

Crime and Punishment for Cognitive Radios

by

Kristen Ann Woyach

B.S. (University of Notre Dame) 2006

A thesis submitted in partial satisfaction
of the requirements for the degree of

Master of Science

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

Fall 2008

The thesis of Kristen Ann Woyach is approved.

Chair

Date

Date

University of California, Berkeley

Fall 2008

Crime and Punishment for Cognitive Radios

Copyright © 2008

by

Kristen Ann Woyach

Abstract

Crime and Punishment for Cognitive Radios

by

Kristen Ann Woyach

Master of Science in Engineering - Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Anant Sahai, Chair

In this thesis, we aim to understand the amount of overhead required to enforce spectrum regulation. The current command-and-control model has a large overhead. The current regulation rules are very easy to enforce, but they pay for this ease by being inflexible and inefficient. Other methods of spectrum regulation have been proposed, but because they rely extensively on technical assumptions that have not been validated, it is not clear how much overhead will be required to enforce them. In this thesis we identify the technical questions that will need to be answered in order to use the proposed schemes. We also note that if we can pose regulation difficulties as technical questions, we may be able to quantitatively understand what overhead we must pay to enforce. Consequently, we may be in a better position to make regulatory decisions.

We begin the process of examining enforcement overhead by introducing a game-theoretic model to capture enforcement of primary-secondary cognitive radio rules. To incentivize following the rules, we introduce a “spectrum jail” which stipulates that users caught cheating must turn off for a specified amount of time. This model is used to understand the interaction of enforcement parameters such as the likelihood of getting caught and the time a radio must spend in jail. We find that this kind of incentive scheme follows well regulatory goals of encouraging technical innovation and being simple to certify. We also find that the single most important parameter for determining

how effectively a cognitive user can fill spectrum holes is the probability that they will be wrongfully sent to jail.

Professor Anant Sahai
Thesis Committee Chair

To my family

Contents

Contents	ii
List of Figures	iv
List of Tables	vi
Acknowledgements	vii
1 Introduction	1
2 Spectrum policy: historical formulation	4
2.1 Command-and-control	4
2.2 The property-rights critique	6
2.2.1 Defining property rights	6
2.2.2 Partial adoption with auctions	8
2.3 The commons critique	9
2.3.1 Arguments against the precursors: constitutionality and the Internet	9
2.3.2 Defining the commons	10
2.4 The hybrids critique	12
2.4.1 Defining hybrid allocation schemes	12
2.4.2 The first steps toward adoption	14
2.5 What the future holds: lessons from the Internet	14
3 Spectrum policy: underlying assumptions	17
3.1 What is scarcity?	19
3.2 Reliance on the infrastructure	21
3.2.1 Broadcast	21
3.2.2 Is it all just the Internet?	23
3.3 Number of players involved	25

3.3.1	Coordination	25
3.3.2	Dispute resolution	27
3.3.3	How can technology help?	28
3.4	Government regulation	29
3.5	Does technology even allow us a choice	30
3.6	Summary of the critique	31
4	Bridging the gap	34
4.1	Code as Law	35
4.2	Porting to spectrum policy	36
4.3	Enforcement for cognitive radio	37
4.3.1	Big Brother	38
4.3.2	More flexible approaches	39
4.4	Summary so far	42
5	Deterrence perspective	43
5.1	Single Band	45
5.1.1	The model and basic regulatory conditions	45
5.1.2	A deeper look into the effect of the enforcement parameters	50
5.1.3	The meaning of β	62
5.2	Multiple Bands with Global Jail	64
5.2.1	Perfect Justice	64
5.2.2	Imperfect Justice	65
6	Concluding remarks	71
	Bibliography	74
	References	74
A	Markov Chain for Simulations	77

List of Figures

5.1	Markov model for enforcement in a single band: the primary user follows a transmission pattern characterized by the two-state chain at the top. In response to this usage, the cognitive user chooses between the actions in the bottom chain. When the cognitive user is cheating (i.e. using the band when the primary is also transmitting), it has a probability of being sent to jail. While in jail, the cognitive user is unable to use the band until a timer (determined by P_{pen}) runs out.	45
5.2	P_{FA} vs P_{MD} tradeoff curves for an energy detector. Bold lines are at SNR = $-6dB$ with number of samples $N = 400$ and $N = 1000$	46
5.3	Typical cheating behavior for a cognitive user. If it is worthwhile to cheat, it is worthwhile to always cheat. The choice of whether to cheat is a function of the probability of being caught, the duration of jail sentences, and the transmission characteristics of the primary. If the primary is always transmitting, jail as a time-out is not an effective deterrent against cheating because an honest cognitive user would not get to transmit anyway.	46
5.4	These lines describe how painful jail must be in order to deter cheating. If β is above these line corresponding to the other enforcement parameters, there is no incentive to cheat, regardless of the primary transmission pattern.	47
5.5	P_{FA} and P_{MD} against channel use when $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 3$. When the penalty is high enough, the cognitive user begins to avoid it by raising its own P_{FA}	51
5.6	Channel utilization if only the primary is present.	51
5.7	Overall utility (percentage of time either the primary or cognitive user is using the channel, without collision) and utility with full collision (if both primary and cognitive users are transmitting, not counted toward utility) in a low penalty regime with $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 0$	52
5.8	Effect of conservative P_{FA} on utility: $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 3$	53
5.9	Utilization with no cheating or extra false alarms $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 0.5$	54
5.10	Utilization when P_{MD} is fixed at 0.05. Overall utilization is not much better than with a purely incentivized scheme.	55
5.11	While keeping $P_{pen} = 0.6$ and $\beta = 0$ fixed, the boundary between always cheating and not cheating migrates with different values of P_{catch}	57
5.12	While keeping $P_{catch} = 0.4$ and $P_{pen} = 0.3$ fixed, the boundary between always cheating and not cheating migrates with different values of β	57

