pdf

[55] M. Nekovee, "Quantifying the availability of TV white spaces for cognitive radio operation in the UK," in *Communications Workshops, 2009. ICC Workshops 2009. IEEE International Conference on*. IEEE, 2009, pp. 1–5.

[56] M. Ma and D. Tsang, "Impact of channel heterogeneity on spectrum sharing in cognitive radio networks," in *Communications, 2008. ICC'08. IEEE International Conference on*. IEEE, 2008, pp. 2377–2382.

[57] H. Zheng and C. Peng, "Collaboration and fairness in opportunistic spectrum access," in *Communications, 2005. ICC 2005. 2005 IEEE International Conference on*, vol. 5. IEEE, 2005, pp. 3132–3136.

[58] M. Kaplan and F. Buzluca, "A dynamic spectrum decision scheme for heterogeneous cognitive radio networks," in *Computer and Information Sciences, 2009. ISCIS 2009. 24th International Symposium on*. IEEE, 2009, pp. 697–702.

[59] K. Woyach, P. Grover, and A. Sahai, "Near vs far field: interference aggregation in TV whitespaces," 2011.

[60] "Implementing Geolocation: Summary of consultation responses and next steps," Sep. 2011. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/geolocation/statement/statement.pdf

[61] V. Muthukumar, A. Daruna, V. Kamble, K. Harrison, and A. Sahai, "Whitespaces after the USA's TV Incentive Auction: A Spectrum Reallocation Case Study," in *2015 IEEE International Conference on Communications*, 2015.

[62] M. Vu, N. Devroye, and V. Tarokh, "On the primary exclusive region of cognitive networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 7, 2009.

[63] J. Bater, H. Tan, K. Brown, and L. Doyle, "Modelling interference temperature constraints for spectrum access in cognitive radio networks," in *Communications, 2007. ICC'07. IEEE International Conference on*.   IEEE, 2007, pp. 6493–6498.

[64] "TV Broadcasters Sue Over FCC Auction of Airwaves." [Online]. Available: http://online.wsj.com/articles/tv-broadcasters-sue-over-fcc-auction-of-airwaves-1408402394

[65] Federal Communications Commission. (2014) Constraint files for repacking. [Online]. Available: http://data.fcc.gov/download/incentive-auctions/Constraint_Files/

[66] E. Kwerel and J. Williams, "A Proposal for a Rapid Transition to Market Allocation of Spectrum," *OPP Working Paper Series No. 38, Federal Communications Commission, Office of Plans and Policy*, 2002.

[67] "Middle Class Tax Relief and Job Creation Act of 2012," Feb. 2012. [Online]. Available: https://beta.congress.gov/bill/112th-congress/house-bill/3630

[68] Federal Communications Commission, "In the Matter of Expanding the Economic and Innovation Opportunities of Spectrum Through Incentive Auctions: Notice of Proposed Rulemaking," Oct. 2012. [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/FCC-12-118A1.pdf

[69] ——, "Comment Sought on Competitive Bidding Procedures for Broadcast Incentive Auction 1000, Including Auctions 1001 and 1002," Dec. 2014. [Online]. Available: http://www.fcc.gov/document/broadcast-incentive-auction-comment-pn

[70] M. Kearns and L. Dworkin, "A Computational Study of Feasible Repackings in the FCC Incentive Auctions," *CoRR*, vol. abs/1406.4837, 2014. [Online]. Available: http://arxiv.org/abs/1406.4837

[71] T. Hazlett, D. Porter, and V. Smith. (2012) Incentive Auctions - Economic and Strategic Issues. [Online]. Available: http://arlingtoneconomics.com/studies/WhitePaper.pdf

[72] T-Mobile. (2014) 600 MHz Repacking Analysis. [Online]. Available: http://apps.fcc.gov/ecfs/document/view?id=7521831699

[73] CTIA, "Broadcast Spectrum Incentive Auctions whitepaper," 2011. [Online]. Available: http://files.ctia.org/pdf/CTIA_CEA_TV_Spectrum_Whitepaper_Summary.pdf

[74] K. Leyton-Brown. (2013) Investigating the viability of exact feasibility testing. [Online]. Available: http://www.cs.ubc.ca/~kevinlb/talk.php?u=2013-Feasibility-Testing.pdf

[75] Federal Communications Commission. (2014) Incentive Auctions - LEARN. [Online]. Available: http://fcc.gov/learn

[76] "LEARN Workshop on Feasibility Checking During Repacking Process," 2014. [Online]. Available: http://www.fcc.gov/events/learn-workshop-feasibility-checking-during-repacking-process

[77] A. Biere, "PicoSAT." [Online]. Available: http://fmv.jku.at/picosat

[78] R. Tandra, A. Sahai, and S. Mishra, "What is a spectrum hole and what does it take to recognize one?" *Proceedings of the IEEE*, vol. 97, no. 5, pp. 824–848, 2009.

[79] K. Harrison and A. Sahai, "A more general whitespace architecture: refactoring the master-client paradigm," in *2015 IEEE International Conference on Communications*, 2015.

[80] P. A. Zandbergen, "Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning," *Transactions in GIS*, vol. 13, no. s1, pp. 5–25, 2009.

[81] E. Paulos, R. Honicky, and B. Hooker, "Citizen science: Enabling participatory urbanism," *Handbook of Research on Urban Informatics*, pp. 414–436, 2008.

[82] B. Liskov and J. Guttag, *Program development in JAVA: abstraction, specification, and object-oriented design*. Pearson Education, 2000.

[83] Jun. 2012, fCC test procedure for whitespace devices. [Online]. Available: https://apps.fcc.gov/oetcf/kdb/forms/FTSSearchResultPage.cfm?switch=P&id=50929

[84] J. H. Saltzer, D. P. Reed, and D. D. Clark, "End-to-end arguments in system design," *ACM Transactions on Computer Systems (TOCS)*, vol. 2, no. 4, pp. 277–288, 1984.

[85] "TV white spaces: A consultation on white space device requirements," Nov. 2012. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/whitespaces/summary/condoc.pdf

[86] K. Harrison and A. Sahai, "Allowing sensing as a supplement: an approach to the weakly-localized whitespace device problem," in *Dynamic Spectrum Access Networks (DySpAN), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 113–124.

[87] "TV white spaces: A consultation on white space device requirements," Nov. 2012. [Online]. Available: http://stakeholders.ofcom.org.uk/binaries/consultations/whitespaces/summary/condoc.pdf

[88] M. B. Weiss, S. Delaere, and W. H. Lehr, "Sensing as a service: An exploration into practical implementations of dsa," in *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on*. IEEE, 2010, pp. 1–8.

[89] "Comments from the Communications Research Centre Canada: In the Matter of Unlicensed Operation in the TV Broadcast Bands," http://apps.fcc.gov/ecfs/document/view?id=6518724125.

[90] S. M. Mishra, "Maximizing available spectrum for cognitive radios," Ph.D. dissertation, University of California, Berkeley, 2009.

[91] [Online]. Available: http://sutrotower.com/tower-technology/

[92] G. Giorgetti, S. K. Gupta, and G. Manes, "Wireless localization using self-organizing maps," in *Proceedings of the 6th international conference on Information processing in sensor networks*.   ACM, 2007, pp. 293–302.

[93] L. Doherty, L. El Ghaoui *et al.*, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3.   IEEE, 2001, pp. 1655–1663.

[94] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'dea, "Relative location estimation in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 8, pp. 2137–2148, 2003.

[95] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *IEE Proceedings-Communications*, vol. 152, no. 5, pp. 580–586, 2005.

[96] J. Riihijarvi, J. Nasreddine, and P. Mahonen, "Demonstrating radio environment map construction from massive data sets," in *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*.   IEEE, 2012, pp. 266–267.

[97] R. Tandra and A. Sahai, "SNR walls for signal detection," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 4–17, 2008.

[98] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 71–82.

[99] V. Atanasovski, J. van de Beek, A. Dejonghe, D. Denkovski, L. Gavrilovska, S. Grimoud, P. Mahonen, M. Pavloski, V. Rakovic, J. Riihijarvi *et al.*, "Constructing radio environment maps with heterogeneous spectrum sensors," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*.   IEEE, 2011, pp. 660–661.

[100] L. Waverman, M. Meschi, and M. Fuss, "The Impact of Telecoms on Economic Growth in Developing Countries," *Africa: the Impact of Mobile Phones*, vol. 2, no. 1980, pp. 1–23, 2005. [Online]. Available: https://www.businessgrowthinitiative.org/KeyPracticeAreas/Documents/Forms/DispForm.aspx?ID=148

[101] R. Jensen, "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector," *The Quarterly Journal of Economics*, vol. 122, no. 3, pp. 879–924, 08 2007.

[102] A. Agüero, H. de Silva, and J. Kang, "Bottom of the Pyramid Expenditure Patterns on Mobile Services in Selected Emerging Asian Countries," *Information Technologies & International Development*, vol. 7, no. 3, 2011. [Online]. Available: http://itidjournal.org/itid/article/view/758

[103] M. Spiegel, "Universal service obligation," in *Expanding Competition in Regulated Industries*, ser. Topics in Regulatory Economics and Policy Series, M. A. Crew, Ed. Springer US, 2000, vol. 37, pp. 83–96. [Online]. Available: http://dx.doi.org/10.1007/978-1-4757-3192-7_5

[104] M. E. Whitman, "A Look at the Telecommunications ACT of 1996." *IS Management*, vol. 14, no. 3, pp. 82–86, 1997. [Online]. Available: http://dblp.uni-trier.de/db/journals/ism/ism14.html#Whitman97

[105] I. T. Union, "Implementation of USO Program in Indonesia," http://www.itu.int/ITU-D/.../Indonesia/Session2-BP3TI.pdf, March 2011.

[106] C. R. Blackman, "Universal Service: Obligation or Opportunity?" *Telecommunications Policy*, vol. 19, no. 3, pp. 171–176, 1995. [Online]. Available: http://EconPapers.repec.org/RePEc:eee:telpol:v:19:y:1995:i:3:p:171-176

[107] P. Chon, L. Flochel, and A. Perrot, "Allocating and Funding Universal Service Obligations in a Competitive Market," *International Journal of Industrial Organization*, vol. 20, no. 9, pp. 1247 – 1276, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167718701000777

[108] D. Weller, "Auctions for Universal Service Obligations," *Telecommunications Policy*, vol. 23, no. 9, pp. 645 – 674, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0308596199000488

[109] A. Joshi and M. Moore, "Institutionalised Co-production: Unorthodox Public Service Delivery in Challenging Environments," *The Journal of Development Studies*, vol. 40, no. 4, pp. 31–49, 2004. [Online]. Available: http://ideas.repec.org/a/taf/jdevst/v40y2004i4p31-49.html

[110] W. F. Lam, "Institutional Design of Public Agencies and Coproduction: A Study of Irrigation Associations in Taiwan," *World Development*, vol. 24, no. 6, pp. 1039–1054, June 1996. [Online]. Available: http://ideas.repec.org/a/eee/wdevel/v24y1996i6p1039-1054.html

[111] C. Fischer, *America Calling: A Social History of the Telephone to 1940*. Berkeley, CA: University of California Press, 1992.

[112] Q. Alam, M. A. Yusuf, and K. Coghill, "Village Phone Program, Commodification Of Mobile Phone Set And Empowerment Of Women," *J. Community Informatics*, vol. 5, no. 3-4, 2009.

[113] H. Galperin and F. Bar, "The Microtelco Opportunity: Evidence from Latin America," in *Information Technologies and International Development*, vol. 3, no. 2, Winter 2006.

[114] T. Egyedi and D. Mehos, *Inverse Infrastructures: Disrupting Networks from Below*. Edward Elgar, 2012. [Online]. Available: http://books.google.com/books?id=wG4c7s8uKjUC

[115] OpenBTS, http://openbts.org, Retrieved 7/2013.

[116] Rhizomatica, http://rhizomatica.org/, Retrieved 4/2013.

[117] M. Zheleva, A. Paul, D. L. Johnson, and E. Belding, "Kwiizya: Local Cellular Network Services in Remote Areas," in *Proceedings of the 11th International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '13. Taipai, Taiwan: ACM, 2013.

[118] S. Sankaranarayanan, P. Papadimitratos, A. Mishra, and S. Hershey, "A Bandwidth Sharing Approach to Improve Licensed Spectrum Utilization," in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*. IEEE, 2005, pp. 279–288.

[119] M. M. Buddhikot, I. Kennedy, F. Mullany, and H. Viswanathan, "Ultra-Broadband Femtocells via Opportunistic Reuse of Multi-Operator and Multi-Service Spectrum," *Bell Labs Technical Journal*, vol. 13, no. 4, pp. 129–143, 2009.

[120] L. Yin, K. Wu, S. Yin, S. Li, and L. M. Ni, "Reuse of GSM White Space Spectrum for Cognitive Femtocell Access," in *Parallel and Distributed Systems (ICPADS), 2012 IEEE 18th International Conference on*. IEEE, 2012, pp. 1–7.

[121] K. Heimerl, K. Ali, J. Blumenstock, B. Gawalt, and E. Brewer, "Expanding Rural Cellular Networks with Virtual Coverage," in *Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation*, ser. NSDI'13. Lombard, Illinois: USENIX Association, 2013.

[122] GSM Association, "Powering Telecoms: West Africa Market Analysis Sizing the Potential for Green Telecoms in Nigeria and Ghana," http://www.gsma.com/mobilefordevelopment/wp-content/uploads/2013/03/GPM-Market-Analysis-West-Africa-.pdf, retrieved 10/2013.

[123] Telkomsel, "Telkomsel, 2006 Annual Report," http://www.telkomsel.com/media/upload/pdf/AR2006.pdf, 2007.

[124] Sensorly, http://sensorly.com, Retrieved 11/2013.

[125] Telkomsel, "Telkomsel, 2012 Annual Report," http://www.telkomsel.com/media/upload/annualreport/%28FINAL%29%20AR%20TELKOMSEL%20CONTENT.pdf, February 2013.

[126] GSMA, "Mobile Development Intelligence, Network Coverage Population," https://mobiledevelopmentintelligence.com.

[127] V. Bharghavan, A. Demers, S. Shenker, and L. Zhang, "MACAW: a Media Access Protocol for Wireless LAN's," in *Proceedings of the Conference on Communications Architectures, Protocols and Applications*, ser. SIGCOMM '94. New York, NY, USA: ACM, 1994, pp. 212–225. [Online]. Available: http://doi.acm.org/10.1145/190314.190334

[128] J. Peha, "Sharing spectrum through spectrum policy reform and cognitive radio," *Proceedings of the IEEE*, vol. 97, no. 4, pp. 708–719, 2009.

[129] 3GPP, "Radio subsystem link control," 3rd Generation Partnership Project (3GPP), TS 05.08, Nov. 2005. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/0508.htm

[130] D. Cabric, S. M. Mishra, D. Willkomm, R. Brodersen, and A. Wolisz, "A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum," in *In Proc. of 14th IST Mobile Wireless Communications Summit 2005*, 2005.

[131] [Online]. Available: http://www.fcc.gov/encyclopedia/white-space-database-administration

[132] [Online]. Available: http://www.fcc.gov/encyclopedia/fm-and-tv-propagation-curves-graphs-sections-73333-73525-and-73699

[133] "Decision Paper Issued by the Info-Communications Development Authority of Singapore: Regulatory Framework for TV White Space Operations in the VHF/UHF Bands," Jun. 2014. [Online]. Available: http://www.ida.gov.sg/~/media/Files/PCDG/Consultations/20130617_whitespace/ExplanatoryMemo.pdf

[134] "TV White Space Devices Standards and Requirements," Feb. 2015. [Online]. Available: http://news.gc.ca/web/article-en.do?nid=928599

[135] [Online]. Available: https://www.google.com/get/spectrumdatabase/

[136] [Online]. Available: http://whitespaces.spectrumbridge.com/whitespaces/home.aspx

[137] [Online]. Available: http://whitespaces.msresearch.us/WSWebGUI/whitespaces.aspx

[138] T. B. Alemu, "Spectrum availability assessment tool for tv white space," Ph.D. dissertation, Aalto University, 2012.

[139] A. Achtzehn, L. Simic, M. Petrova, P. Mahonen, V. Rakovic, P. Latkoski, and L. Gavrilovska, "Software tool for assessing secondary system opportunities in spectrum whitespaces," in *World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2013 IEEE 14th International Symposium and Workshops on a*. IEEE, 2013, pp. 1–3.