5.13	When P_{pen} is very low ($P_{pen} = 0$ here), with $P_{catch} = 0.8$ and $\beta = 1$, the boundary between cheating and not cheating is determined only by the chance of legally transmitting in the next time slot, p	58
5.14	When P_{pen} is very high ($P_{pen} = 0.99$ here), with $P_{catch} = 0.02$ and $\beta = 0$, the choice to cheat or not cheat depends on the long term average channel availability, so it is determined almost entirely by the steady state probability of the primary transmitting	58
5.15	With $P_{catch} = 0.8, P_{pen} = 0.6$, and $\beta = 3$, a better P_{FA} - P_{MD} curve means that the cognitive user does not have to decrease its P_{MD} (thereby increasing P_{FA}) much to avoid penalty.	60
5.16	With $P_{catch} = 0.8, P_{pen} = 0.6, \beta = 3$, and a better P_{FA} - P_{MD} curve, the overall utility is increased.	61
5.17	Depiction of the Markov chain for multiple bands. Here, each band has an independent primary user with an independent transmission pattern. The cognitive user may choose how to cheat on each band independently, but if it is caught cheating in any of the bands, it is sent to a Global Jail. While in jail, the cognitive user cannot transmit in its home band or in any of the other bands.	63
5.18	Requirement on the number of home bands to deter cheating regardless of primary transmission pattern when multiple bands are considered. Each new band introduces one more possible unit of utility, and thus more temptation. The required β is now B times that required for a single band.	63
5.19	The required β to deter cheating rises dramatically when $P_{wrong} \neq 0$. Of particular note is the behavior when P_{wrong} is very close to P_{catch} . If you will be sent to jail with the same probability anyway, you might as well cheat and get some utility for it.	65
5.20	The regulator sets P_{pen} such that for a given value of home band, β and a particular expansion B , there is no incentive to cheat.	66
5.21	As P_{pen} grows, so does the amount of time spent in jail for wrongful convictions. This influences the utility an honest cognitive user will gain from extra bands. Shown here is the cognitive user's utility as a function of the number of expansion bands, along with the percentage of time the user spends in jail. The utility peaks around the time the user spends 50% of its time in jail.	66
5.22	Assuming P_{pen} is as low as possible to satisfy the primary, the cognitive user can choose how many bands to expand into to maximize its overall utility. Notice that for expansion to be large, P_{wrong} must be very small.	66
5.23	We define overhead as the proportion of utility available that the cognitive user cannot gain due to spending time in jail. When plotted against the B which maximizes utility, we see that the cognitive user will never have an overhead greater than 50%.	67
5.24	The overhead incurred while varying P_{catch} . $P_{wrong} = 0.02$, and $P_{TX} = 0.55$ Notice that the maximum amount of expansion for any given amount of overhead is determined by P_{wrong} . P_{catch} can only make the possible expansion worse.	67
A.1	The Markov chain used for illustration.	78
A.2	The Markov chain actually used for simulations.	78

List of Tables

2.1	Timeline of radio regulation in the U.S.	5
3.1	Summary of differences between different regulation models	18
3.2	Summary of technical questions identified in this chapter	33

Acknowledgements

I would like to thank Professor Anant Sahai for his guidance and patience. Thanks to Professor Jean Walrand for his time and helpful comments. Thanks to my family and friends for their support. And thanks to all those who have read and commented on parts of this thesis at one point or another, especially Rahul Tandra, Pulkit Grover, and Mubaraq Mishra.

Finally, many thanks go to the National Science Foundation which has funded this research through a Graduate Research Fellowship.

Chapter 1

Introduction

How should spectrum be regulated? The current method of spectrum regulation, referred to as command-and-control, is a remnant of the first days of radio regulation. As time and technology progress, however, this model seems ever more outdated and inefficient. As evidence of the inefficiency, we note that currently the entire spectrum is allocated [1], but very little of the spectrum is actually used [2]–[4]. So, what is the problem, and why was such an inefficient solution adopted in the first place?

Command-and-control was designed to strike a particular balance between the conflicting objectives of efficiency of use and enforceability. When the model was first adopted, the fear was that interference would render spectrum unusable to anyone. So, the system was designed with large margins in frequency and in area to stop interference [5]. There was fear of spectrum not being used for socially desired purposes, so the government decided to control access in the interest of the community [5]. There was fear that chaos would occur with unregulated use, so the designations of proper use were made specific enough to allow a simple directional antenna to find misbehaving transmitters [5]. There was, finally, fear of the government losing the ability to leverage this important resource, and so large swaths of spectrum were reserved for government use [6]. The command-and-control model grew naturally out of the conditions at the time.

However, now that technology and understanding of policy has progressed, the balance attained by command-and-control seems undesirable. For the ability to enforce its will, the government trades away flexibility of use and encouragement of innovation. New ideas, such as privatization [5], [7] and commons [8], [9], have offered alternatives that may trade off freedom and enforcement in a less

costly way. The progression of technology offers possibilities for other balances still by providing the opportunity to use a hybrid of the two schemes to leverage good ideas from both [10]. However, while the privatization and commons schemes offer a new way to balance the objectives, it is not clear where that balance actually lies. With the added flexibility of these new regulation schemes, what will we have to do in order to enforce them? Will the required overhead be so much that we lose flexibility of use and spectrum efficiency anyway?

This thesis explores the overheads that will likely be required in order to support changes to regulation. This is a large, complex problem which cannot possibly be solved in this document alone. So, this thesis serves a dual purpose of identifying questions that will need to be answered to understand the regulation tradeoffs involved, and of taking a deeper look at one particularly important scenario and a proposed method of enforcement.

The body of this thesis is divided into four chapters. Chapter 2 explores the development of different policy strategies from a historical perspective. We try to understand the technical and political situations that gave rise to the original command-and-control model, as well as the new privatization, commons, and hybrid schemes. We also examine the arguments for and against each scheme to determine what are the important ideals we need to be concerned with when reforming spectrum regulation. The chapter closes with a discussion of how frequency-agile devices may change the situation, looking to the Internet as an example. We see that there is great potential for frequency-agile devices to revolutionize how we use spectrum. Furthermore, we have the opportunity to change regulation in a way that will support this technical advancement.

Chapter 3 takes a closer look at the privatization and commons regimes to try to understand the balance between our ideals of freedom and the realities of enforcement. We note that the two regimes rely on basic assumptions about the underlying medium and technology that may or may not be appropriate. We point out the technical questions that will need to be answered to understand the validity of these assumptions and what implications they have for how much freedom these new schemes will actually provide.

In Chapter 4 we engage in a more subtle discussion of how regulation and liability rules need to be designed in order to ensure both explicit and implicit protection of freedoms. For example, we may design regulation that allows complete freedom of spectrum use, but holds the manufacturers of devices liable for any disputes that may arise. In theory, this scheme promotes freedom. In practice, the manufacturers will severely limit what the end users are allowed to do in order to

protect themselves against litigation. In practice, there is no freedom of spectrum use. We use this to motivate a light-handed regulation for cognitive radio enforcement based on a “spectrum jail”: users who are caught not following sharing rules are denied use of the band for a specified amount of time.

Chapter 5 mathematically evaluates the effectiveness of a spectrum jail system to deter cheating. We do this by formulating the problem as a game, and investigating the effects of different enforcement parameters on the cognitive user’s decision to cheat. The thesis closes in Chapter 6 with an overview of the main themes and a discussion of where we go from here.