[140] "Licensed broadcasting transmitters." [Online]. Available: http://www.acma.gov.au/Industry/Spectrum/Radiocomms-licensing/Apparatus-licences/list-of-licensed-broadcasting-transmitters

[141] [Online]. Available: http://www.statsilk.com/files/country/StatPlanet_Australia.zip

[142] "CDATA Online, 2006." [Online]. Available: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Lookup/2064.0Main+Features12006

[143] "Census of Population and Housing: Census Geographic Areas Digital Boundaries, Australia, 2006." [Online]. Available: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/2923.0.30.0012006

[144] "Suggested Technical Rules and Regulations for the Use of Television White Spaces." [Online]. Available: http://www.dynamicspectrumalliance.org/assets/submissions/SuggestedTechnicalRulesandRegulationsfortheuseofTVWS.pdf

[145] "Google Spectrum Database Data Download." [Online]. Available: http://www.google.com/get/spectrumdatabase/data/

[146] "Census Profile - Comprehensive download files for a selected geographic level." [Online]. Available: http://www12.statcan.gc.ca/census-recensement/2011/dp-pd/prof/details/download-telecharger/comprehensive/comp-csv-tab-dwnld-tlchrgr.cfm

[147] "2011 Census - Boundary files." [Online]. Available: http://www12.statcan.gc.ca/census-recensement/2011/geo/bound-limit/bound-limit-2011-eng.cfm

[148] Y. Luo, L. Gao, and J. Huang, "Trade information, not spectrum: A novel tv white space information market model," in *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2014 12th International Symposium on*. IEEE, 2014, pp. 405–412.

[149] Greenhill, "Incentive Auction Opportunities for Broadcasters." [Online]. Available: http://wireless.fcc.gov/incentiveauctions/learn-program/docs/ia-opportunities-book.pdf

[150] [Online]. Available: http://www.fcc.gov/encyclopedia/wireless-medical-telemetry-service-wmts

[151] [Online]. Available: https://itunes.apple.com/us/app/micfrequency/id422935696

[152] [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/DA-11-131A1.doc

[153] K. Harrison, V. Muthukumar, V. Kamble, A. Daruna, and A. Sahai, "Appendices for "Whitespaces after the USA's TV incentive auction: a spectrum reallocation case study"," UC Berkeley, Tech. Rep., 2015. [Online]. Available: http://www.eecs.berkeley.edu/~sahai/icc2015-incentive-auction-appendices.pdf

[154] C. Sinz, "Towards an optimal CNF encoding of boolean cardinality constraints," in *Principles and Practice of Constraint Programming-CP 2005*. Springer, 2005, pp. 827–831.

[155] "Google Spectrum Database Data Download." [Online]. Available: http://www.google.com/get/spectrumdatabase/data/

[156] Federal Communications Commission. TV service contour data points. [Online]. Available: http://www.fcc.gov/encyclopedia/tv-service-contour-data-points

[157] ——. Antenna Height Above Average Terrain (HAAT) Calculator. [Online]. Available: http://www.fcc.gov/encyclopedia/antenna-height-above-average-terrain-haat-calculator

[158] ——. FM and TV propagation curves. [Online]. Available: http://www.fcc.gov/encyclopedia/fm-and-tv-propagation-curves

[159] P. R. Milgrom, *Putting auction theory to work*. Cambridge University Press, 2004.

[160] Federal Communications Commission, "Technical Appendix: Constraint Files Generation." [Online]. Available: https://apps.fcc.gov/edocs_public/attachmatch/DA-13-1613A2.pdf

Appendix A

Comments to regulatory bodies

A.1 [FCC] Preliminary findings on the cost of the FCC's upcoming incentive auction

This work was done with significant help from Angel Daruna and Vijay Kamble who created the incentive auction repacker used for these results.

The following was submitted on 10 December 2014 via email to Ira Keltz at the Federal Communications Commission.

A.1.1 Executive Summary

We present preliminary findings which suggest that the 84 MHz band plan is achievable if the forward auction brings in $45 billion, even if broadcasters are compensated at the maximum amounts listed in the Greenhill report [149]. If more modest prices are used, the 114 MHz band plan may be achievable. In these cases, the forward auction winners would need to pay at least $1.50-2 per MHz-pop. For comparison, the average price per MHz-pop in the recent AWS-3 auction was over $2 (and over $2.50 for paired spectrum).

Additionally, we propose a simpler reassignment methodology that fits within the constraints of the incentive auction while increasing certainty for broadcasters, reducing post-auction transition time and overhead, and potentially reducing the overall cost of the auction.
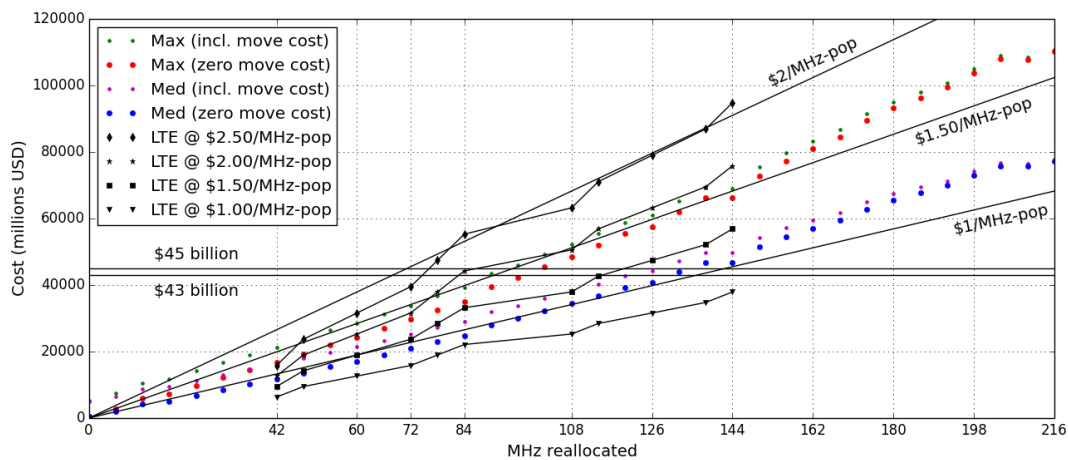
Figure A.1: Potential cumulative cost of compensating broadcasters for relinquishing spectrum usage rights.

A.1.2   Brief overview of methods and data

A.1.2.1   Auction assumptions

We make the following assumptions on the behavior of the auction and its participants:

- Only bids to relinquish spectrum usage rights will be accepted. (Specifically: no channel sharing bids will be accepted. No UHF-to-VHF bids will be accepted. If such bids were accepted, it may be possible to achieve larger band plans.)

- *All* broadcasters participate in the auction and are willing to relinquish their spectrum usage rights at precisely the prices given in the Greenhill report[1].

- It is sufficient to minimize the number of broadcasters which will relinquish spectrum rights, as opposed to minimizing auction costs directly. Due to computational limitations, minimizing auction costs is significantly more difficult than minimizing the number of broadcasters which relinquish their rights.

A.1.2.2   Reassignment methodology

Potential reassignments were created using the FCC's latest repacking constraints and data available on its LEARN website since 20 May 2014. Satisfiable solutions were found using PycoSAT, a Python interface for picoSAT. PicoSAT has been used by most other groups studying the incentive auctions, including a study commissioned by AT&T at the University of Pennsylvania [70].

In each reassignment, effort was made to minimize the number of broadcasters that would relinquish their spectrum usage rights. We do *not* consider minimizing the cost of the reverse auction because this is a fundamentally more difficult problem.

For practical reasons we are (1) unable to find the true minimum number for most band plans and (2) unable to generate a large quantity of potential reassignments for the minimum numbers we have obtained. Thus in each reassignment a small number of "extra" broadcasters are counted among those who must relinquish their spectrum usage rights.

All reasonable efforts were made to uniformly sample the space of satisfiable solutions. However, our research suggests that truly uniform sampling is a difficult problem requiring further study.

Below we present aggregated data. Each point represents the results from 100 potential reallocations. We present data for the FCC's identified band plans as well other natural band plans (i.e. N channels removed).

---

[1]We have very preliminary work which considers alternatives to this assumption.

A.1.2.3 Pricing

Potential costs of broadcasters' spectrum usage rights are taken from the Greenhill report [149]. Each auction-eligible broadcaster is assigned a price based on its DMA. We consider both the maximum and median prices reported by Greenhill.

We define the **potential cost of a specific reallocation** as the cumulative cost of compensating broadcasters for relinquishing spectrum usage rights. These values are shown by the red and blue dots in Figures A.1 and A.2.

We also consider the additional per-station cost of relocating broadcasters within their home bands. This is captured by the smaller green and magenta dots in the figures for a nominal move cost of $3 million per station.

A.1.3 Preliminary findings

A.1.3.1 Description

We present our preliminary findings in Figure A.1 which shows the potential auction costs. **Each set of dots** represents the *average* cost of the auction across many potential reverse auction outcomes and reassignments. We found that the variation in cost between assignments was minimal (approximately ±$1.5 billion).

The **horizontal lines** mark $43 and $45 billion. Some have predicted that the auction will raise as much as $45 billion[2], which leaves $43 billion for compensating auction winners after the following amounts are removed[3]:

- $250 million to cover the costs of running the auction

- $1.75 billion to cover the costs of moving stations to a new channel

The **solid angled lines** are used to mark various $/MHz-pop values for reference. These values are calculated as if *all* repurposed spectrum is purchased. The **jagged black lines** indicate the price per MHz-pop values when only LTE spectrum is purchased.

We are not in a position to predict the demand curve of the forward auction. However, we do *not* expect it to follow any of these reference lines exactly.

---

[2]EOBC Notice of Oral Ex Parte (June 13, 2014) according to [149].
[3] Note that it is no longer necessary to raise money for the Public Safety Trust Fund due to the overwhelming success of the AWS-3 auction.

A.1.3.2 Discussion

If broadcasters are compensated using the *maximum* numbers reported by Greenhill, it works out to be approximately $1.50 per MHz-pop (or approximately $2 per LTE-MHz-pop, depending on the band plan). The median numbers predict something more like $1 per MHz-pop ($1.25 per LTE-MHz-pop).

If forward auction participants are indeed willing to collectively pay $45 billion (at about $2/MHz-pop), our analysis suggests that the 84 MHz band plan (7 paired LTE channels) is the most likely outcome if broadcasters are compensated as richly as suggested in the Greenhill report. However, more modest compensation (i.e. the median rather than maximum value per DMA) points to the 114 MHz band plan (9 paired LTE channels) as most likely. The 138 and 144 MHz band plans appear to be out of reach if the Greenhill estimates are accurate and only $45 billion is raised. In order to reach the 144 MHz band plan, forward auction participants may need to raise as much as $70 billion (at $2/MHz-pop). We wish to remind the reader that this analysis is done under the assumption that all broadcasters fully participate in the reverse auction.

A.1.4 Proposal: a simpler repack algorithm

Our methods above made no attempt to minimize the number of station reassignments. After performing some exploratory analysis, we suggest considering a new reassignment method which is within the scope of the FCC's current incentive auction rules. Our proposed methodology has the following important properties:

- There will be no change to the set of bidding options offered to each broadcaster; however

- No stations below the spectrum clearing target (i.e. in post-auction broadcast TV spectrum) will be reassigned without participating in the auction and receiving direct, upfront compensation as the result of a winning channel-sharing bid or relinquishment bid.

- The cost of the reverse auction is not significantly impacted.

In particular, we propose artificially limiting the domain of stations *below* the spectrum clearing target to each station's currently-assigned channel. That is, stations which are currently using spectrum that will remain allocated for broadcast television will *not* be reassigned to a new channel.

Minimizing the number of reassignments is not strictly necessary since the TV Broadcaster Relocation Fund ("Reimbursement Fund") will cover a maximum of $1.75 billion of reallocation costs [13, ¶648]. However, we see several advantages to this method:

1. **Broadcaster certainty.** Broadcasters which do not participate in the auction are provided the certainty that their operations will be minimally affected by the incentive auctions. While they may still experience small increases in interference, broadcasters which do not participate will at least have the peace of mind that they will not incur any kind of reassignment cost. This would allay any concerns that broadcasters may have about the $1.75 billion fund being insufficient to cover the cost of all reassignments: non-participating broadcasters will incur no reassignment costs, while any participating broadcasters which truly believe it is insufficient will simply increase the values of their reverse auction bids accordingly.

2. **Faster post-auction transition.** Since far fewer stations will be reallocated, the post-auction transition is more likely to occur on schedule. For example, achieving the 84 MHz band plan could require over 1300 broadcaster reassignments versus 500 with our proposed methodology.

3. **Reduced overhead.** With fewer reallocated stations, the amount of processing needed for broadcaster reimbursement is significantly reduced.

4. **Potentially lower overall cost.** Figure A.3 shows the number of stations that would be need to be reallocated under our proposed scheme. At a modest estimate of 3 million per move, the total cost of reassignment alone for the 84 MHz band plan could be as much as $3.9 billion with the existing methodology vs. $1.5 billion with our proposed methodology. This is below the $1.75 billion reallocation fund amount, thus producing additional revenue.

Although our method does limit the flexibility of the auction process, we find that sufficient flexibility remains—especially with the larger band plans—to achieve approximately the same total broadcaster compensation costs (the red and blue dots in Figure A.2). The compensation costs rise slightly but are more than made up for by the reduced reassignment reimbursements. Note: these results may be affected by severe broadcaster non-participation which is a topic of future study. However, our results imply that this is a worthwhile methodology to explore.
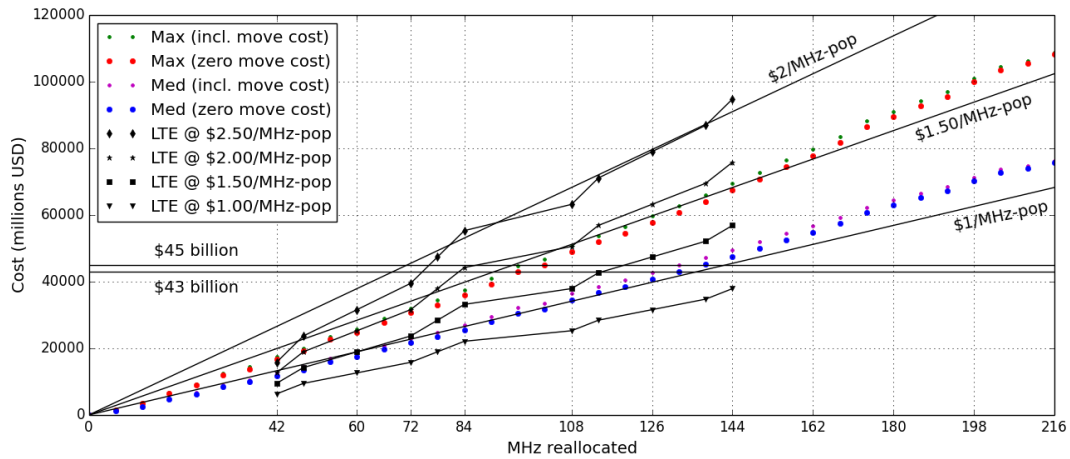
Figure A.2: Potential cumulative cost of compensating broadcasters for relinquishing spectrum usage rights *using our proposed method*.
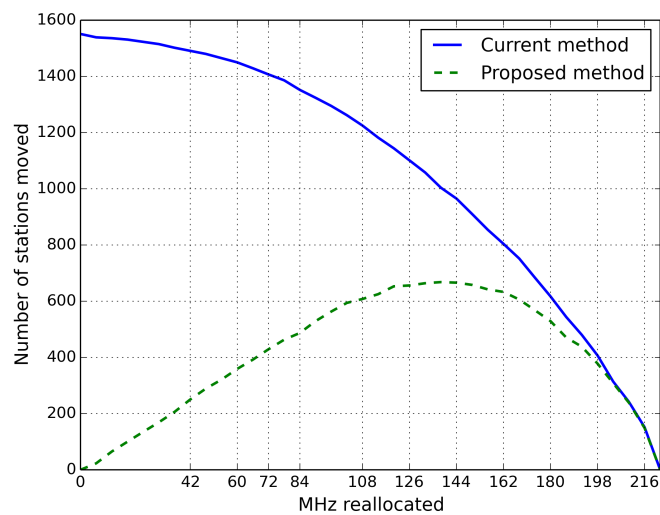
Figure A.3: Average number of stations which are *reallocated* after the incentive auction.

A.2   [FCC] Comments on the Part 15 NPRM

The following was submitted on 10 December 2014 to the Federal Communications Commission as a comment on proceedings 12-268 and 14-165. It can be viewed at `http://apps.fcc.gov/ecfs/document/view?id=60001008741`.

To whom it may concern:

As researchers, we feel we have an important position as a neutral third party in these proceedings. Our goals are to promote the most efficient and innovative use of spectrum while providing sufficient protection to incumbents. As such, we have the following comments:

Channels of operation for white space devices

1. We believe that the TV white space (TVWS) rules should **only allow portable devices to use channels 14 through 20** *if this operation is approved by a white space database* (WSDB). We agree wholeheartedly with the spirit of the Commission's proposal and suggest this change only to fix an apparent oversight which would allow sensing-only devices to use channels containing PLMRS equipment.

2. We believe that it is prudent to also **open up channels 2 through 13 for portable use**. We see no reason to do otherwise. Portable device manufacturers may or may not make use of these channels, but innovators should not need to wait for special regulatory approval to do so.

3. We support the proposal to **allow white space devices (fixed and portable) to operate on channels 3 and 4**.

Operating parameters for white space devices

1. We **strongly support the FCC's move to provide a variety of (emissions limit, separation distance) pairs**. We believe that this approach offers more flexibility to white space devices while adequately protecting the incumbents.

2. We believe that **the FCC should define a methodology for calculating the required separation distance given an emissions level and HAAT of the WSD**. We believe database providers should be given the choice to support a small or large number of emissions levels, based on their computational abilities[4]. This approach will allow for maximum flexibility while imposing no extra burden on the database providers.

---

[4]More precisely, what we mean is that database providers should be free to choose how many different emissions levels they wish to support. For example, a resource-poor database may only support a few levels (e.g. the ones currently defined in the NPRM) while a resource-rich database may support many levels. Since devices will always use the larger of two separation distances if they fall between two predefined emissions levels, this approach is still safe for incumbents. At the same time, it allows for database providers to differentiate themselves in the level of service that they can provide.

3. We believe that **white space devices should only be required to communicate their location to the database, not a proposed power**. In return, the database should provide a set of (max. power, channel) pairs describing what is available at the device's location. This is already done by database providers via PAWS. Furthermore, there is no reason for a device to pre-commit to a power level: operational decisions should be made on the device in real time after all available information has been collected. If databases want to support an additional query mode, e.g. one in which the device proposes a power level, then they should be able to do so.

Rural operation

1. We strongly support the idea of **increasing height and emissions limits in "rural" areas**. However, **we caution against using the proposed definition of rural**. Please see our extended comment below on this issue. Briefly, we instead propose simply expanding the set of (emissions limits, separation distances) to include at least a point at 10 Watts. We see no reason to artificially limit this to some areas.

2. We **recommend expanding this set further to include even higher emissions limits**. As long as the separation distances are adequate for the power and expected device density, this should cause no additional potential for interference. Additionally, since the separation distances are defined in the databases rather than the devices, these separation distances can be adjusted over time if they are found to be inadequate.

Location accuracy

1. We **strongly support the Commission's proposal to consider geolocation methods other than GPS and methods with variable accuracy**. We believe moving to a more general notion of geolocation will promote innovation and competition.

2. We believe that **devices should be allowed to report their location uncertainty as a polygon or set of polygons**. In the future, devices may be able to prove that they are located within a generic shape. Since location uncertainty directly factors into the amount of recoverable white space, allowing the use of more generic shapes would drive innovation and improve access to spectrum. However, we also propose that support for this type of uncertainty should be a choice left to the database providers. In this way, we are not held back by the "lowest common denominator" among database providers but can move forward to provide the best service possible. Please see our extended comment below regarding alternative geolocation methods.

Allowing non-white space devices to support white space devices

1. We believe that **white space devices should be allowed to "tether" to a non-white space device for the purposes of using its Internet connection to contact the database**. Since the communication to the database will be encrypted end-to-end, the device providing the connection need not be certified. For example, a fixed or geolocated portable device may "tether" to a nearby smartphone rather than requiring its own Internet connection.

2. We also believe that **non-white space devices should be able to help geolocate white space devices**. For example, a user's personal/portable white space device may use the geolocation functionality (e.g. GPS) on the user's smartphone rather than including a costly GPS receiver of its own. Alternatively, indoor devices may use their proximity to a location beacon–which need not be a fixed or Mode II white space device–to approximate their true location. We believe that this proposal has significant potential to promote innovation and economic growth in the white space device market.

Alterations to the channel re-check interval

1. We **agree that a long re-check interval for WSDs may encourage protected entities to over-register** if they feel unable to predict their actual usage needs far in advance. We believe that shortening the re-check interval, even for a few channels, may help with the problem of over-registration.

2. However, we also believe that **the solution to over-registration of protected entities should not be limited to changing the behavior or permissions of white space devices**. Instead, abuse should be addressed head-on by the Commission.

3. We believe that **limiting the number of hours a protected entity may reserve without a waiver may help to curb this abuse**. Furthermore, registrants which continually over-reserve – as shown by reported usage statistics – should be directly admonished. If abuse still continues, registration privileges may be suspended or withdrawn, or fines may be levied.

4. We **agree with the precedent set by the change of the re-check interval**. We believe that spectrum availability and protection parameters should be defined in a dynamic way rather than encouraging the assumption that they will not change once enacted.

5. We **propose that the re-check interval be defined as a variable parameter in the response from the white space database**. While mandating this would be heavy-handed, the Commission does have the flexibility to adopt rules which encourage this behavior. For example, a small number of channels could be designated as "fast response channels," requiring a shorter re-check interval, while others require infrequent re-checks. This not only supports device diversity but it also promotes the adoption of more flexible protocols. In this way, the Commission would be able to change the re-check interval(s) in the future without needing to adopt an explicit transition plan (as proposed in §15.37(i)) and thereby hastening the obsolescence of otherwise-compliant devices.

Comments on the proposed definition of "rural area" In ¶45 of the Part 15 NPRM, the following definition of white space "rural area" was proposed:

> We propose to identify rural areas for white space devices as those where at least half of the TV channels are unused for broadcast services and available for white space use.

We have identified several problems with this definition for "white space rural," supported by our figures below:

1. *Use of the word "rural" is confusing* because although there is a strong correlation between truly rural areas and "white space rural" areas, they are not the same. This can be seen by looking at the maps below showing the "white space rural" areas in blue.

2. *The definition as stated is unnecessarily complicated.* As we understand it, it would be sufficient to say "at least half of the TV channels are available for white space use."

3. Although the "white space rural" areas tend to be in areas with few TV stations, *these areas are not necessarily far from the protected contour of TV stations*. We support the idea of defining an additional separation distance for "rural" operation for this reason.

4. The *locations which are considered to be "white space rural" are erratic in their nature*. This complicates system planning for white space device operators which wish to take advantage of the "rural" areas.

5. Finally, we believe that *emissions limits should depend only on the distance to co- and adjacent-channel contours rather than the number of TV channels which are available in the region*. Creating a dependency conflates the two.

Rather than define specific rural areas, we propose to simply allow white space device operation at 10 Watts if the device is sufficiently far from co- and adjacent-channel contours. This separation distance can be made conservatively large at first and adjusted over time to suit the needs of the primary and secondary systems. This proposal fits within the existing framework for defining protections and requires no new logic within the white space databases.

The figures above were produced using WEST (https://github.com/kate-harrison/west), an open-source tool for evaluating the amount of white space available under different regulatory assumptions. For the simple calculations above, we have included only TV station protections (i.e. we have ignored PLMRS/CMRS, wireless microphone, etc. protections). We have also used idealized circular contours based on F-curves for the sake of quick computation. The fixed white space device was assumed to have a 30 meter HAAT.
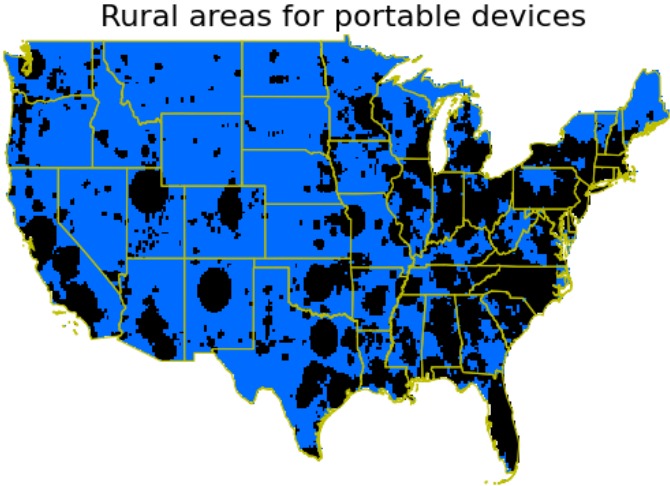
Figure A.4: Rural areas for portable devices
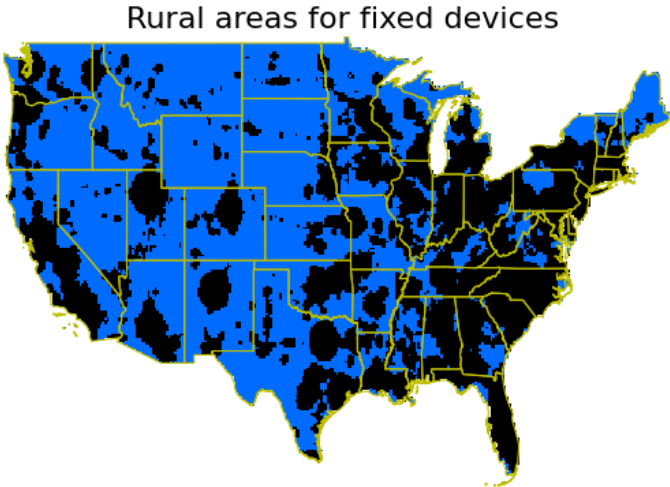
Figure A.5: Rural areas for fixed devices

Comments on alternative geolocation methods

As part of our research at UC Berkeley, we have conducted studies on how to support non-geolocated or weakly-geolocated devices. In particular, we compare the FCC's approach, Ofcom's approach, and a new proposed method which achieves the same protection goals while increasing the amount of recoverable white space for these devices.

This research was published and presented at IEEE DySpAN 2014 in April and can be found at http://inst.eecs.berkeley.edu/˜harriska/docs/2014_DySpAN_localization.pdf. We believe that our proposals may help address some of issues related to the indoor operation of devices.

At a high level, our proposal can be summarized as follows:

1. We disagree with the FCC's assumption that slave devices will be near enough to their master that the set of available channels is identical.

2. We agree with Ofcom's approach of generic operating parameters but find them unnecessarily restrictive.

3. White space databases need only return safe responses to white space devices, not necessarily the most permissive safe responses. (This sentiment is already reflected in e.g. Ofcom's generic operational parameters.)

4. Hypothetically, if a device knows with high confidence that it is either in Cory Hall on the UC Berkeley campus or in the FCC's headquarters, it is clear how the database should respond: it should allow the white space device to transmit only on channels which are available at both locations. This admittedly silly example demonstrates the heart of our proposal:

   a) The white space access problem is not necessarily a geolocation problem.

   b) White space databases can give answers that protect the primary, even when the geolocation uncertainty is high.

5. White space devices can generate their set of possible locations in many ways. For example, the device may use a combination of max-range-to-master and RF fingerprinting (e.g. WiFi fingerprinting or TV signal fingerprinting). The latter uses measurements of strong TV signals to rule out potential slave locations (e.g. "I am definitely in the service area of a station on channel 14, so I cannot be in locations where channel 14 is not viewable").

   a) Note that we do not actually suggest allowing the white space device to generate the set of possible locations. Instead, we envision this functionality sitting in the cloud, either as part of the WSDB or as a separate add-on service. In addition to it being nearly impossible for the device to perform this task itself, we see many benefits in using software as a service that can be continuously evolved as requirements change.

    b) We also envision a richer set of location-relevant data (not just WiFi- and TV signal fingerprinting) which evolves over time to meet consumers' and the Commission's needs and reflects the capabilities of modern technology.

If you would like to learn more about our proposals or our work, please do not hesitate to contact us.

Sincerely,

Kate Harrison and Anant Sahai
harriska@eecs.berkeley.edu, sahai@eecs.berkeley.edu

A.3 [Ofcom] Proposal to address bootstrapping and non-geolocated device problems

The following was submitted via email on 8 November 2014 to Ofcom as a comment on their white spaces consultation.

To whom it may concern:

Purpose: To promote innovation by suggesting additions and modifications to Ofcom's TV white space rules which increase the amount of recoverable white space for non-geolocated or weakly-geolocated devices as well as geolocated devices without a direct connection to a white space database.

Proposals:

1. Allow white space devices to "tether" to a non-white space device for the purposes of using its Internet connection to contact the database. Since the communication to the database will be encrypted end-to-end, the device providing the connection need not be certified. This will allow geolocated slave devices to bootstrap a connection to the master over white space.

2. Allow white space devices to prove their location using methods other than GPS or communication with a master WSD. For example, a device could as well prove that it is within the (smaller) service area of a non-white space device, possibly leading to more permissive generic operating parameters.

3. Allow white space devices to report their location uncertainty as a polygon or set of polygons, rather than a rectangle. In the future, devices may be able to prove that they are located within a more generic shape. Since location uncertainty directly factors into the amount of recoverable white space, allowing more generic shapes would drive innovation and improve access to spectrum.

Studies: As part of our research at UC Berkeley, we have conducted studies on how to support non-geolocated or weakly-geolocated devices. In particular, we compare the FCC's approach, Ofcom's approach, and a new proposed method which achieves the same protection goals while increasing the amount of recoverable white space for these devices.

This research was published and presented at IEEE DySpAN 2014 in April and can be found at http://inst.eecs.berkeley.edu/~harriska/docs/2014_DySpAN_localization.pdf. We believe that our proposals may help address some of the issues that Ofcom is facing.

If you would like to learn more about our proposal or our work, please do not hesitate to contact us.

Sincerely,

Kate Harrison and Anant Sahai
harriska@eecs.berkeley.edu, sahai@eecs.berkeley.edu

A.4 [FCC] Register everyone: on the whitespace use of wireless microphone channels, channel 37, and the soon-to-be guard bands

The following was submitted on 25 January 2013 to the Federal Communications Commission as a comment on proceeding 12-268. It can be viewed at `http://apps.fcc.gov/ecfs/document/view?id=7022112348`.

Abstract

This report is intended as a response to some of the questions posed by the FCC regarding the upcoming TV-band incentive auction, given in their NPRM [68], as they relate to the television whitespaces.

In particular, we argue (1) that channel 37 should be made available for whitespace use; (2) that the channels reserved for wireless microphones should be reserved on an as-used basis only; and (3) that the guard bands which will be created via the incentive auction must be considered as database-registration-requiring whitespace if unlicensed devices are authorized to use them.

These three proposals have two common themes: (1) they each work toward the goal of making otherwise-wasted spectrum available as whitespace; and (2) in each case, the key concept is that the involved parties can (and in some cases *must*) register their devices and use geolocation of some sort.

We will sketch each of our proposals and show how together they can make whitespace available for up to 10 million *more* Americans with minimal overhead while ensuring that licensed users receive the quality of service that they expect. As a result, essentially no one would be left without whitespace access.

A.4.1 Introduction

On October 2, 2012, the FCC released its Notice of Proposed Rulemaking (NPRM) for incentive auctions in the television bands [68]. The intent is to monetarily incentivize broadcasting licensees to modify their spectrum usage rights (either by giving them up completely or moving to another band) in order to reorganize ("repack") the spectrum currently occupied by over-the-air television broadcast.

At the same time, unlicensed TV-bands devices (TVBDs, a.k.a. whitespace devices) may operate in the same frequencies under the FCC's rules [1, 9, 23]. These TVBDs represent an enormous economic and innovative opportunity. To that end, the FCC recognizes the need for more whitespaces:

Currently, some urban markets do not have channels available for white space use. To address this issue, the National Broadband Plan recommended that, as the FCC seeks to provide additional spectrum for broadband services, it make available for exclusive or predominant use by unlicensed devices sufficient spectrum to enable innovators to try new ideas for increasing broadband access and efficiency, and to enable new unlicensed broadband access providers to serve rural and unserved communities. [68, ¶231]

Furthermore, the FCC expects and intends to maintain the whitespaces as a place for innovation and development:

Given that there is considerable white space available now in many areas—more than 100 megahertz in some markets—we expect that there will still be a substantial amount of spectrum available for use by these devices in the remaining broadcast television channels after the incentive auction. We also expect that there will continue to be more spectrum available in areas outside of the central urban areas of the largest markets than within those areas. We seek comment on these views. It is our intent to continue to allow both the use of white space devices and the development of devices for various applications that operate in the broadcast television bands after the incentive auction. [68, ¶233]

Furthermore, the FCC is contemplating allowing TVBDs to operate on channel 37[5] as well as the two channels currently reserved (from TVBDs) for wireless microphone use[6].