Chapter 2

Spectrum policy: historical formulation

Almost 20 years after Marconi demonstrated the first radio [11], federal regulation of radio spectrum was born in the form of the 1912 Federal Radio Act [12]. The spirit of this first form of regulation did not shift until 80 years later when the first radio auctions began. Now, almost 20 years after the introduction of auctions, the rights in spectrum are changing again with the approval of TV white space devices [13]. But where regulation should be headed in the future remains a heated debate. This chapter casts the debate in a historical setting to understand the technical and policy arguments that have directed radio regulation. We also investigate future regulation needs by looking to the Internet as an example of what is possible with common, standard components.¹

2.1 Command-and-control

Around the time of the 1912 Federal Radio Act, wireless communication was used primarily for Naval ship-to-ship and ship-to-shore communications. The 1912 Federal Radio Act was passed to protect this Naval communication from the interference caused by developing commercial radio systems. This Act established that any radio company could use the spectrum so long as it registered with the Secretary of Commerce, Herbert Hoover. Realizing that this free-for-all approach would

¹Although individual events will have appropriate citations, [5]–[7], [9], [14] were all very helpful in developing this story.

1895	Marconi demonstrates first radio
1909	First recorded listing of amateur radio stations
1912	Federal Radio Act gives regulation authority to Secretary of Commerce
1926	United States v. Zenith Radio Corporation et al. [15]
1927	Tribune Co. vs Oak Leaves Broadcasting Station court case [6]
1927	Federal Radio Act establishes the precursor to the FCC
1928	Philo Farnsworth demonstrates the first TV
1934	Federal Communications Act creates the FCC as we know it today
1947	Bell Labs designs first cellular system
1948	Claude Shannon introduces Information Theory [16]
1951	Leo Herzel proposes auctions for spectrum [17]
1959	Coase argues for market allocation of spectrum [5]
1969	De Vany et al define property rights for spectrum [7]
1976	543 subscribers on first cellular system in New York
1985	FCC Part 15 establishes ISM bands
1985	FCC allows spread spectrum
1993	Omnibus Budget Reconciliation Act establishes ability of FCC to auction spectrum [18]
1994	Gilder gives first critique against property rights [19]
1994	Jaap Haartsen and Sven Mattisson design Bluetooth specification
1995	First FCC auction (PCS bands)
1995	Noam gives first discussion of open access [8]
1997	FCC establishes U-NII bands [20]
1997	Balanced Budget Act expands auction authority of FCC
1999	802.11a, 802.11b released
2000	Mitola introduces Cognitive Radio [10]
2002	FCC authorizes UWB use [21]
2002	Wimedia alliance forms
2002	FCC issues Spectrum Task Force Report [2]
2006	White spaces coalition forms
2006	First chip sets approved for UWB use
2007	First Wimedia PHY devices certified
2008	FCC allows unlicensed use in TV white spaces
2009	transition to DTV

Table 2.1. Timeline of radio regulation in the U.S.

cause unnecessary disputes, Hoover parceled spectrum and allocated pieces to those who came to register.

In 1926, Zenith Radio [15] Corporation pursued litigation against Hoover because the law stated that Hoover could only register, not allocate. Zenith's success prompted an explosion of wireless use, with hundreds of radio stations coming online virtually overnight. In the wake of such active use, disputes arose over who should be allowed to transmit and cause interference. In 1927, the problem of interference went to the courts in the *Tribune Co. vs Oak Leaves Broadcasting Station* [6] case. To arbitrate this dispute, the courts began the process of assigning property-like rights to particular radio operators.

The government was quick to respond to this threat to its exclusive "ownership" of spectrum, with the 1927 Federal Radio Act [22]. This act legalized the process Hoover had begun in 1912. It allocated the same parcels Hoover had registered and established the precursor to the current Federal Communication Commission (FCC) to allocate remaining spectrum on the basis of "public interest, convenience, or necessity." This act was followed in 1934 by the Federal Communications Act [23] which created the FCC as we know it today. It also entrenched the "command-and-control" model for radio regulation.

In this model, the FCC retains the right to allocate spectrum in accordance with the "public interest, convenience, or necessity." The FCC interpreted this as a mandate to allocate spectrum based on comparative hearings, and subjectively evaluating the worth of any applicants. It allocated spectrum to those whose operations it deemed to be most in the public good. It also defined how the applicant may use the spectrum, including specifying the location, height, and transmit powers of broadcast towers.

2.2 The property-rights critique

2.2.1 Defining property rights

The command-and-control model met resistance in the 1950s, when Herzel [17] proposed markets as a better means of allocating spectrum. In 1959, Coase [5] wrote his landmark paper arguing for spectrum allocation through markets. De Vany et al, in 1969 [7], presented a comprehensive plan for property rights that is still used as the basic model today.

The argument against command-and-control is that the process is subjective, unfriendly to development, and reserved too much of the precious spectrum for government use. The criterion of “in the public interest” gives the FCC no objective standards by which to judge applicants; the comparative hearings even earned the nickname “beauty contests.” The subjectivity provides an opaqueness to the allocation process that discourages new entrants because there is a high risk of being denied spectrum for no foreseeable reason. The system is further debilitating to new technology because of the static allocations. An incumbent has no incentive to develop new technology because his current service is protected. A developer who wants to introduce a new technology faces the same risky bureaucracy as the new entrants, and therefore sees a barrier.

Finally, the command-and-control model allocated large swaths of the most useful spectrum to military and government use without thought or evaluation of how much that spectrum could be worth in the private sector. Property rights activists argue that spectrum is really no different than other scarce resources, so it should not be treated any differently by the government, and certainly should not be arbitrarily set aside for government use.

The primary problem with the old system was a lack of transparency and lack of evaluation of worth of government bands; markets offer a solution to these problems. If the method of receiving spectrum is a price, the barrier to entry is determined by market value, a transparent and available quantity. Likewise, the decision of who will get a piece of spectrum (when spectrum to allocate is relatively scarce) is determined by who values the spectrum the most. In other words, it goes to those who are willing to pay the most and therefore to those most capable of turning a profit. The property rights advocates assume that profitable uses are those that are most desired by the public, and therefore most in the public good. They consider the market mechanism as a fair and efficient way to distribute a scarce resource.

Note that the market mechanism is inefficient in that it cannot inherently limit monopolies. Monopolies pose a problem because they limit competition and therefore skew market price and create a large barrier to new entrants. The current command-and-control system can control monopolies by simply refusing to allocate too much spectrum to a single user. With markets, there is no immediate control on the ability to form monopolies. The assumption seems to be that if monopolies form, they can be broken up in the same way monopolies in other industries are: with anti-trust litigation. Again, we see privatization proponents arguing that spectrum should be treated no differently than other resources.

De Vany et al then provided the specifics of how the spectrum distribution would work: spectrum is divided into parcels according to time, geographic area, and frequency. The owner has the ability to use the spectrum however they wished within their parcel so long as they respected power limitations at the boundaries. This would give future entrants the ability to develop completely new kinds of networks with new equipment without having to go through the red tape of FCC approval.

After an initial distribution, owners have the right to participate in Coasian bargains and secondary markets. This means that they can negotiate with neighboring users to change the boundary conditions to mutually benefit both. They can also divide and sell their parcel of spectrum, so long as the new owners follow the boundary restrictions.