However, it is difficult to exactly predict what the results of the incentive auction—more specifically, the repacking—will be or what they will mean for the white spaces. A number of things may change as a result:

- The sets of channels which are available (UHF, VHF) may change; this is significant given the variation in propagation characteristics between these bands.

---

[5] "In addition, there may be an opportunity for unlicensed devices to operate in channel 37 (an additional 6 megahertz of spectrum), whether or not we relocate the WMTS and the Radio Astronomy Service now using channel 37. As discussed in section VII, the rules require that locations of WMTS operations be registered with the American Society for Healthcare Engineering (ASHE), and there are relatively few radio astronomy operations, all at specified locations. Therefore, we may be able to protect these services by establishing appropriate protection areas in the white space database. We propose to make channel 37 available for unlicensed use, while protecting WMTS and the Radio Astronomy Service." [68, ¶237]

[6]"The current rules for white space devices in the television bands designate two channels (when available) in all locations for use by wireless microphones. White space devices are not permitted to operate on these channels, preventing them from using 12 megahertz of spectrum that could otherwise be available for their use. We invite comment as to whether the Commission should maintain the designation of two channels for wireless microphones following the broadcast television spectrum incentive auction or whether this spectrum should be made available for unlicensed use." [68, ¶238]

- The number of cochannel TV transmitters will likely decrease which improves the noise floor for TVBDs.

- The effect of adjacent-channel restrictions may increase or decrease depending on the repacking.

- The total amount of whitespace available (in MHz) will almost certainly decrease as a result of the forward auction[7].

- Channel 37 may be available as whitespace.

- The two channels reserved for wireless microphone may be once again available for whitespace use.

- The existence of guard bands which will potentially provide "clean" whitespace for TVBDs.

The possibilities are endless; thus for clarity we only focus on the last three items in this report. We do not attempt to predict the results of the repacking and instead use the current allocations [28] and rules [23] (which will be valid for the next few years when many TVBDs are being developed). We fully expect that the details of these results will change after the incentive auction; however, we expect that our general message will not change.

We expect that, through the use of modern technology, we can easily protect wireless microphones and channel 37 services while using the remainder of these channels for TVBDs. This has the opportunity to provide huge gains in whitespace, including opening it up for use by up to 10 million people who currently do not have access to whitespace spectrum.

A.4.2 Registration

As the title implies, one of our central themes is that everyone seeking protection from TVBDs— television licensees, wireless microphone operators (both licensed and unlicensed), channel 37 occupants, **and the winners of the forward auction**—should register their sites in a real-time database in order to minimize interference to protected services while maximizing whitespace availability. The FCC rules for whitespace use in the TV bands already mandate such a database [1,9,23] for TV licensees and licensed wireless microphones; it is reasonable and technically easy to add to the list of protected services.

---

[7]The repacking is intended, in part, to increase the efficiency with which the spectrum is used. However, increasing the efficiency means that there are fewer "cracks" (a.k.a. whitespaces) for TVBDs to use.

A.4.2.1   Registration benefits for protected services

When a protected entity (e.g. TV licensee) registers in the database, the database administrator calculates the corresponding protected region in which no TVBDs may transmit co-channel. The protected region is designed from the parameters of the entity (e.g. transmit power, location) to minimize the amount of interference to the protected entity's receivers (e.g. consumers' television sets).

There is always a protected region cochannel to the protected entity and sometimes (as in the case of TV licensees) there are adjacent-channel protected regions as well. TVBDs must[8] request permission from the database before they can transmit; this ensures that they do not erroneously transmit so as to cause harm inside a protected region.

The burden on the protected entity depends on its type. Many such entities are at fixed locations and have relatively static needs, e.g. TV licensees, the winners of the forward auction[9], and the channel 37 occupants (described in more detail in Section A.4.3). In these cases, registration is a one-time process and could even be done automatically by the FCC when the entities receive or modify their operating permits.

Some entities, e.g. wireless microphones, may be used intermittently and at a wide variety of locations. Due to this unpredictability, the level of involvement required is higher but still not unreasonable, as argued in Section A.4.4.

A.4.2.2   Registration benefits for TVBDs

The FCC's current rules for the TV whitespaces [23] allow TVBDs to either (1) contact a database to request permission to transmit; or (2) sense with high sensitivity to determine if the channel is currently occupied by a protected entity. While advances in sensing, especially collaborative sensing, are promising, it has been shown that there is a significant loss of whitespace when sensing instead of contacting a database [48]. This is as a result of the safety margin that *must* be included to ensure adequate protection for the protected services.

Furthermore, the FCC allows for limited "chaining" of devices: a device *without* location services may receive a list of allowed channels from a device which is able to communicate directly with the database. This allows for the deployment of many smaller, cheaper devices in the presence of one more sophisticated device, further reducing the cost of database access.

We believe that the promise of increased spectrum availability will outweigh the increased cost of location-estimation equipment[10] and thus that most manufacturers of TVBDs will opt to use the database method for whitespace discovery.

---

[8]The caveat to this is explained in the following subsection.

[9]We expect that the winners of the forward auction will implement LTE *or* will deploy an LTE-like system given the characteristics of the spectrum.

[10]One low-cost alternative—which is not currently in the FCC's rules—is to allow devices to submit a "fuzzy location." This is a region (of any size) in which the device is guaranteed to be located. The database would authorize the use of only those channels which are available at *all* points inside the region. In this way, manufacturers can individually choose a point on the tradeoff curve between equipment cost and spectrum availability.

A.4.3 Argument for Channel 37

Both radioastronomy and WMTS (wireless medical telemetry service) are very important and should continue to operate normally. However, we believe that there is also room for whitespace devices in these channels.

A.4.3.1 Radioastronomy sites

The few radioastronomy sites in the United States are already registered with the FCC whitespace database. They are currently afforded 2.4 km of separation from whitespace devices on all channels. Although we recognize that this equipment is very sensitive, we argue that there exists a distance at which it is safe to operate wireless devices on the same channel. Indeed, this already happens with WMTS devices (which are prohibited from operating near the radioastronomy sites but are otherwise allowed to operate on channel 37). We do not pretend to know[11] what this distance is or should be, but for illustrative purposes we assume that this distance is 50 km. We then see the following excluded areas (in black) in Figure A.6. Here we see that even though 50 km is a large separation distance, a large portion of the nation (shown in blue) is still available for whitespace use on channel 37. This number should be carefully chosen so that interference is unlikely. However, it is important to realize that we can *easily* and *quickly* make modifications via the databases should the need arise.

Channel 37 exclusions
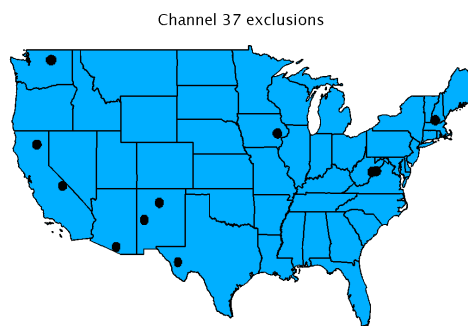
Figure A.6: Exclusions on channel 37 due to radioastronomy sites: black indicates locations where secondaries are not permitted to transmit.

A.4.3.2 Wireless medical telemetry service (WMTS) devices

WMTS devices are devices which are used in hospitals to help monitor a patient's condition:

> WMTS spectrum is used for remote monitoring of a patients health. Wireless medical
> telemetry systems include devices to measure patients' vital signs and other important

---

[11]This question is definitely answerable but we have not done the calculations ourselves.

health parameters (e.g., pulse and respiration rates) and devices that transport the data via a radio link to a remote location, such as a nurses' station, equipped with a specialized radio receiver. For example, wireless cardiac monitors are often used to monitor patients following surgery. [150]

WMTS devices are already required to register with the FCC[12]. Not only are they registered but their locations (or at least locations which tend to use such devices) are predictable and relatively static over time. It would be no harder to accommodate WMTS devices than it is to accommodate TV stations.

Note that since we do not have the locations nor the protection criteria of these devices/hospitals, we have not included them in our calculations. However, we expect that the impact would be relatively small compared to the benefit derived from allowing TVBDs to transmit on channel 37.

A.4.4   Argument for wireless microphone channels

There are two characteristics that significantly set wireless microphones apart from radioastronomy equipment and WMTS devices:

1. Not all wireless microphones are licensed.

2. Registration of wireless microphones is more difficult.

We will tackle each of these issues in order in the following sections. In general, we argue that wireless microphones (or locations which utilize them) should register in the whitespace database in order to receive protection from TVBDs, thus allowing TVBDs to otherwise operate on these two channels.

A.4.4.1   Not all wireless microphones are licensed

In the FCC's 2010 regulations, they expanded the PLMRS/CMRS exclusions[13] to include the "reservation of two channels in the range of 14-51 to all markets nationwide as suggested by several petitioners" [9, ¶29]. The intent was to create a "safe haven" for unlicensed wireless microphones since they stated that it would be inappropriate to allow unlicensed wireless microphones to register as primaries in the database due to their unlicensed status:

---

[12]"WMTS devices must be registered with the FCC's designated frequency coordinator, the American Society for Healthcare Engineering of the American Hospital Association (ASHE/AHA)." [150]

[13]These reserve 1-3 channels in 13 major metropolitan areas.

> With regard to registration of unlicensed devices in the TV bands database, we first observe that unlicensed wireless microphones operate under the same general conditions of operation in Section 15.5 of the rules as TV bands devices, meaning they may not cause interference to authorized services and must accept any interference received, including interference from other non-licensed devices. As a general matter, we therefore find that it would be inappropriate to protected unlicensed wireless microphones against harmful interference from other unlicensed devices, and in particular TV bands devices. [9, ¶31]

The most natural solution, given their unlicensed status, is to turn unlicensed wireless microphones into TVBDs. However, this is impractical due to the difference in requirements: wireless microphones do not typically have Internet access and they need to be low-latency given their real-time use [9, ¶30].

It is likewise impractical to move wireless microphones to alternative bands or leave them unprotected: too many venues, ranging from professional sports events to local theatrical productions, rely on wireless microphones and replacing this equipment could be very costly.

Therefore rather than reserving two *full* channels nationwide for these unlicensed devices, we should simply allow unlicensed microphones to be registered for protection *on these two channels* using the whitespace databases. Reserving two full channels nationwide *already* "protect[s] unlicensed wireless microphones against harmful interference from other unlicensed devices"—the FCC's stated reason for denying them registration rights—so allowing them to register (only within these two channels) would be no worse. Some of these unlicensed wireless microphones are further protected since "unlicensed microphones at event sites qualifying for registration in TV bands databases will be afforded the same geographic spacing from TVBDs as licensed microphones" [9, ¶32]

A.4.4.2 Registration of wireless microphones only seems hard

There are two reasons that the registration of wireless microphones seem more difficult than that of WMTS or radioastronomy sites:

1. Wireless microphones are more numerous and dispersed

2. Wireless microphones are not necessarily operated by professionals, organizations, etc.

However, we believe that the prevalence of smart phones today can help solve this problem. We propose the development of an application for iPhones, Android devices, etc. that would make the registration procedure for wireless microphones quick and easy for any operator. In fact, the FCC has already suggested that operators of wireless microphones consult the whitespace database for a list of available channels:

Entities desiring to operate wireless microphones on an unlicensed basis without potential for interference from TVBDs may use the two channels in each market area where TVBDs are not allowed to operate, as well as other TV channels that will be available in the vast majority of locations. *Such entities may consult with a TV bands database to identify the reserved channels at their location, as well as the TV channels that may not be available for TV bands devices.* [9, ¶32]

Furthermore, Spectrum Bridge has produced an application [151] which helps wireless microphone operators determine the best channels on which to transmit. Inspired by Spectrum Bridge's application, we have created a mock-up our proposed application, shown in Figures A.7(a) and A.7(b).
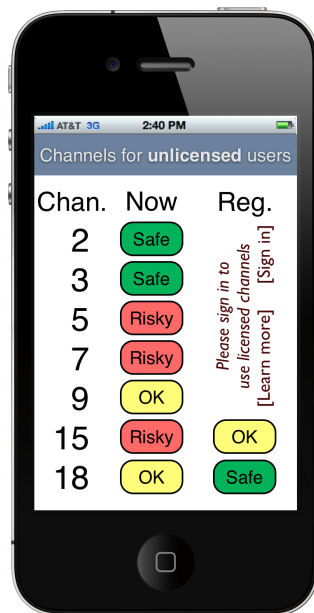
(a) Interface for *unlicensed* wireless microphone operators
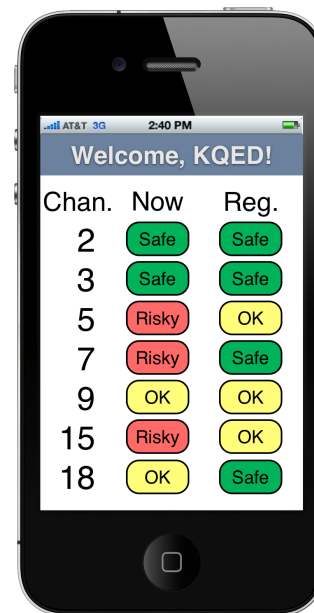
(b) Interface for *licensed* wireless microphone operators

Figure A.7: Mockups of an iPhone app

This application would use the location services (roughly, GPS) available on iPhones and Android devices to automatically determine the location of the venue. This would be sent to a TV whitespace database which would then feed back information on the current local channel availability and quality, shown in the middle column of Figures A.7(a) and A.7(b). In addition to the current channel quality, the database would also provide information on the channel quality that could result from venue registration, shown in the far right column of Figures A.7(a) and A.7(b). Note that operators of unlicensed microphones will see the screen in Figure A.7(a). Users who sign in with the appropriate credentials (i.e. are confirmed operators of licensed microphones) will see a

screen like that in Figure A.7(b). This provides operators with a gentle reminder that unlicensed microphones are not allowed to register on more than the two provided channels.

If the operator determines that he can operate reasonably without registration, then no further action is required. Until such time as it is necessary to register in the database, he can simply use this application to help inform his choice in channels, similar to the function that Spectrum Bridge provides with their application.

However, if he wishes to improve his operating conditions he can simply tap on the entry corresponding to the desired channel (e.g. the green "Safe" button for channel 18) to register his location (venue) in the database. The application then transmits this information to the TV whitespace database which acknowledges receipt and forwards the information to the other TV whitespace databases[14].

There are clearly some questions regarding the design of this application, for example:

- How long should the registration last? Some events last only a few hours while others last days.

- What area should be protected? Just 1 km around the location at which the request was sent?

- Should the reservation start immediately or after some delay? How long should this delay be?

However, the main point is that it is easy to create such an application which can be utilized by most Americans *today*. Furthermore, if demand is present, multiple similar applications can be created to suit the needs of different types of operators and venues.

Thus we have demonstrated one possible way to overcome the difficulties inherent in wireless microphone registration.

A.4.4.3 Additional benefits of a wireless microphone registration app

One issue which we have not yet seen addressed is the potential operation of the wireless microphones within auctioned bands. We are unaware of any mechanism by which wireless microphones will be migrated out of these bands, thus by default they will be technically able to cause interference to licensed users. If left completely unaddressed, this has the potential to decrease the value of these bands in the forward auction which in turn threatens the success of the entire auction.

While it is by no means a complete solution, we suggest that the registration application proposed above could also be used to notify wireless microphone operators that some channels are unavailable (perhaps by always reporting them as low-quality), thus helping to nudge their operation to more acceptable bands.

---

[14]The existence of multiple database providers [152] only slightly complicates the situation. These databases are already required to exchange information on a daily basis and this requires only a change in timescale.

A.4.5   The guard bands must have spatial holes

Through the incentive auction, the FCC will soon create two guard bands, each 6 to 10 MHz wide, which separate up- and down-link spectrum from broadcast television in order to prevent harmful interference. The FCC intends[15] to allow unlicensed devices to transmit in these guard bands [68, ¶126]:

> ...this approach would create a uniform downlink band plan to help ensure interoperability, and nationwide guard bands that could be used by unlicensed white space devices, at least on a secondary basis. [68, ¶182]

What does operating on a secondary basis mean? In order to adequately protect licensed services (e.g. broadcast television and the winners of the forward auction) from undue amounts of interference, we must enforce some protected regions in space within the guard bands whenever a protected service is operating on an adjacent channel. The only alternative is to impose such an extremely low power limit that adjacent-channel interference is inconsequential; however, with the advent of the whitespace technology this is clearly wasting an opportunity.