But how much do we need to give up to enforce this new flexibility? The rigid constraints in the command-and-control model acted as overhead paid for simplicity of enforcement: if you can regulate exactly how one is supposed to be using spectrum, it is very easy to determine who is not following the prescribed rules and tell them to stop. De Vany et al defined property rights with enforcement in mind: they defined boundary conditions under the assumption that they were the most easily enforceable restrictions that maintained an acceptable level of freedom. Disputes were left to the owners to prove harm, partly to encourage resolutions through Coasian bargains. Because the judicial system was adept at dealing with civil disputes, those that could not be solved with bargaining would go to normal courts to award appropriate damages. But are these enforcement provisions sufficient? We will return to this question in the next chapter.

Although a better understanding of the underlying physical environment has prompted criticism, [24], and legal scholars have weighed in on the realistic abilities of the courts, [25], the basic tenets of property rights in spectrum has changed little since Coase and de Vany. Matheson [26] gave a more precise model of spectrum division, including seven dimensions, adding transmitter height and two angles of arrival to de Vany's original four. But, the concept of a boundary constraint defined by power levels has not changed.

2.2.2 Partial adoption with auctions

In the 1980s, the number of applicants for cellular services began to change how the government allocated spectrum. Because of the sheer volume and similar service, the government began running a lottery for cellular allocations [27]. The Commerce Department estimated that the licenses given away in the 1980s were worth around \$46 billion [28]. The revenue potential of selling spectrum

could not be ignored and prompted the Omnibus Budget Reconciliation Act [18] of 1993 to give the FCC the right to auction spectrum.

The ability to auction spectrum prompted a flurry of papers [27], [29], [30] on how these auctions should be run in order to be effective. [27] gives an argument that in other countries, the auction design itself created unnecessary difficulties. For example, in Australia, a poor auction design allowed two companies to artificially raise the bid by bidding against themselves. When the auction ended, the companies merely defaulted on a series of bids, until they ended up paying far less for the license than it was worth. Other competitors had simply not participated because the price had been inflated so egregiously. The design of the auction itself is clearly very important. One of these early commentators in 2002 even suggested that in order to effectively allocate while allowing bidders to aggregate spectrum pieces and eliminate inefficient hold-outs, all of spectrum must be auctioned at the same time [31]. With a clean slate and correct property rights, spectrum regulation would proceed much more smoothly in the future. This idea is commonly referred to as the “big bang” theory of auction design.

A big bang auction is obviously not feasible, and the first auctions occurred in 1995 to allocate the PCS bands. However, these auctions were far from those envisioned by Coase. Although the spectrum swaths were now being sold by auction, remnants of command-and-control still remained: the auction was not simply for a piece of spectrum, it was for a piece of spectrum designated for cellular use. It also did not include provisions for Coasian bargains or secondary markets, so spectrum owners cannot change their allocation without explicit FCC approval. This prompted economists to submit a reiteration of Coase’s arguments in 2001 [32]. It is not yet clear whether the FCC has listened.

2.3 The commons critique

2.3.1 Arguments against the precursors: constitutionality and the Internet

While property rights ideas were filtering through the government, a radically different spectrum regulation concept was also growing. The stage was set in 1985, when the FCC established Part 15 rules to create the unlicensed ISM bands [33]: the same frequency bands that would later host WiFi. At the same time, the FCC permitted spread spectrum technology to be used [33].

Just before the first auctions took place, Gilder [19] presented the first rumblings of dissent from auctions. He claimed that auctioning off pieces of spectrum was akin to France selling server time with Minitels instead of opting for personal computers. The traditional commons critique was then introduced by Noam in 1995 [8] and crystallized by Benkler in 1998 [9]. Taking its inspiration from spread spectrum and the success of the Internet, the commons proponents argued that technology is advancing in such a way as to make devices capable of sharing spectrum resources. It is therefore questionable for the government to be involved at all in the process of distributing spectrum resources. When it is possible to have Free Speech by letting anyone transmit, government allocation is against the First Amendment, and is therefore unconstitutional [34].

The commons proponents agree with the privatization camp that the command-and-control model is reprehensible for its subjective beauty contests and wasting large amounts of spectrum for government use. However, a privatized scheme merely replaces the judges of the beauty contest with large corporations. The only ones capable of purchasing large swaths of spectrum are those with mass approval. This is obtained by angering the least number of people, not by offering the most desired services. Large companies simply cannot provide the level of innovation that a large number of new entrants can. New entrants come from the pool of end users, and therefore end users should have the ability to introduce new services. The success of the Internet gives evidence for this claim.

The original Internet was developed for military and academic use. When networked connectivity was first being provided for general consumers, the original model was for private companies like AOL or Prodigy to provide their own private bundles of networked services. These companies began to offer connectivity to the Internet as a small component of the bundle. However, the ability of a massive distributed set of individuals to provide services blew away the competition presented by the network providers. Consumers began to use their network providers as gateways to the Internet instead of for their other bundled services. Internet Service Providers grew out of these original network providers.

2.3.2 Defining the commons

Using the Internet as an example, commons proponents argue that the social value of end-user content is far greater than that of services which could be provided by large companies. So, the best use of spectrum would be to remove all gatekeepers and give end users the ability to access spectrum and transmit at will.

Notice here that there is a distinction between end users having the ability to create content, and having the ability to provide transport. On the one hand, there could be “open” spectrum in which companies are controlling the transport, but must carry anything the end users wish to say. This is essentially the model of the Internet. The argument of the commons proponents is slightly different. They argue that it is unclear whether the transport and content are coupled. To understand why this may be the case, consider multi-terminal information theory. Although it can be shown that for a point-to-point link the source and channel can be separated, for networked scenarios, the two must be coupled for optimal rate [35], [36]. In spectrum, it is not yet clear whether allowing freedom of content is sufficient for optimal innovation. Therefore, the commons proponents argue that the transport, too, should be available to end users so that services that are transport-dependent may also develop.

Therefore, instead of a market in spectrum, the market should be at the level of equipment. This would encourage technology that could share more efficiently as the equipment manufacturers will sell more devices if they are able to share the spectrum better. It would lower the barrier to entry and ensure a maximum level of flexibility as anyone who wants to transmit information may do so as long as their devices are certified to follow sharing rules. It also would extend the protections of the First Amendment to spectrum as anyone would be able to make a statement, not just large companies vying for mass approval.

With even more flexibility than a privatization scheme, the concern again focuses on enforcement. Everyone cannot possibly transmit whenever they please or a Tragedy of the Commons is likely to occur [37]. In this situation, no one who uses the resources has incentive to protect it and so it becomes unusable to everyone. The classic example is of over-fishing the ocean: fishermen get more money for more fish, so there is incentive to fish more. At some point, the resource cannot replenish itself fast enough, and there are no longer any fish to catch. So, the question becomes how much flexibility do we have to give up in order to keep the commons running smoothly?