We wholeheartedly support the FCC's decision to treat the guard bands as whitespace but we wish to remind all interested parties that one consequence of this decision is that the bands will *not* be guaranteed to be available nationwide[16]. The guard bands cannot be thought of as equivalent to the 2.4 GHz ISM band because of these numerous spatial holes.

A.4.5.1   Protected TV licensees

As is currently required in the TV whitespaces, unlicensed devices should not be allowed to operate in the 6 MHz adjacent to a broadcast TV station's channel while within its service area[17]. Therefore it is reasonable to require that devices operating in the guard bands will contact the TV whitespace databases; this then imposes no additional overhead for any of the involved parties.

---

[15]Although the quotation below comes from a section on an alternative band plan, we believe that the FCC intends for the guard bands to be used by whitespace devices rather than regular unlicensed devices. This is particularly evident in footnote 198 (from ¶126): "This unlicensed spectrum is in addition to (rather than in lieu of) the white space spectrum that exists today in the UHF band, and will continue to exist after the repacking of the broadcast services."

[16]Depending on the size of the guard bands, it may be possible to exclude devices on only a portion of a guard band, e.g. the 6 MHz nearest the relevant protected service. However, it is still possible for the entire guard band to be off-limits to whitespace devices at that location since protections may overlap, e.g. when a TV transmitter and an cellular tower are near one another.

[17]These should be the same protections afforded in the FCC's rules for the TV whitespaces. In particular, the separation distance—an additional spatial buffer between the service area and transmitting whitespace devices—ranges from 0.4 to 2.4 km, depending on the height of the whitespace device [23, §15.712].

A.4.5.2   Protecting the winners of the forward auction

For the purposes of this discussion, we will assume that the winners will deploy an LTE-like network[18].

Up-link

Protecting the up-link portion of an LTE-like network is simple: the receivers (i.e. cellular towers, pico- or femtocells) are at fixed locations and are long-lived enough that it is worth it to characterize the receiver specifications. From this, a reasonably-sized exclusion region can be created around each tower and registered in the database. The relatively static nature of these receivers keeps the overhead to a minimum.

Down-link

The problem of protecting the down-link portion of an LTE-like network is very similar to the problem of protected TV receivers: their exact locations are unknown but they are necessarily within some range of their tower(s) of interest. The sensible solution in this case is also the same as that for the TV receivers: create a protected region around the tower which is large enough to encompass all associated receivers plus an additional spatial buffer. Carriers, while planning their network, already compute these spatial footprints in order to determine their coverage maps so registering them in the whitespace database is not a burden for them.

A.4.5.3   Whitespace in the guard bands should merge with the TV whitespaces

The FCC remarks in footnote 198 (from [68, ¶126]): "This unlicensed spectrum [in the guard bands] is in addition to (rather than in lieu of) the white space spectrum that exists today in the UHF band, and will continue to exist after the repacking of the broadcast services." We interpret this to mean that the whitespace in the guard bands will be merged with the TV whitespaces and we wish to reinforce this idea. The whitespace in the guard bands cannot stand on its own for several reasons which we discuss below.

Coverage holes with no alternative are unacceptable

As argued above, the guard bands will necessarily have holes (as a consequence of providing adequate protection for licensed services) and therefore cannot provide full nationwide coverage *on their own*. In particular, coverage will be sparse in populated areas since there are more TV broadcast towers [38] and LTE-like networks in populated areas. With such holes, this spectrum will be attractive to only a very few manufacturers of unlicensed devices.

However, when coupled with the TV whitespaces we reach a critical mass of spectrum: with so much potential spectrum, no location will be completely bereft of whitespace. In particular, opening up the whitespaces in channel 37 will benefit metropolitan areas in which spectrum is typically scarce.

---

[18]In this case, location estimates are essentially free: a cellular-style network and any device that has access to such a network will be able to estimate its location with reasonable fidelity.

Aggregating whitespaces reduces overhead
Adding the whitespace in the guard bands to the TV whitespaces reduces the number of proceedings and simplifies the entire process. While the individual protections may differ (e.g. different protected regions and separation distances for LTE-like networks as opposed to TV licensees), using the same style and protocol allows for reuse of regulations and (more importantly) technology. These differences represent merely a slight change in the databases, *not* in the main text[19] of the regulations.

A.4.6   Proposed modifications to regulations

We propose the following changes to the FCC's regulations [23], also given in Figure A.8:

- Rather than reserving two channels nationwide for unlicensed microphones, just allow the operators of these devices to register on one of these two. Thus locations without currently operational unlicensed microphones would be available for whitespace on those two channels.

- Radioastronomy sites would be assigned a cochannel protected region which is large enough to avoid interference to their operations. These sites are already registered in the whitespace databases and protected regions could be easily modified if interference is observed.

- Locations where WMTS devices are operated (e.g. hospitals) would be assigned protection regions. These locations should already be registered with the FCC and could easily be added to the whitespace databases.

- The whitespace in the soon-to-be guard bands should be designated as part of the TV whitespaces. Furthermore, devices operating in the guard bands should be subject to adjacent-channel exclusions for protected services operating in nearby bands.

A.4.7   Impact of regulations on the whitespaces

The benefits of the whitespaces are already well-known to the reader: innovation spurred via a low barrier-to-entry courtesy of the fact that use of whitespace spectrum is free. However, it is also important to ensure that new entrants will have a large and inviting market for their products. To that end, we need to make sure that whitespace is readily available in populated areas such as the east and west coasts of the United States. We see in Figure A.9(a) that these are precisely the regions which have the least available whitespace. In fact, roughly 3% of US residents currently have no access to whitespace channels.

As shown in Figure A.9(b), some locations are losing up to 8 channels due to the metropolitan area exclusions:

---

[19]Instead, these changes could be reflected in updated or expanded tables of values, e.g. separation distances.

	Current FCC regulations	Proposed regulations
Television station protections	Protected region + separation margin cochannel and adjacent channel	*No change*
Wireless microphone protections	• Two channels nationwide • Venue registration	Venue registration only (extended to unlicensed on only these 2 channels)
Radioastronomy site protections	• Channel 37 nationwide • 2.4 km on all channels	• Large protected region on channel 37 • 2.4 km on all channels
WMTS (medical) protections	Channel 37 nationwide, except radioastronomy sites	Protected region on channel 37
PLMRS/CMRS protections	1-3 channels in each of 13 major metropolitan areas	*No change*
Proposed guard bands	Designated as whitespace	Part of the TV whitespaces and subject to adjacent-channel protections

Figure A.8: Table comparing the current and proposed protections.
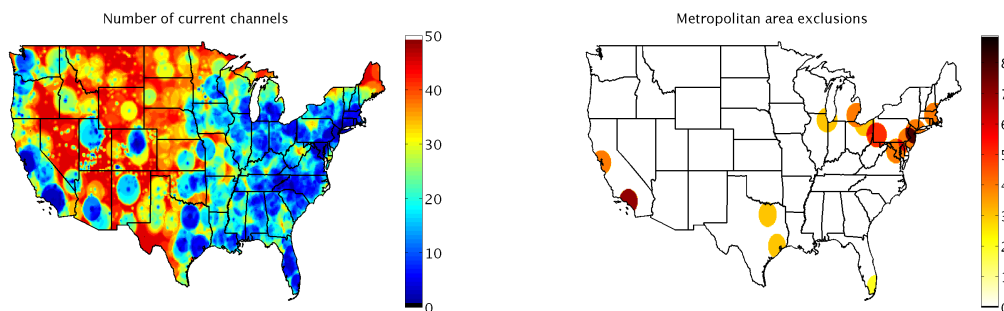
(a) Number of whitespace channels currently available.

(b) Number of channels excluded specifically by the metropolitan area exclusions.

Figure A.9:

> TVBDs may not operate at distances less than 134 km for co-channel operations and 131 km for adjacent channel operations from the coordinates of the metropolitan areas and on the channels listed in §90.303(a)[20] of this chapter. For PLMRS/CMRS operations authorized by waiver outside of the metropolitan areas listed in §90.303(a) of this chapter, co-channel and adjacent channel TVBDs may not operate closer than 54 km and 51 km, respectively from a base station. [9, §15.712(d)]

The FCC's wireless microphone protections further reduce whitespace availability by reserving two channels (approximately) nationwide for unlicensed microphones:

> All TVBDs are permitted to operate available channels in the frequency bands 512-608 MHz (TV channels 21-36) and 614-698 MHz (TV channels 38-51), subject to the interference protection requirements in §§15.711 and 15.712, except that the operations of TVBDs is prohibited on the first channel above and the first channel below TV channel 37 (608-614 MHz) that are available, i.e. not occupied by an authorized service. If a channel is not available both above and below channel 37, operation is prohibited on the first two channels nearest to channel 37. These channels will be identified and protected in the TV bands database(s). [9, §15.707(a)]

Analysis of the modifications proposed above (*excluding* the guard bands since their size is currently unknown nor do we know the deployments of the primary systems which will impose spatial exclusions on them) yields the following results, represented by the blue lines in Figures A.10(a) and A.10(b):

- Up to 10 million people (mostly in New York City and Miami) would now be able to use the whitespaces.

- Only approximately 650 people (in a rural area in central California) would remain without whitespace access of any kind.

- Over 30% of people would see *at least* a 50% increase in the number of whitespace channels (fractional gain = 0.5).

- Over 10% of people would see *at least* a doubling of the number of whitespace channels (fractional gain = 1).

The red and green lines in Figures A.10(a) and A.10(b) represent the individual effects of modifying *either* the channel 37 rules *or* the wireless microphone rules (but not both simultaneously). Notice that adding back the microphone channels has a more significant impact than opening up channel 37. This is because most places will gain two channels from a change in wireless microphone rules whereas they can gain a maximum of one channel by using the whitespace on channel 37.

---

[20]This document can be found at http://www.hallikainen.org/FCC/FccRules/ 2008/90/303/section.pdf (footnote added by authors of this report).
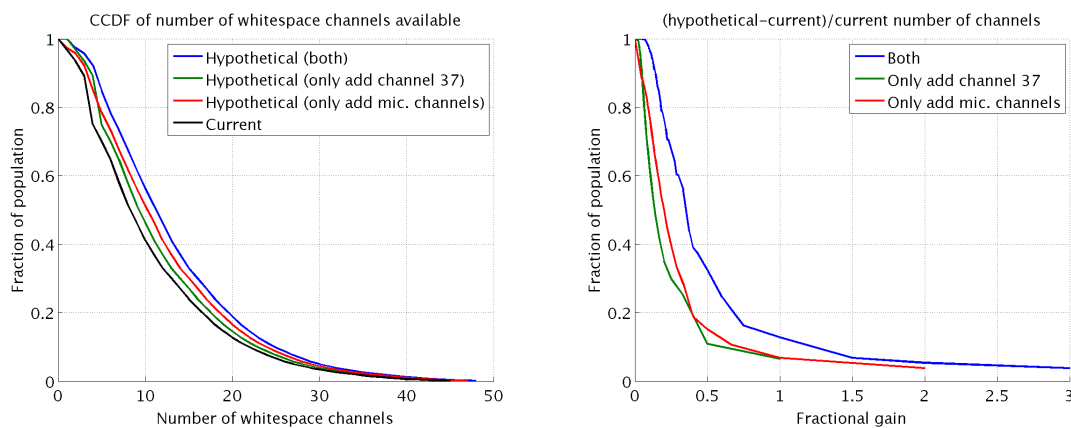
(a) Number of channels available under each scheme. (b) Fractional gain in the number of channels.

Figure A.10:

A.4.8 Methodology and assumptions

A.4.8.1 Wireless microphone channels

There appears to be a potential contradiction regarding which channels will be reserved for wireless microphones. In the discussion, it is stated that "[the FCC is] herein expanding the reservation of two channels in the range 14-51 to all markets nationwide" [9, ¶29]. However, [9, §15.707(a)] (quoted in the previous section) seems to imply the reserved channels are in the range 21-51.

For the purposes of this report, we assume that channels 2-13 are *never* reserved for unlicensed wireless microphones, even in the cases where this means that fewer than two channels are reserved for these microphones.

A.4.8.2 Wireless microphone channels and PLMRS/CMRS exclusions

In the discussion in [9, ¶29], the FCC notes that it is "expanding the reservation of two channels in the range 14-51 to all markets nationwide" which could be taken to imply that locations with PLMRS/CMRS exclusions will reserve no further channels for unlicensed wireless microphones.

However, the regulations themselves state wireless microphone channels should be reserved on available channels [9, §15.707(a)] (also quoted in the previous section). "Available" is defined [9, §15.703(a)], in part, as "acceptable for use by an unlicensed device" and subject to §15.711, which includes the PLMRS/CMRS exclusions [9, §15.711(d)]. Thus these rules are additive and may reserve up to a total of 5 channels.

For the purposes of this report, we use the latter of these two interpretations.

A.4.9   Future work and other questions

The analysis presented above is not intended to be 100% complete or accurate. In particular, the following questions remain unanswered:

- How large will the guard bands be?

- How much white space will be left in the guard bands once all major players have their equipment in place?

- What kind of protection do WMTS devices need? What about radioastronomy sites?

- Exactly how much which space will be left in channel 37 once we protect these sites and devices?

- How much whitespace will be available once wireless microphone operators routinely register their devices?

Appendix B

Background on repacking and the incentive auctions

Some or all of the work in this chapter appeared in [153].

A follow-on to Chapter 6, this chapter serves to detail our computational methods, describe the datasets we used, provide additional results, and to give the reader more detail on the incentive auction. We also describe some of the problems that arise when working with real-world data and how we overcome them.

B.1 Methods

In this section, we discuss the methods used for the calculations and analysis presented in Chapter 6. We begin by outlining the design and methodology of our repacker. In particular, we describe how we used the FCC's public datasets and a well-known SAT solver to output multiple feasible reallocations for each band plan. We also briefly touch upon some technicalities with respect to the SAT solver and how they impact the output of our repacker.

Next, we discuss the assumptions made while designing the repacker, and explain the reasoning behind these assumptions.

Finally, we describe our method for determining the amount of available whitespace, as well as viewable TV, for a particular reallocation. We also describe in detail how the figures in Chapter 6 were created. We make use of the FCC's public datasets and our open-source software, WEST, for these computations.

B.1.1 Details of our repacker

We used the FCC's domain, interference, and baseline files—released on 20 May 2014 at [65]—as a starting point. We then used PycoSAT, a Python wrapper for the well-known SAT solver library PicoSAT [77], in order to synthesize the constraints and output a feasible repacking. Two other studies on repacking have used PicoSAT [70, 72] and it was also featured in an FCC workshop on the topic of repacking in the incentive auctions [76].

The repacking problem can be naturally cast as a boolean satisfiability problem. The problem is to determine whether a set of stations can be packed into a given set of channels under the various interference and domain constraints. These constraints can be encoded as boolean expressions as follows. Let the set of stations be labelled $i = 1, \cdots, N$ and the set of channels be labelled $j = 1, \cdots, M$. For each station i and channel j, define the binary variable x_{ij}, which takes value 1 if station i is assigned to channel j and 0 otherwise. We have the following type of constraints on these variables:

- **Domain constraints:** For each station i, there is a set of channels $M_i \subseteq \{1, \cdots, M\}$ that it can be feasibly assigned to. These constraints can be expressed as follows:

$$\neg x_{ij} \text{ for each } i \text{ and each } j \notin M_i. \tag{B.1}$$

  Further, each station can be assigned to at most one channel in its set of feasible channels. This can be expressed as:

$$\neg x_{ij} \vee \neg x_{ir} \text{ for each station } i \text{ and every pair of channels } j, r \in M_i. \tag{B.2}$$

  Finally, each station must be assigned to *at least* one channel in its set of feasible channels. This is expressed as

$$\bigvee_{j \in M_i} x_{ij}. \tag{B.3}$$

- **Co-channel interference constraints:** For each channel j, there is a list of stations $S_j \subseteq \{1, \cdots, N\}$ such that no two stations in S_j can be assigned to channel j together. These constraints can be expressed as follows:

$$\neg x_{ij} \vee \neg x_{kj} \text{ for each channel } j \text{ and stations } i, k \in S_j. \tag{B.4}$$

- **channel interference constraints:** For each pair of adjacent channels j and $j + 1$ where $j \in \{1, \cdots, M-1\}$, there is a set of stations $A_{j,j+1} \subseteq \{1, \cdots, N\}$ such that no two stations in $A_{j,j+1}$ can be assigned to channel j and channel $j + 1$ or vice versa. These constraints can be expressed as:

$$(\neg x_{ij} \vee \neg x_{k(j+1)}) \wedge (\neg x_{i(j+1)} \vee \neg x_{kj}) \text{ for each pair of channels } j, j+1 \text{ and stations } i, k \in A_{j,j+1}. \tag{B.5}$$

The boolean expression that needs to be satisfied by the SAT solver is the conjunction of all clauses listed above. Note that this produces a boolean expression in conjunctive normal form (CNF), i.e. a conjunction of a series of disjunctions, which is the standard format in which SAT solvers accept their input.