The commons solution to enforcement and avoidance of a tragedy of the commons lies in regulating equipment: just as a party can be easily regulated by having all attendees lower their voices, spectrum can be regulated by imposing a transmit power constraint [38]. It is assumed that any further coordination will naturally occur because device manufacturers have an incentive to build equipment that can effectively share. Commons proponents point to the success of the Internet as evidence that large, effective networks can grow naturally from unlicensed use with few constraints.²

²Note that the Internet works because of TCP/IP. This common standard allows interoperability. In spectrum,

2.4 The hybrids critique

2.4.1 Defining hybrid allocation schemes

Despite a difference in ideological stands, both the property-rights critique and the commons critique agree that the command-and-control model presents an unnecessarily high barrier to entry with its subjective “beauty contest” approach to allocation. They also agree that the current practice of holding vast swaths of spectrum for military and public service use is wasteful: these frequencies are the most likely to be underused. Although property rights and commons camps each claim that their regime should encompass most if not all of spectrum, this is not actually feasible in the near future. They are both actually arguing that the government should release military bands to their use. Indeed, recent economic papers try to evaluate not which would be better overall, but whether a privatized regime or a commons regime would produce the most social value with the next available unit of free spectrum [39].

Even though the evaluation is often done in terms of social value, in reality, spectrum is allocated by politics; the two camps do not have equal amounts of political sway. The privatized scheme presents an appealing choice for policy makers because in return for releasing military bands, the government would be compensated by increased revenue. If new bands are allocated for commons use, however, the government gets no compensation for the loss of resources. Although a commons model may produce other kinds of social value, like better education through cheaper and better wireless Internet access, and although taxes on devices used in a commons might raise revenue in the long term, the argument is still not as politically appealing as trading spectrum bands for immediate revenue.

A much more politically feasible solution is a hybrid regime which provides an opportunity for the commons to gain spectrum without anyone else losing access to the band. The social value is therefore raised without loss of resource.

One such hybrid regime is ultra-wideband (UWB). Conceptually developed in the 1980s, UWB was rejected for military use in 1990 [40], but approved by the FCC as a commons underlay in 2002 [21].³ UWB spreads the communications signal over a wide range of frequencies, allowing the

the situation is likely different. Although the Internet shows that networks can evolve, in spectrum there would likely not be just one such network that had one common standard. Different services may require different kinds of architecture. The rules that would be required to govern a commons, then, are not necessarily the single common standard, but a set of rules that allow different standards to coexist peacefully.

³An underlay is a use of spectrum that can exist underneath regular spectrum use. Its presence is only observable as a mild increase in the noise floor.

signal to become part of the noise, only mildly raising the noise floor to a level that would not disrupt other communications in the band. So, any users of UWB could coexist with protected (licensed in this case) users without harmful interference. We will call these protected users “primary users.” Upon approval in 2002, the UWB Forum and WiMedia Alliance formed inside the IEEE to develop standards for UWB deployment [41].

UWB is operating in frequencies that have very little other use, so the primary interference is from other UWB devices. Therefore, the MAC layer is very important. Different ideas on how to design the MAC deadlocked the standards debate within the IEEE [41]. WiMedia eventually split off and independently developed standards for UWB use outside of the IEEE. It approved the first chipsets for use in 2006 [42], and the first WiMedia devices were certified in 2007 [42].

But UWB was intended for short distance, high rate services like wireless USB, not the kind of pervasive networks the commons proponents had originally envisioned. These kinds of networks have a more promising future in Cognitive Radio. Originally proposed in 2000 by Mitola [10], the idea is that smart radios would be able to react intelligently to their surroundings, finding and transmitting on the best available frequencies. This prompts three versions of what cognitive radio can mean: radios could participate in a real-time auction, choosing bands based on their current market price. But if a radio can choose frequencies based on price, it can just as easily choose frequencies based on sensing. Therefore, cognitive radios could also be used to share a common unlicensed frequency band by sensing and avoiding other users. They could, finally, operate in bands occupied by licensed systems. This would allow cognitive devices to fill in the holes left by licensed services [43], in particular in the military and public service bands where so much spectrum is currently wasted.

Again we question what we need to pay for the increased flexibility? Clearly with the licensed and cognitive users sharing the same band, there must be tradeoffs between how much spectrum can be recovered and how much protection can be offered to licensed users. Indeed, [44] gives a holistic view of the overhead required to implement a cognitive system that can protect licensed users. Overhead here is the amount of spectrum that is available that cannot be recovered because of the need to protect the licensed user. This is incurred by several operations: time and space margins must be maintained because of uncertainty in sensing the licensed user. This overhead can be reduced by cooperation [45], but cannot be entirely eliminated. We must also give up something in order to enforce sharing rules. We will return to this kind of overhead later.

2.4.2 The first steps toward adoption

In 2002, the Spectrum Policy Task Force [2] was formed within the FCC to evaluate the future direction of spectrum policy. It indicated that white space devices were a possibility. The television bands currently have licenses to occupy large swaths of spectrum, but few channels are actually used for transmission. This creates “white space” between the used channels where unused spectrum is available. As television stations are relatively static (it takes quite a bit of infrastructure and political bargaining to host a new station), they provide a good opportunity to test and develop cognitive ideas which could later be transitioned into other bands. In response to the possibility of white space devices, the Wireless Innovation Alliance, made up of industry leaders, was created in 2006 to encourage legally permitting cognitive devices [46].

On November 4, 2008, the FCC approved white space devices in the television bands [13]. The decision was to allow devices with a combination of GPS and sensing technology to operate in locations and bands that are not currently registered as occupied. The FCC also left the door open to devices that were capable of choosing transmit locations and frequencies based on sensing alone. This will require further FCC action as the FCC believes that these devices are not yet mature enough for commercial use.

But these hybrid technical solutions are currently incomplete. Although they provide techniques for regulating spectrum to fill current usage holes, they do not yet present a plan for certification or enforcement. The FCC has made clear that, for the time being, certification for white space must involve GPS and a television tower registry; methods of enforcement are still an open question. In order to migrate out of the television bands into the more fertile government frequencies, cognitive radio must transition to sensing alone. When this happens, even certification is unclear.

2.5 What the future holds: lessons from the Internet

Much of the policy debate rests on decidedly ideological grounds: the command-and-control model is inefficient and subjective but also easy to certify and enforce. Thus far, the inadequacies of the current model are annoying, but are not debilitating to technological development. Although the time required to get an allocation and technical approval is long, so is the time required to develop custom hardware for particular frequency bands. Consider UWB: although the FCC approved the technology in 2002, the development of technology, standards, and chipsets lasted from the 1980s to

2006. Time lags in adoption are also subject to network effects: the unlicensed band around 5 GHz is not well populated. It should be a prime target for use given that WiFi fills the 2.4 GHz band. However, 802.11a devices that operate at 5 GHz are more expensive than their 2.4 GHz counterparts and do not offer a significantly better service. So, the band is relatively unused.