In Figure 6.2 of Chapter 6 we also found the minimum number of stations that must be removed in order to meet a particular spectrum clearing target. This entails asking the following question: can a set of stations be packed into a given set of channels under the interference and domain constraints, assuming that up to a number K ($< N$) of these stations need not be assigned to any channel? One can start from a large value for K to obtain a "SAT" from the SAT-solver and subsequently decrease the value of K to find out the minimum value at which a "SAT" output is no longer obtained within a reasonable amount of time. We have borrowed this technique from [70]. The next section discusses why we use the "timeout" method as opposed to obtaining "UNSAT" directly.

This new condition can be encoded as a clause as follows. For each station $i \in \{1, \cdots, N\}$, one defines a binary variable x_{iH}, which takes value 1 if the station remains unassigned and 0 otherwise. In order to ensure that a station is not simultaneously assigned to some channel and $x_{iH} = 1$ (i.e. marked unassigned), we need to add the following set of expressions:

$$\neg x_{ij} \vee \neg x_{iH} \text{ for each station } i \text{ and every channel } j \in M_i. \tag{B.6}$$

Also since each station can now also remain unassigned, the clause $(B.3)$ gets modified to

$$\left(\vee_{j \in M_i} x_{ij}\right) \vee x_{iH}. \tag{B.7}$$

Further, not more than K of the set of variables $\{x_{iH} : i \in \{1, \cdots N\}\}$ can be assigned a value of 1, i.e.

$$\sum_{i=1}^{N} x_{iH} \leq K. \tag{B.8}$$

There are many ways of translating this type of a *boolean cardinality constraint* into CNF clauses. Following the precedent set in [70], we use the encoding described in [154]. This encoding essentially encodes a binary-logic adder into CNF clauses by adding intermediate variables. Interested readers should refer to [154] for details.

B.1.2 Timeout vs. UNSAT

SAT solvers typically yield an answer of either "SAT" or "UNSAT," indicating whether the problem was satisfiable or not. In the case of a "SAT", an actual assignment meeting the constraints is given as proof of satisfiability. If "UNSAT", the solver has found a proof of unsatisfiability and offers up information related to that proof.

However, the size and complexity of the repacking problem (see Section B.5) is such that both types of solutions may take a long time to obtain. In this case, we "time-out" the SAT solver (i.e.

halt its execution early) in the interest of time. This method has precedent: [70] set a time-out of 60 seconds (much shorter than any of ours). Typical run time of different SAT solvers on instances of repacking problems has also been studied in [74]. In our repacker, we set a time-out threshold that varied from 150 to 300 seconds, depending on the difficulty of the repacking problem.

We have further observed a well known quirk of SAT solvers: randomizing the order in which the variables x_{ij} are created considerably changes the time taken by the solver to find a satisfying assignment (if one exists). Because of this, for a particular time-out threshold, one attempt may lead to no output, but another attempt after randomization may lead to a "SAT". Hence, while trying to solve an instance, we make multiple randomized attempts (typically 5-10) with a fixed time-out threshold. If the SAT-solver still does not output a solution, we consider this to mean that the instance was *probably* unsatisfiable. Using this approach, the numbers that we obtain for the minimum number of stations to remove to clear spectrum for the different band plans are considerably lower than those reported in [70]. This indicates that the true minimum numbers may in fact be even lower. The only way to ascertain the true minimum numbers is to obtain an "UNSAT." But in all of our experiments, except for specific small instances, we very rarely obtained an "UNSAT" from the SAT solver, even with time-out thresholds as long as 30 minutes.

B.1.3   Repacking assumptions

We note the following assumptions made while designing our repacker:

- Stations currently assigned to a channel in UHF, whether the channel is above or below the spectrum clearing target, may be moved to any band; that is, to VHF or UHF channels.

- Stations currently assigned to a VHF channel may remain in the VHF band or be moved to the UHF band.

- Stations, whether above or below the spectrum clearing target, can either be repacked or relinquish their spectrum usage rights.

- All stations, regardless of their current channel assignment, are equally willing to relinquish their spectrum usage rights.[1]

In essence, we assume that all stations fully participate in the auction and we eliminate the distinction between VHF and UHF. Although this may not accurately reflect the incentive auction scenario, it accurately represents the generic spectrum reallocation scenario which is the focus of Chapter 6.

---

[1]In reality, this is not entirely true. Many stations may not even be willing to participate in the incentive auction. It is difficult to predict this behavior, and so we do not consider the possibility here.

If we wish to mimic or study the outcome of the incentive auctions, there is a wide variety of assumptions that can be made. For instance, many opine that during the incentive auction, stations will be reluctant to move to lower bands, in particular from UHF to VHF. This restriction is easily implemented in our repacker. As another example, leaving stations below the spectrum clearing target untouched and attempting to repack only stations above the clearing target would minimize the additional costs required to relocate repacked stations after the auction.

However, we focus on the efficient clearing method with the assumptions listed above because it places minimal restrictions on the repacking process. This repacking flexibility results in the fewest possible stations being taken off the air. This represents a best-case scenario for viewable TV, and a worst-case scenario for available whitespace as stations will now be efficiently packed into the channels below the clearing target, reducing the number of spectrum holes.

We compare the efficient clearing method with a naive clearing method, where stations above the clearing target simply relinquish their spectrum usage rights, and stations below the clearing target are left completely untouched. The naive clearing method results in the most stations being taken off the air because there is no attempt at repacking. This represents the worst-case scenario for viewable TV, and a best-case scenario for available whitespace as the inefficient packing of stations in the channels below the clearing target will create a larger number of spectrum holes.

With these two methods, we have explored the best-case and worst-case scenarios for both whitespace and TV channel availability. Although they may not perfectly mimic what is expected to happen in the actual incentive auction, they prove useful in understanding the general problem of spectrum reallocation. Future work includes modifying our repacker to study the incentive auctions themselves.

B.1.4   Description of datasets and figures

In this section, we describe our methodology for computing the amount of available whitespace (in channels or MHz) and the amount of viewable TV (in channels) for a generic allocation of stations in a region. To do this, we have built extensions for our in-house, open-source software, Whitespace Evaluation SofTware. As described in Chapter 11, WEST provides general-purpose tools for computing the amount of whitespace and TV for a generic region, and a generic set of protected entities (e.g. TV stations and Private Land Mobile Radio Services (PLMRS)). In this section we provide a brief description of the datasets used for our calculations, the building blocks of our extensions to WEST, and the process by which we generated the figures in Chapter 6.

Datasets used

Our work utilizes a variety of publicly-available datasets. This section describes these datasets, how we used them, and where to find them.
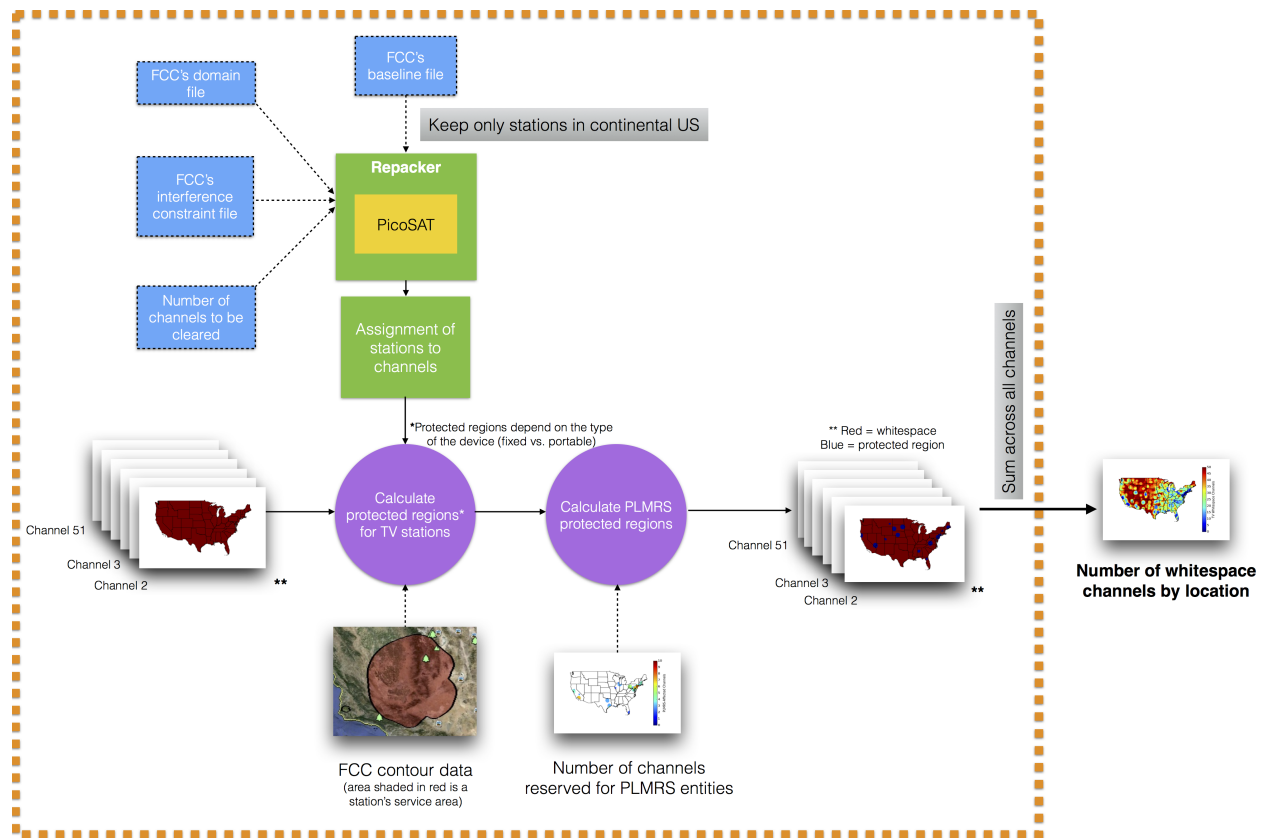
Figure A.1: Data flow diagram showing the process used to evaluate the number of available whitespace channels at every location in the US. Note that this process is executed for a particular band plan and for a particular repacking scenario.

- **FCC's incentive auction baseline file [65]:** This file, released by the FCC, contains basic information about the TV stations which are eligible to participate in the incentive auction.

- **Datasets for PLMRS entities [155]:** We also consider restrictions due to other incumbent services, specifically PLMRS (Private Land Mobile Radio Services). These restrictions primarily affect metropolitan areas.

- **FCC's domain and interference constraint files [65]:** These FCC datasets give us the explicit domain and interference constraints for each TV station as described in Section B.1.1.

- **FCC's TV station service area data [156]:** The FCC has also made available the official service area polygons of TV stations (defined by 360 pairs of latitude and longitude coordi-

nates). From this, we can determine the protected contour polygons by adding a buffer of the appropriate size (separation distance) for the whitespace device in question[2].

- **United States 2010 census data [10,11]:** We use the census data to create a discretized map representing the population of the United States. This population map is used in weighted statistical metrics (for example, mean, median, and CDF) of available whitespace or TV in the United States.

Basic whitespace- and TV-availability calculations

As mentioned in Chapter 11, WEST is a free, open-source, and general-purpose software for whitespace studies. It has well-defined modules that are designed to independently support generic geographic regions, protected entities, and rulesets. These modules are easily customized for the US, and for pre- and post-incentive auction scenarios.

Figure A.1 outlines our core algorithm for computing whitespace availability. We begin by reading from the FCC's baseline file [65], considering only stations within the continental US. We then provide these stations, along with the FCC's domain and interference constraints [65] to our repacker, a thin wrapper on top of PicoSAT. The repacker outputs a feasible repacking for the given band plan. This process is shown in the top half of Figure A.1.

From the re-assignment of TV stations obtained from the repacker, and the service area polygons in [156], we can determine the protected regions for each TV station, and therefore protected locations on each channel. We also mark off regions that are further protected for PLMRS entities. Finally, we obtain maps that show how much whitespace is available at each location. This process is shown in the bottom half of Figure A.1.

The process for evaluating the amount of viewable TV is extremely similar to the above. On every channel, locations within the service area of a TV station are considered "TV-viewable." Therefore, we can similarly use the FCC's service area data to determine TV-viewable locations on every channel, and thereby obtain maps showing how many TV channels are viewable across the US.

Both of these processes are used extensively as the baseline data for the generation of our other figures. **Figure 5 in Chapter 6** shows how much whitespace is available at each location in the continental United States after an efficient re-allocation. We can also create "loss maps" that show us where and how much whitespace/TV is lost post-incentive auction, as seen in **Figure 3 of Chapter 6**.

---

[2] We were able to make use of these for almost all stations. For about 2% of the stations, there were discrepancies in the data which forced us to fall back to an inferior method of calculating the protected region. We discuss this in more detail in Section B.3.

Figure generation

Here, we briefly describe the process used to generate Figures 6-10 in Chapter 6. These figures are crucial in understanding the effect of repacking on available whitespace and viewable TV.

- **Figures 6, 7, and 8 in Chapter 6** are 2-D histograms that represent the distribution of people who have a certain number of available whitespace channels pre-auction and a certain number of available whitespace channels post-auction. Figure A.2 in this section describes the process by which these figures were made. For every possible ordered pair describing the number of whitespace channels pre- and post-auction, (x, y), we determine the number of people that have x available whitespace channels pre-auction and y whitespace channels post-auction. We then color the pixel at position (x, y) according to the number of people. For robustness, we calculate the pixel-wise average of these 2-D histograms over multiple repackings for a given band plan to obtain the final histograms presented in Chapter 6.

- **Figures 9 and 10 in Chapter 6** show plots of the amount of viewable TV versus the amount of available whitespace for various spectrum clearing targets. Figure A.3 in this section describes the process by which these figures were made. We use the whitespace availability map obtained according to Figure A.1 along with population data to determine how much whitespace is available to the median person in the US in a particular repacking scenario. Similarly, we use the TV availability map to determine how much TV is viewable to the median person in the US for a particular repacking scenario. We repeat this process for several potential repackings for a given spectrum clearing target, and then for all possible spectrum clearing targets, and plot the amount of viewable TV versus the amount of available whitespace spectrum (in MHz).

B.2 Television availability after repacking

One way of looking at the impact of the incentive auction is to examine which places lose access to TV under various repacking scenarios. Figure 3 in Chapter 6 showed which places lose at least one TV channel under the naive clearing method (orange), the efficient clearing method (green), or both (blue) when 14 channels are to be cleared for LTE.

We present here more detailed maps showing how many TV channels were lost in each market. Figure A.4 makes the difference between the naive clearing method and the efficient clearing method even more stark. With the efficient clearing method, only San Francisco, Los Angeles, and the most populous portions of the East Coast lose more than four channels. With the naive clearing method, a variety of regions (including Utah, for example) lose six or more TV channels.