The money required to develop custom hardware for, and the effort required to encourage adoption of a service are significant barriers to entry for any new entrants to the spectrum market. In fact, with the time scale of development being on the order of the time scale required to get approval, it is unclear which is the true bottleneck. Does development halt for the time required to get approval, or is it running in parallel?

The development of cheap, frequency-agile devices has the potential to change this situation dramatically. With the development of RF MEMS [47] comes the possibility of creating a standard hardware component that can be tuned in software. When this is true, custom components are no longer necessary, and development can occur at a much faster rate. What was once unusable spectrum for lack of proper development will be prime for use. Regulation clearly will be the bottleneck. The Internet can give us an idea of how much of a detriment slow regulation may be.

The Internet is a good example of what can happen when a common standard is introduced with few operational constraints. For example, putting together a new broadcast television station requires millions of dollars to get the approval, construct new towers, produce content, and run the station. Compare this to the success of YouTube, where the service was built very quickly by a few individuals without significant financial support. Content is virtually free beyond server storage and network connection because it is user-provided. Likewise, consider the infrastructure required to put together a cellular system, with towers, custom handsets, and a wired backbone. Compare that to Skype, where the necessary development costs are simply in software: the customers can use off-the-shelf computers, microphones, and webcams.

The beauty of the Internet is not just in the services that can migrate there and develop quickly at low cost. The true value of the Internet lies in those services that no one could predict. Social networking is a prime example of this: websites like Facebook or Myspace have revolutionized the way people meet and keep in touch and have no good analog in the real world. The common platform of the Internet allowed them to not only be developed by otherwise unknown individuals, it also allowed an extremely fast adoption of these services.

Social networking services on the Internet began in the late 1990s but did not immediately catch

on. They grew more popular with the introduction of Orkut and Friendster, and they achieved ubiquity with Facebook and Myspace. In 2008, millions of people rely on these services to help maintain contacts and inform friends. Compare this to the adoption of cell phones: the first cellular network in New York was created in the 1970s [11]. It was not until the turn of the century that cell phones became a ubiquitous and necessary part of modern life.

Once frequency agile radios enable common, off-the-shelf components for wireless use, spectrum holds the same potential as the Internet. The analogy is imperfect: the Internet provided a common transport for innovation in content to bloom. Innovation in spectrum will likely involve both transport and content, with different services requiring different protocols but commodity hardware. The Internet can also grow by simply adding fiber and servers while wireless spectrum remains fixed in extent. The difficulty in managing a limited resource with different, incompatible networks is much more difficult than a single, extensible, interoperable network. But, once development time and cost are no longer significant barriers to entry, regulation that inefficiently deals with this difficult problem will be a serious bottleneck. The reality of common components is not that far off, and regulation takes many years to change (auctions took almost 30 years to become a reality), so we need to start investigating the correct direction for regulation now.

If we do not create regulation that can deal with frequency-agile devices, the most probable course of action will be to ban them because there is no way to control them. This would not be a good idea. Challenging incumbent interest to introduce new services is already difficult [48]. If frequency-agile devices are banned, the incumbents are in an even stronger position to keep regulation as it is right now.⁴ Luckily, unlike with the Internet, we can predict how these frequency-agile devices might change the use of spectrum. Therefore, we have the opportunity to anticipate their effect and design regulation that can not only deal with them, but also encourage their use to realize their potential.

⁴Note that the incumbents likely prefer the current system. It caters well to their needs and considers them first in trying to introduce new services.

Chapter 3

Spectrum policy: underlying assumptions

The history of spectrum regulation shows that the command-and-control model grew naturally out of the situation at that time. However, as time and technology progress, this model grows increasingly outdated and more at risk of severely limiting technical advancement. In response, several schools of thought have offered alternate regulation systems. Despite their ideological differences, all of the reformers are actually trying to achieve the same goal. They are trying to find a solution that is flexible, encourages technical innovation, has a low barrier to entry, and effectively utilizes the available spectrum. At the same time, the solution ideally is easy to certify and enforce. The solution should also keep in mind general anti-trust ideals and First Amendment rights. The command-and-control, privatization, and commons models are compared in Table 3.1 with respect to the goals of spectrum regulation.

If the goals are the same, why are the methods of solving the regulation problem so different?

Area of Concern	Command and Control
Restrictions on use	FCC Approval
Encourages innovation	No
Barrier to entry	Political
Markets	None
Enforcement	Directional antennas
Dispute resolution by	FCC committee
Anti-trust ensured by	FCC
Who can speak?	Winners of “beauty contest”

Area of Concern	Privatization	Commons
Restrictions on use	Boundary constraints	Power limit
Encourages innovation	Protects investments	Low barrier to entry
Barrier to entry	Market price	Equipment price, congestion
Markets	Primary and Secondary in spectrum	Equipment
Enforcement	Owner proves harm	Victim proves harm
Dispute resolution by	Courts	Courts
Anti-trust ensured by	Litigation, auction design	Decentralized model
Who can speak?	Those with enough money	Anyone

Table 3.1. Summary of differences between different regulation models

Even Benkler admits that it may not be possible for the regulation debate to have a satisfying conclusion and the best option may be to run a 10-year experiment between private and commons schemes to see which more effectively meets the stated goals [9].

De Vries identifies a possible source of the problem in [49] where he points out that the metaphors we use to describe spectrum determine what we think regulation should look like. For example, the privatization critique relies on the analogy of spectrum as land. If we assume spectrum acts like land, and interference looks like conflicting uses of that land, then the only way to resolve the dispute is to parcel spectrum so that conflicting uses never have to share the same band. However, this analogy is imperfect: radio waves do not respect drawn boundaries in the same way houses or crops do. Interference also does not have to be avoided entirely in order for two users to share spectrum.

De Vries offers an alternate analogy, trademark law, in order to reframe the discussion so that solutions not otherwise obvious may arise. Trademark law originally evolved to keep people and companies from trying to impersonate each other in order to steal credibility and mislead customers. For example, it is illegal to make your coffee label look too much like that of Starbucks because customers may think you *are* Starbucks. Applied to spectrum, the trademark analogy brings interesting issues to the forefront: regulation would entail a registry in which users would claim rights to a particular signal. Others may then use any signal that does not look enough like yours to confuse receivers. This accurately captures the concept that interference is a byproduct of receiver design, not a quality inherent to the medium. We need to protect receivers, not space itself. Trademark also allows adoption of a “use it or lose it” rule so that you can only claim the rights to a particular signal if you are actually using it. This would help alleviate problems of spectrum being allocated but not used as anything unused would in fact be unallocated.

However, this analogy, too, has shortcomings: a registry-like registration process would perpetuate and exacerbate the command-and-control tendency toward incumbent interest at the expense

of new entrants. The number of possible marks to register in the real world and in spectrum is vastly different. Consider the wavelengths of marks in the two regimes: in the real world, the wavelengths of marks are those of light and matter. Extremely tiny wavelengths mean many can coexist without interference. However, radio waves have much longer wavelengths and so the number of uniquely identifiable radio waves is greatly reduced. This means that there are a limited number of users who can register a signal, and the first ones to register get to keep it. It also begs the question who gets to decide what will cause interference? If an early registrant has a receiver very susceptible to interference, should all future registrants have to respect this bad technology?