Beyond answering the question of how much TV coverage will be lost, these more detailed figures are important because they give insight into how the repacking will work. In particular, we see that
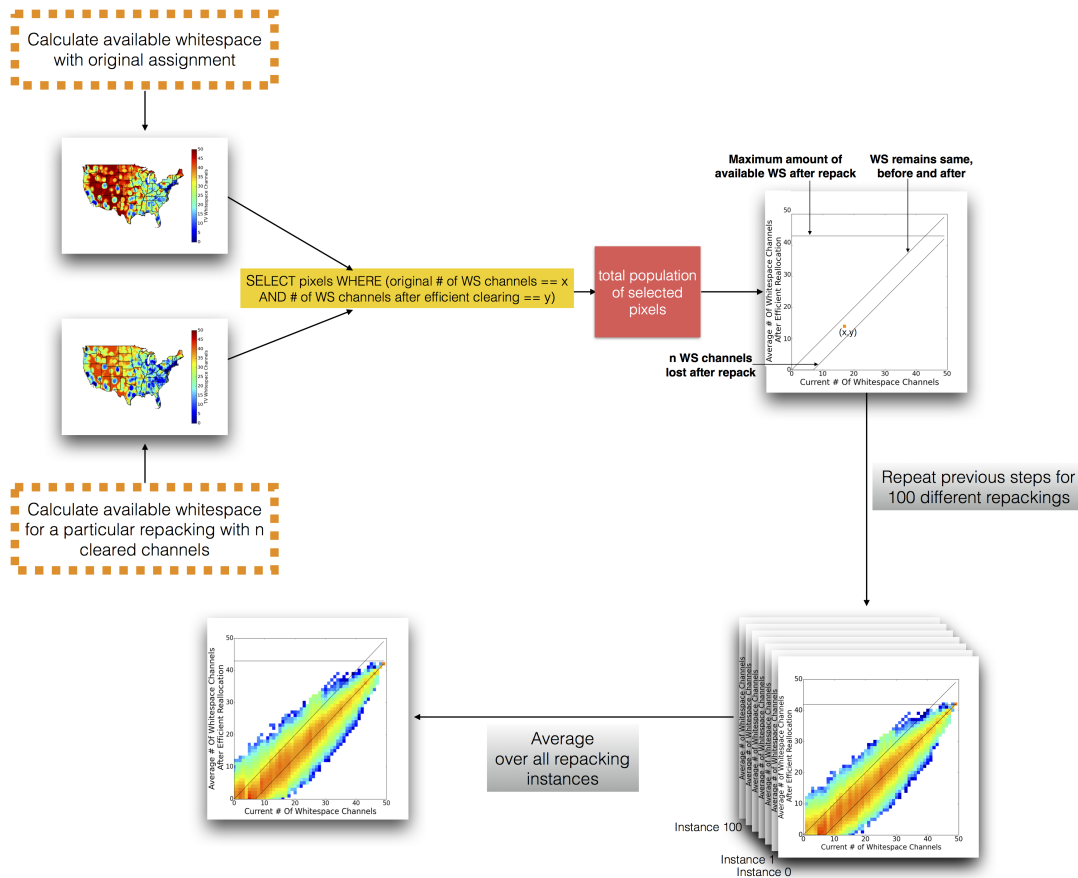
Figure A.2: Data flow diagram showing the process used to generate Figures 6, 7 and 8 in Chapter 6. Note that this process is done for a particular band plan.

the areas that currently have a lot of TV stations (e.g. New York City, Los Angeles) have so many that some must be removed rather than repacked. However, most of the country is not brimming with TV stations (as evidenced by the current amount of whitespace in these regions) and so no stations would need to be removed to meet most clearing targets.

B.3   Addressing the problems in real-world datasets

In this section, we discuss apparent discrepancies and errors in datasets provided by the FCC. The errors mainly affect our assessment of the service areas (and hence protected contours) of TV stations, and therefore our estimate of the available whitespace and/or viewable TV. We describe techniques we have used to eliminate, or at least minimize, the effects of these errors and discrepancies.
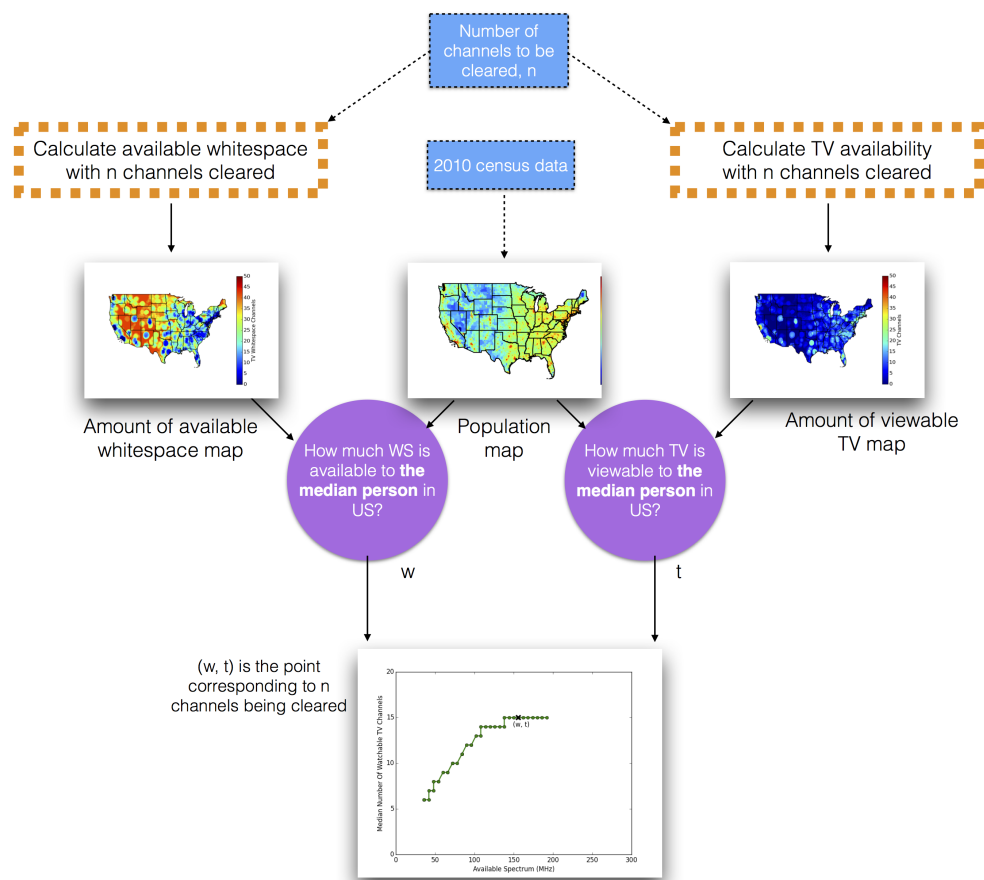
Figure A.3: Data flow diagram showing the process used to create Figures 9 and 10 in Chapter 6.

B.3.1   TV station records with missing HAAT data

About 18% of the records in our TV station dataset had no height above average terrain (HAAT) data listed. According to an FAQ entry posted on Aug. 23, 2011 at http://www.fcc.gov/encyclopedia/white-space-database-administration-q-page, the HAAT for records missing this value should be calculated from the RCAMSL (Height of Antenna Radiation Center Above Mean Sea Level) in conjunction with the terrain database. Thus we updated the HAAT entries for the relevant records using the FCC's HAAT calculator at [157].

B.3.2   Stations without service contour data

The baseline file [65] and the FCC's service contour data points file [156] contain information for slightly different sets of stations. The baseline file contains the list of stations eligible to participate in the incentive auction whereas the contour data points are given for all active records in the FCC's TV Query database. Furthermore, the baseline file is a snapshot from 20 May 2014 while the

(a) Efficient clearing method   (b) Naive clearing method
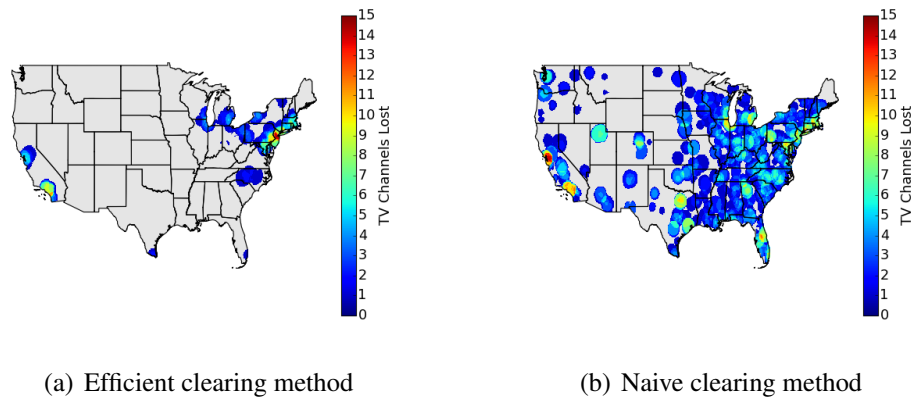
Figure A.4: Map showing how much TV coverage is likely to be lost across the US after the incentive auctions if 14 TV channels are repurposed using the efficient clearing method. Gray denotes areas whose TV coverage is not affected.

service contour data is updated daily and our copy is from 17 June 2014. About 2% of the stations in the baseline file updated their licenses between these two dates (e.g. some stations have changed physical location or transmit power), meaning that the service contours may not correspond to the station parameters given in the baseline file and used when generating the repacking constraint files. Thus it is not possible to obtain official FCC service contours for all auction-eligible stations.
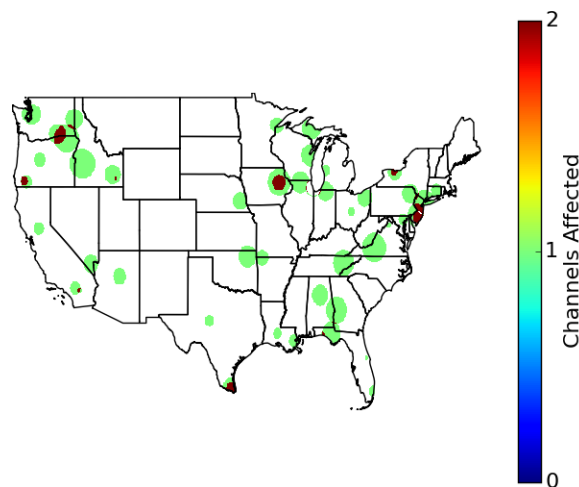
Figure A.5: Map indicating areas in the United States which are serviced on one or more channels by stations where we use circular contours.

For these stations without official FCC service contours, we create approximate contours. Al-

though we use the FCC's F(50,90) propagation curves [158] to create these approximations, we do not incorporate the terrain data which is necessary for reproducing the official contours[3]. Instead, we assume flat terrain which results in circular service contours. In relatively flat areas, our approximate contours will closely match the official FCC contours. In areas with rough or mountainous terrain, the official contours are generally not circular and thus our approximations will have greater error. The difference is shown for two example stations *with* official contours in Figure A.6. Light green and dark blue indicate disagreement between the approximate and official contours whereas gray and red indicate agreement.

Figure A.5 indicates the locations which are serviced by these stations *according to the approximate (circular) contours*. We see that about 29% of the population is serviced by these circular contours.

(a) Flat terrain   (b) Rough/mountainous terrain
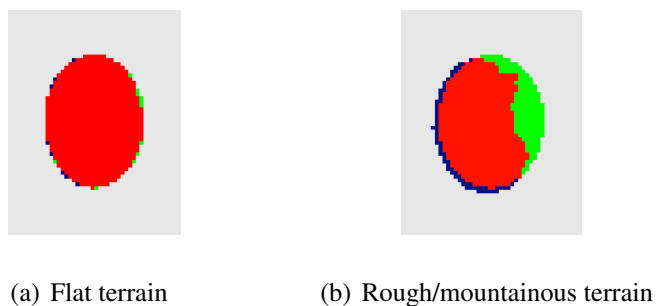
Figure A.6: Comparison of protected contours provided by the FCC and approximate circular contours for different types of terrain.

B.3.3   Overlap in stations' service contours

While calculating protected regions for stations across the US, we observed that a few service contours overlapped for stations on the same channel. We believe this is because the FCC model for television reception allows the receiver to point his antenna towards the station he wishes to receive. This allows for the theoretical possibility of locations where two or more stations can be received on the same channel by pointing the receiver's antenna in different directions. At such locations, there are two or more viewable TV stations on the same channel. Figure A.7 highlights these locations in blue and makes it clear that the "overlap effect" occurs in a small portion of the United States.

When evaluating the amount of viewable TV at each location in the US, we take the effect of this overlap into account and count all visible stations at that location. The effect also shows up in locations where there was an overlap in the current allocation but no overlap after repacking. These locations will now "lose" an extra whitespace channel, even in cases where *no* channels

---

[3]Although WEST supports the use of terrain data, we have not obtained and integrated terrain data into our code yet.
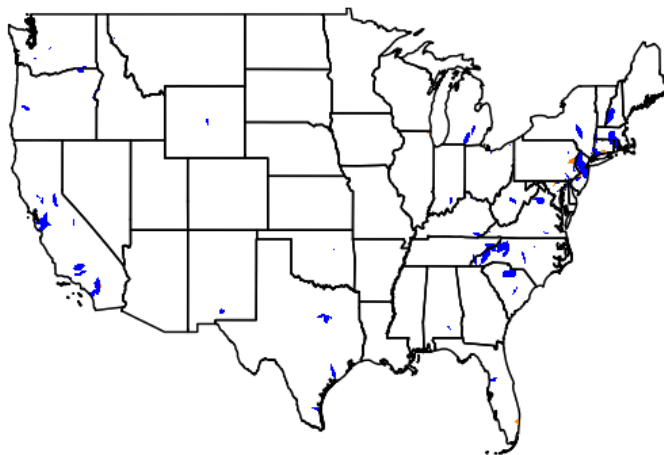
Figure A.7: Map indicating locations in the United States where there are two or more effectively viewable TV stations on any channel. Blue represents overlap between stations' contours given by the FCC, and orange represents an overlap between at least one circular contour.

are cleared. Conversely, locations where there was originally no overlap on any channel but at least one overlap present after repacking will "gain" at least one extra whitespace channel. These overlap effects explain why a few places lose more than n channels of available whitespace when n channels are cleared for LTE, and why a few places gain whitespace channels. These can be seen in the points that lie outside of the diagonal lines in Figures 7 and 8 of Chapter 6.

B.4   Overview of the incentive auction

We provide a brief overview of the incentive auction background and procedures as a convenience to the interested reader. Although our work in Chapter 6 heavily uses the data released in connection with the incentive auction, it does not study the auction itself. Studies of the incentive auction are intentionally separated from Chapter 6, which is a study of the more general problem of spectrum reallocation.

B.4.1   Background

Until the 1990s, spectrum allocations in the United States were assigned by lottery. When demand began to exceed supply and buyers demanded greater certainty about allocation outcomes, the Federal Communications Commission (FCC) switched to an auction-based system. In this system, spectrum leases were auctioned off to the highest bidder. When the lease expired, the spectrum went back on the market.

However, we have recently seen that the turnover rate in this system is too slow. Rather than wait for new spectrum to come to auction, companies are instead purchasing smaller companies purely for their spectrum holdings. When spectrum *is* available, it commands a hefty price as we saw in the AWS-3 auction that concluded in early 2015, in which carriers collectively paid almost $45 billion for 65 MHz of nationwide spectrum [19].

This, along with the oft-touted wireless boom, demonstrates the great need for new LTE-friendly spectrum. However, there are no upcoming bands where older licenses are expiring, especially at lower frequencies, so the FCC had to get a little creative. In particular, they turned their eye to over-the-air television broadcasting spectrum which is licensed in a fundamentally different way from most spectrum (briefly: broadcasters did not pay for their licenses because they are serving the public good). This left the door open to repurpose their spectrum with minimal cooperation from the broadcasters.

However, the FCC cannot simply break the broadcasters' leases. Instead, an act of Congress [67] passed in February 2012 gave them the authority to conduct a novel double-auction. In November 2012, the FCC put out a Notice of Proposed Rulemaking (NPRM) [68] for these so-called "incentive auctions" and they have since received over 2000 comments. A Report and Order [13] with further auction details was published in June 2014. Although the R&O established the overall framework for the auction, the smallest yet most important details were left for Public Notices (PNs). The first of these PNs was published in December 2014 [69] and the second is pending. In the FCC's own words:

> Well in advance of the auction, also by Public Notice, the Commission will resolve these implementation issues, and provide detailed explanations and instructions for potential auction participants ("Incentive Auction Procedures PN" or "Procedures PN"). [13, ¶15]

This section briefly explains what is currently known about the auction process. The auction procedure is presented as an algorithm in Algorithm 2.

B.4.2 Reverse auction

The first half of the incentive auction is the "reverse auction" in which TV broadcasters indicate the price at which they would be willing to relinquish their spectrum usage rights[4]. The auction will have a descending clock format [13, ¶447], meaning that bidders (broadcasters) will be offered successively lower bids for their spectrum[5]. If the price drops too low for a broadcaster, it can

---

[4]This is not precisely true. Broadcasters are actually given several choices [13, ¶365] but here we simplify to the most important choice.

[5] In a reverse auction, multiple potential sellers compete to *sell* their goods or services to a single buyer, as opposed to the better-known case of the forward auction, in which multiple buyers compete to *buy* goods or services from a single seller. The descending-clock reverse auction is just the reverse auction counterpart of the well known ascending clock forward auction, also known as the English auction, in which bidding according to true preferences is a weakly dominant strategy. See [159] for more details.

Algorithm 2 Incentive auction procedure

Require: M_{excess}: the amount of money needed in addition to the cost for the reverse auction (mandatory expenses)
Require: $M_{MHz-pop}$: reserve MHz-pop price ([69], ¶47] suggests \$1.25/MHz-pop)
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ INITIALIZATION

1: Broadcasters given initial bids from the FCC
2: Broadcasters submit applications to participate in the auction
3: Set $S_{\text{clearing target}}$ (in number of paired LTE channels) based on broadcaster interest
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ REVERSE AUCTION

4: **function** REVERSEAUCTION($S_{\text{clearing target}}$)
5: $\quad$ Offer broadcasters large amounts of money
6: $\quad$ **while** clearing target met **do**
7: $\qquad$ Offer each broadcaster slightly less money (customized offers)
8: $\qquad$ Receive bids from broadcasters
9: $\qquad$ Determine which bids to provisionally accept
10: $\quad$ **end while**
11: $\quad$ Accept the last bids for which the clearing target was met
12: $\quad$ **return** ReverseMoneyNeeded
13: **end function**
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ FORWARD AUCTION

14: **function** FORWARDAUCTION($S_{\text{clearing target}}$)
15: $\quad$ **while** demand $>$ supply **do**
16: $\qquad$ Increase prices from previous round
17: $\qquad$ Each bidder indicates the number of generic LTE bands it wants in each market at that price
18: $\quad$ **end while**
19: $\quad$ **return** ForwardMoneyRaised
20: **end function**
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ FINAL CONDITION FOR AUCTION SUCCESS

21: **function** FINALCONDITIONMET(ReverseMoneyNeeded, ForwardMoneyRaised, $S_{\text{clearing target}}$)
22: $\quad$ **if** (ReverseMoneyNeeded + $M_{excess} \leq$ ForwardMoneyRaised) AND (MHz/pop($S_{\text{clearing target}}$) $> M_{MHz-pop}$) **then**
23: $\qquad$ **return** True
24: $\quad$ **else**
25: $\qquad$ **return** False
26: $\quad$ **end if**
27: **end function**
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ COMBINING THE REVERSE AND FORWARD AUCTIONS

28: AuctionSucceeded $\leftarrow$ False
29: **while** $S_{\text{clearing target}} \geq 2$ LTE channels AND not AuctionSucceeded **do**
30: $\quad$ ReverseMoneyNeeded $\leftarrow$ REVERSEAUCTION($S_{\text{clearing target}}$)
31: $\quad$ ForwardMoneyRaised $\leftarrow$ FORWARDAUCTION($S_{\text{clearing target}}$)
32: $\quad$ **if** FINALCONDITIONMET(ReverseMoneyNeeded, ForwardMoneyRaised, $S_{\text{clearing target}}$) **then**
33: $\qquad$ AuctionSucceeded $\leftarrow$ True
34: $\quad$ **else**
35: $\qquad$ $S_{\text{clearing target}} \leftarrow S_{\text{clearing target}} - 1$
36: $\quad$ **end if**
37: **end while**
$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ ▷ POST-AUCTION

38: **if** AuctionSucceeded **then**
39: $\quad$ Forward auction winners bid for specific spectrum blocks
40: $\quad$ Repacking optimized
41: **else**
42: $\quad$ Apply egg directly to face
43: **end if**

"drop out" of the auction, meaning that it will receive no payment but will keep its spectrum usage rights. If the broadcaster's bid is eventually accepted by the FCC, it will be paid and will be required to cease transmission within three months.

Broadcasters that "drop out" of the auction will be treated equivalently to broadcasters which choose not to participate in the auction. In both cases, the FCC has the authority to "repack" (i.e. relocate in frequency) these stations, compensating broadcasters only for the cost of the move itself. To minimize the negative impact on broadcasters and their viewers, a station will only be moved within its original band (e.g. UHF) and will have its transmission parameters adjusted so that its coverage area remains approximately the same.

B.4.3 Forward auction

The second half of the incentive auctions, called the "forward auction," is in many ways a standard spectrum auction; that is, it will have an ascending-clock format. At each round, bidders will indicate how many 2×5 MHz bands (i.e. LTE uplink + downlink) of spectrum they would buy and at what price. Expected bidders include AT&T, Verizon, T-Mobile, Sprint, and possibly Dish Network.

Again, we have slightly simplified our description of this stage to improve readability. For example, the FCC will be using two types of generic spectrum blocks ("impaired" and "unimpaired") and will use a post-auction auction to determine which blocks go to which winning bidders. Interested readers should read the source documents for full details [13, 68, 69].

B.4.4 Combining the reverse and forward auctions

Since Congress did not authorize the FCC to make any payments itself, the reverse and forward auctions must be sufficiently intertwined to ensure that the proceeds from the forward auction can pay for the reverse auction[6]. In particular, the FCC will set a spectrum clearing target (i.e. number of TV channels to be repurposed) based on initial interest indicated by broadcasters. They will then conduct a reverse auction, provisionally "buying out" enough broadcasters to meet the clearing target. A forward auction is then held with the goal of auctioning the provisionally-cleared spectrum for enough money to actually fund the reverse auction. If the forward auction successfully raises the funds to pay the provisional reverse auction bids, all bids are officially accepted and the auction concludes. If not, the FCC reduces the spectrum clearing target and begins again, starting with the reverse auction.

---

[6]In addition, the proceeds must cover reallocation costs and costs of running the auction [13, ¶341], amounting to $2 billion.

B.4.5 Remaining details

B.4.5.1 Specific band plans

The FCC has created specific spectrum clearing targets, each called a "band plan," as can be seen in Figure A.8, which is Figure 23 in [13]. Each row represents a unique band plan, with the lowest clearing target at the top and the highest at the bottom. The paired LTE channels are shown in blue with the higher frequencies being designated for uplink and the lower frequencies for downlink. Remaining TV channels are shown in white (channels 2-20 are not shown but will remain allocated for TV). Band plans include spectrum for the duplex gap (11 MHz between the LTE up- and downlink bands), a guard band between the LTE downlink and remaining TV channels (7-11 MHz, depending on the plan), and a 3-MHz buffer around channel 37 (used by sensitive radioastronomy equipment as well as some medical devices deployed in hospitals) as needed.

2	42	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	11	A	B	11	A	B	700 MHz UL
3	48	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	7	A	B	C	11	A	B	C	700 MHz UL
4	60	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	9	A	B	C	D	11	A	B	C	D	700 MHz UL
5	72	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	11	A	B	C	D	E	11	A	B	C	D	E	700 MHz UL
6	78	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	7	A	B	C	D	E	F	11	A	B	C	D	E	F	700 MHz UL
7	84	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	3	A	B	C	D	E	F	G	11	A	B	C	D	E	F	G	700 MHz UL
8	108	21	22	23	24	25	26	27	28	29	30	31	32	11	A	B	3	37	3	C	D	F	F	G	H	11	A	B	C	D	E	F	G	H	700 MHz UL
9	114	21	22	23	24	25	26	27	28	29	30	31	7	A	B	C	D	3	37	3	E	F	G	H	I	11	A	B	C	D	E	F	G	H	I	700 MHz UL
10	126	21	22	23	24	25	26	27	28	29	9	A	B	C	D	E	F	3	37	3	G	H	I	J	11	A	B	C	D	E	F	G	H	I	J	700 MHz UL
11	138	21	22	23	24	25	26	27	11	A	B	C	D	E	F	G	H	3	37	3	I	J	K	11	A	B	C	D	E	F	G	H	I	J	K	700 MHz UL
12	144	21	22	23	24	25	26	A	B	C	D	E	F	G	H	I	J	3	37	3	K	L	11	A	B	C	D	E	F	G	H	I	J	K	L	700 MHz UL

Figure A.8: FCC band plans from [13]. Each row represents a specific band plan. The left column of numbers represent the number of 2×5 LTE channels created. The second column lists the number of repurposed MHz. TV spectrum is shown in white and LTE spectrum in blue.

B.4.5.2 "Repacking" TV stations

A major component of the reverse auction that is at first hidden from the reader is that it is essential to be able to relocate TV stations to another channel (within reason) without paying them more than relocation costs. For example, consider a TV station on channel 51, the highest-frequency TV channel. Removing this TV station from channel 51 (either buying its spectrum usage rights or relocating it to another channel) is absolutely necessary for *any* of the band plans to be feasible. If given the option to name a price, the broadcaster could essentially hold the entire auction hostage. This same reasoning applies to any broadcaster in a channel that is included in the spectrum clearing target.

As such, the Spectrum Act also gives the FCC the authority to relocate TV stations within their "home band" (e.g. if a station is currently allocated a channel in the UHF band, it must either be paid or be relocated within the UHF band) [67, §6403(b)(1)(B)(i)]. Although it sounds simple, the

repacking is actually critical to the auction.

Repacking also provides a great opportunity for researchers. In order to be able to conduct the auction in real time, the FCC preprocessed a large amount of data about the potential for co- or adjacent-channel interference between TV stations. They also analyzed which stations have "domain constraints." For example, a station near the Canadian border cannot be assigned to a channel being used by a nearby Canadian station (non-US stations may not participate in the auctions). It is precisely this data which we repurposed in Chapter 6, since it completely enumerates the constraints on the efficient repacking of TV stations into existing spectrum. The following section gives further details about the data.

B.5 Exploring incentive auction data

While working closely with the FCC's incentive auction data, we noticed several interesting phenomena. Although our work in Chapter 6 is intended to be somewhat separate from the incentive auction itself, these phenomena will affect our results. Furthermore, we feel strongly that any interesting patterns or potential errors in the incentive auction data should be made public.

B.5.1 Domain constraints

A station's domain is the set of channels to which it *may* be assigned during the repacking. For many stations, the domain will be the entire set of TV channels. However, some stations, particularly those along the borders, will have constrained domains:

> In order to protect LM base stations, LMW base stations, Mexican allotments, Canadian allotments and Channel 37, the FCC staff had to consider these as fixed constraints, which will limit the channels on which any U.S. television station can be assigned or reassigned in the incentive auction repacking process. [160, §3.3]

The FCC has published the domains for all stations that may participate in the auction at [65]. Note that the domain data is independent of the spectrum clearing target since it only identifies channels on which a station could be placed *in the context of protecting non-TV incumbents and non-US TV stations*.

Figure A.9 shows the number of stations with domain of size D. The majority of stations (approximately 57%) have no domain restrictions. All but about 12.5% have a domain size greater than 40.

Figure A.10 illustrates the domain size as a function of location. Each marker represents a station. White stations are completely unconstrained and may be placed on any TV channel. Blue stations are nearly unconstrained whereas red stations are extremely constrained.
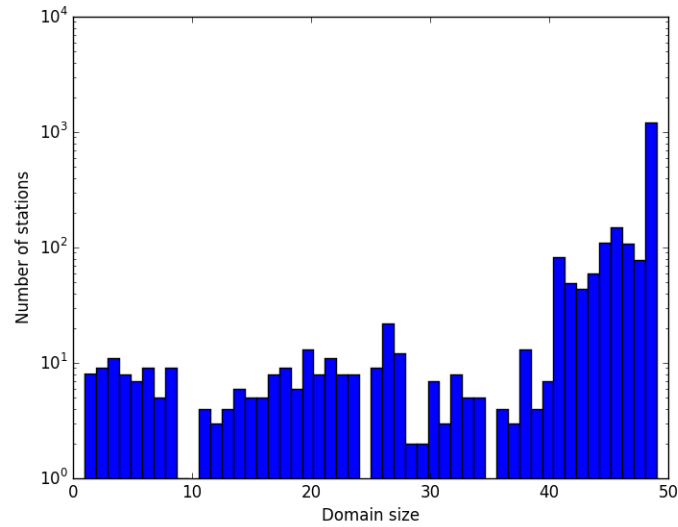
Figure A.9: Frequency distribution of domain sizes of stations that may participate in the incentive auction.
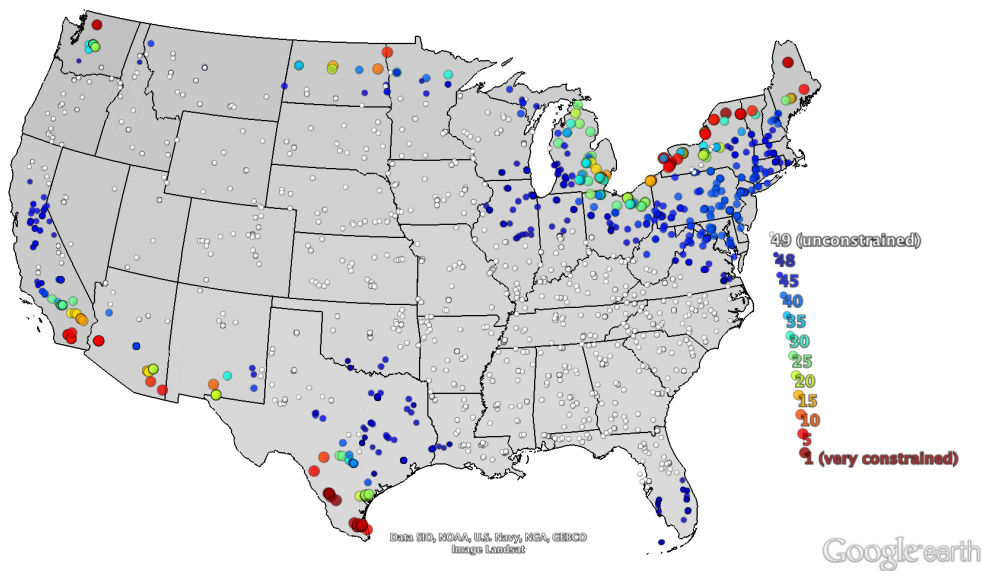

Figure A.10: Color-coded sizes of domains of stations plotted with Google Earth. The source KML can be found at http://www.eecs.berkeley.edu/~sahai/icc2015-incentive-auction-domain-constraints.kml

We see that the majority of the constraints are coming from the border protections with Canada and Mexico. For example, Los Angeles, southern Texas, and upstate New York are particularly affected.

In other places such as San Francisco and New York City, most if not all domain constraints come from the LM and LMW protections. Data detailing these protections can be downloaded from [65]. These affect only channels 14-20. This can be seen most clearly in Figure A.11. While the protected stations in Canada and Mexico have no particular affinity for one channel over another, there is an obvious decrease in the number of stations that have channels 14-20 in their domain.

Note that channel 37 is not in the domain for any TV station since it is used for radioastronomy, which involves receivers that are too sensitive to coexist with TV stations.
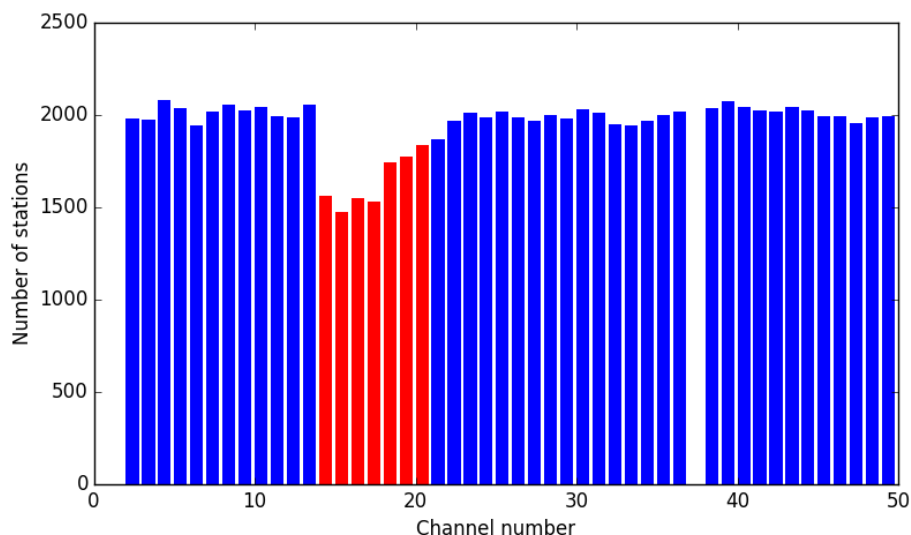

Figure A.11:   Number of stations with a particular channel C in their domain. Channels 14-20, which are affected by the LM and LMW protections, are shown in red.

B.5.2   Possible errors in the domain constraints

Figure A.10 displays some data that appears to be inconsistent with the protection criteria. For example, there are several white (i.e. completely unconstrained in terms of domain) stations in New York, situated among many rather constrained stations. One of the white stations (on the NY-Pennsylvania border) is even located on the same tower as a constrained station, although the difference could be attributed to the relatively low transmit power of the unconstrained station. There are similar apparent anomalies in the Upper Peninsula of Michigan, the North Dakota-Minnesota border, northern Idaho, and near Austin, Texas. Differences in transmit power cannot

explain all of these apparent anomalies: for example, the transmit power of the unconstrained station in the Upper Peninsula of Michigan (facility ID 81448) is over three times that of the nearest constrained station (facility ID 59281) and it is approximately the same distance from the border.