The point of the trademark analogy is not to offer a perfect solution; there are no perfect analogies. Instead, it is meant to help identify why the regulation debate seems to be deadlocked. The different sides have very different ideas about what spectrum actually is and how it should be treated. Although the analogies help in understanding an entity that we do not really know what to do with, they also hinder the discussion because they abstract from the technical characteristics of spectrum and wireless devices.

Both commons and property rights regimes have workable plans for how to allocate spectrum. Both have evidence that their system can successfully operate. However, there are several assumptions both sides tend to make with respect to the technical reality that effect their thoughts about what is possible to enforce. Our goal in this chapter is to identify the underlying assumptions related to the meaning of scarcity, the required technical infrastructure, the number of supportable players, and the amount of necessary government regulation. In explaining these assumptions we will point out the technical questions that need to be answered in order to determine what are the costs and feasibility of privatization and commons schemes. The chapter closes with a discussion of the implications of frequency-agile devices on the choice of future regulation and whether it is really a choice at all.

3.1 What is scarcity?

At the heart of the spectrum debate is the question “is spectrum scarce?” If spectrum is indeed scarce, then methods need to be developed to get rid of the least socially valuable uses. Command-and-control gets rid of them through administrative decisions. Privatization removes them through pricing. For example, consider cell phones: there is lots of congestion during the day, so prices are higher to exclude those who do not value their conversation enough to pay extra for it.

If spectrum is not scarce, however, there is no reason to restrict access as all uses can be accommodated. For this case, think of the Internet. The Internet can be expanded by adding more fiber and servers, so there is no reason to restrict access.¹ YouTube would likely have much less content if use were priced to weed out those videos with little social value.

In spectrum, scarcity means a lack of enough usable spectrum to accommodate everyone who wants to use it. There are four sources of spectrum scarcity: physical limits to the utility of the medium, the government gatekeeper, technical limitations, and network effects.

The physical limits are those developed in information theory as introduced by Claude Shannon [16]. Regardless of technical advancements, there are fundamental limits to the capacity of a wireless network.

The government as gatekeeper is the primary critique of both commons and privatization camps. If so much of spectrum is reserved for military and government use, there is not enough left to accommodate all the other services. This will hopefully change as regulation evolves.

Technical limitations create scarcity in terms of how much available spectrum is actually usable. For this kind of scarcity, the land analogy is appropriate: when the Midwest was developing, land was available for farming, but most of it was unusable until trees and prairies were cleared. Therefore, the unusable land was given to anyone who was willing to develop it, but the cultured farmland was private property that could exclude new arrivals. In spectrum, bands are usable if the chipsets exist to effectively communicate on them. Some of the cellular bands currently remain idle because the technology to use them has not been developed. The land analogy suggests that this kind of scarcity should be dealt with by allowing commons on unusable spectrum until it is developed. Exclusive rights would then be employed so that future entrants will be encouraged to invest in developing a new band. Indeed, this kind of “spectrum homesteading” was the original method of allocating spectrum. The 1912 Federal Radio Act gave allocations of spectrum to those who were already squatting there [6]. But, this is a decidedly technical question that will likely change as frequency agile devices become available: how many different services can actually coexist and what kinds of exclusion are really necessary?

The final type of scarcity is on the consumer end with network effects. If a service is not well

¹Note that this is an idealized example. In reality, it costs a lot of money to increase capacity, so access providers do try to limit less desirable content in times of congestion. Comcast recently noticed that most of their congestion came from P2P traffic. Therefore, during times of high use, they would drop the P2P packets. The FCC intervened to make them reduce congestion based on load, not on content. But, use has to be limited somehow since even the Internet is not of infinite extent.

used, like 802.11a, the devices will be more expensive, which will hurt adoption and perpetuate the cycle. Even if the technology for service in a band exists, if consumers are not willing to adopt the service, the band will be underutilized.

Deciding whether spectrum is scarce has serious implications for what regulation should look like in the future. In order to determine whether spectrum is scarce, we need to determine the bottlenecks and determine how these bottlenecks will evolve with technology and regulation.

3.2 Reliance on the infrastructure

The commons and property rights models reflect different ideas about how to best encourage innovation. However, both rely on questionable assumptions regarding what technology and the medium itself will allow. Privatization assumes exclusion will easily encourage innovation by protecting current investments. However, exclusion is actually only simple with a tightly controlled broadcast scheme like that in the command-and-control model. Likewise, commons proponents believe a decentralized scheme without a gatekeeper will best allow niche services and new innovation to occur. But, this kind of model relies on a pervasive infrastructure that is not likely to form on its own. It is not obvious that either of the models will be feasible in their current form and what overhead will need to be added in order to make them enforceable. These feasibility and overhead questions are inherently technical.

3.2.1 Broadcast

When Coase first proposed property rights, he adopted the concept that exclusion was an efficient and easy way to encourage innovation. De Vany et al proposed TAS (Time-Area-Space) packages with boundary power constraints with the idea that spectrum parceled in this way would be only marginally harder to enforce than command-and-control. Enforcement was assumed to require only a different definition of what was allowed and a porting to normal courts. However, enforcement was only simple in the command-and-control model because the government tightly constrained what the users were allowed to do. If a broadcast tower was causing too much interference, a directional antenna could tell you where the interference was coming from and a database could tell you who was causing it. The ease of enforcement is not inherent to an exclusion model. Indeed, movement

toward smaller, mobile devices begs the question whether the exclusion-based regime defined by de Vany et al is even possible.

Hatfield and Weiser [24] gave the first critique that this method of boundary definition may be difficult due to the nature of the physical medium. Radio waves do not propagate deterministically. Indeed, the propagation distance can depend greatly even on the time of day or the state of the ionosphere. Moreover, different frequencies vary in different ways: AM bands are much more variable than PCS bands because the larger wavelengths bounce more easily off terrain and the atmosphere. Therefore, boundary power levels will be very hard to predict and impossible to maintain. This can lead to users who either cannot prove harm from neighbors at all or users who act as “spectrum trolls” [24]. Like their patent troll analogs, these users do not productively use spectrum but rather sit idle, wait for their neighbors to overstep the boundary and threaten litigation if they do not pay a bribe.

Hatfield and Weiser suggest that these problems can be overcome by establishing stochastic boundary conditions and safe-harbor protection. Stochastic boundary conditions would say, for example, that the power observed at the boundary must be below a threshold a certain percentage of the time. Stochastic boundary conditions would mitigate some of the difficulty with uncertain propagation, but it would make proving innocence very difficult as an owner cannot be expected to constantly monitor the boundaries waiting for environmental disturbances. This is the same problem with deterministic conditions, except stochastic conditions allow simple safe harbors: if the offender’s certified models indicate that the boundary conditions will be met, the owner is protected against litigation. Weiser and Hatfield [50] take the argument even further to suggest that spectrum should be zoned so that appropriate boundaries and models are used to accommodate different spectrum characteristics.

Hatfield and Weiser capture well the problem that the physical medium presents, but they fail to address how the medium and progression of technology work together to make boundary definitions a nearly impossible problem. When transmitters are broadcast towers, fading and path loss with distance from the transmitter suggests a stochastic model of the boundary might be feasible. You can even track the tower with a directional antenna. Now, consider a network of smaller transmitters. These transmitters could beamform to produce a larger signal strength at a particular point. Even if they do not intentionally beamform, the signals will still constructively add at particular locations. A model for boundary conditions is now much more difficult because of the large number of variables; enforcement is nearly impossible as the location of culprits can no longer be determined from the

signal characteristics themselves. Now allow these small transmitters to be mobile, with consumer-determined patterns of motion. At this point, even modeling boundary conditions seems impossible because there are just too many variables changing in unpredictable ways.

Exclusion with de Vany's model is very difficult, and so we need to question what exclusion even means and what kinds of exclusion technology supports. This ties back to what is causing the scarcity that makes exclusion necessary. For example, consider when spectrum is scarce because technology has not yet been developed to use all of the bands. Then, whoever does develop the technology can protect exclusive use through patents, and not have to enforce their actual band at all. No one else will have the capability to use it. Unfortunately, sidestepping the problem is only possible in the short term. When more bands are available for use, if we want to preserve exclusion, we need to find technical answers to the problems with de Vany's model. In particular, how will we catch the small, mobile players? How should we design boundaries to be fair and to minimize disputes? Should they be defined by power levels, or do we have to impose something like an electric fence that causes user devices to shut down when they pass outside of their designated area? Finally, if we can appropriately define the rules and identify groups of offending devices, can this functionality be easily certified such that there is no incentive to cheat?

3.2.2 Is it all just the Internet?

On the other hand, the commons proponents adopt a massively decentralized scheme, assuming that market forces on equipment are sufficient to achieve efficient use of spectrum. The questions here are related to what is actually required to make this system run and whether there are incentives and ability for the equipment manufacturers to make it happen.

Commons proponents claim that a commons can be efficiently managed like one would manage a party [38]: if everyone keeps their voices (transmit power) low, everyone will be able to hear their own conversation. However, this also means that everyone can only hear their *local* conversation. In spectrum, this would require an abandonment of any services that relied on large distance broadcast, like radio or television. It would also require that if you needed long distance connectivity, the information would have to be routed through either a wired backbone or a pervasive mesh network. The latter seems to be the intended infrastructure, as arguments exist like Noam's [33] that in the proposed "abundant networks" [9], signals will be able to pay for their own way as they travel

through the ether. This requires not only routing, but even universal decodability and billing. It would require a single standard for all wireless use.

The intent is to eliminate the gatekeepers and the barriers to entry by allowing users to control both content and transport. The commons advocates like the freedom of the Internet. Free speech is protected by letting end users control the content. However, the fear is that if the networks themselves lie in the hands of a few small players these gatekeepers can limit freedom. This has recently come up with Comcast limiting P2P traffic during times of congestion. Although the FCC ruled that Comcast must stop, the fact that they tried gives credibility to the argument. If the transport is also in the hands of many small companies, it is much harder for freedom to be limited. Furthermore, to the extent that some content may rely on its method of transport, giving transport control to smaller individuals may also help innovation because different networks may arise serving different purposes and needs.

But is the plan even feasible? Innovation can occur at several levels: the consumer device manufacturers, the company device manufacturers, the service providers, and users who also presumably act as the content providers. The commons approach seeks innovation from the content and service providers but also requires innovation from the equipment manufacturers. Is there enough incentive to build the pervasive networks required for connectivity? That would depend on how hard it would be to produce the network.

Unfortunately, setting up and maintaining such networks seems very difficult. The local users would need signaling and routing protocols over a mobile network. Even UWB was deadlocked for years trying to develop a MAC protocol that was likely simpler than what the commons would require. In terms of enforcement, we encounter problems not just with loud, jamming nodes you have no provision to detect and stop, but also with greedy nodes trying to conserve their own power by not routing other people's signals.

The commons ubiquitous network faces serious technical problems. In order to determine its feasibility or deploy it, we will need to answer whether we can make sufficient universal routing rules. We must figure out how to deal with greedy or malfunctioning nodes. We finally must figure out what density of deployment is required for this mobile network to maintain reliable connectivity.

3.3 Number of players involved

We have already discussed some of the technical questions we will need to answer to implement either a commons or a privatization scheme because of the realities of the medium and the realities of the technology. But, what kind of constraints exist because of the realities of the players and the need for dispute resolution through the law?

Both commons and privatization are trying to keep in mind values beyond the technical efficiency of usage. They are trying to protect First Amendment rights, which require an ability for a large numbers of people to be able to speak. They are also protecting anti-trust ideals, which require that the system create opportunity for diversity and competition. The lowering of barriers to entry are partly to encourage technical innovation but also are aimed at these softer goals. In order for a scheme to support these goals, it must be able to handle a large number of players. This includes both service providers who direct the spectrum use and the end users who control use at least through their mobility if not through power control or channel choice.

However, a large number of players is not easily supported by current methods of coordination or dispute resolution. Indeed, when either of these involves a large number of people with different objectives, they are usually relegated to politics. Our goal in this section is to identify the reasons that these schemes actually require a small number of players for coordination and dispute resolution. We will then try to identify technical questions whose answers may help the number of players grow without requiring politics.

3.3.1 Coordination

The limit to the number of players shows up in both schemes through the need for coordination among parties. In a privatization scheme, this comes up when we allow the freedom of Coasian bargains. In a commons, as mentioned before, coordination is likely to be the activity on which the feasibility of the scheme depends. In both cases, successful coordination requires a small number of players.

Hazlett, in [51], argues that the solution to the difficulty in designing good borders lies in taking ideas from contract law. Contracts are often written to cover the situations most likely to occur. Other contingencies are often left undefined. The idea is that it is inefficient in both time and money to specify the details when there are other means of resolving the dispute if it happens

to arise. Later agreements between individuals, perhaps with the help of lawyers, will work out anything the contract happens to miss. In spectrum, this would mean that borders would be ill-defined but neighboring spectrum owners should work out a solution. So, Coasian-like bargains would be required by the very definition of the spectrum parcel.

Even if the boundaries are well-defined, the strength of the property rights scheme lies in the ability to strike Coasian bargains. But successfully completing a bargain requires a very small number of players. Even Coase [5] himself admits that with too many players, it would be more efficient for the government to regulate everything. This is because people tend to disagree when they are all trying to optimize conflicting interests. It may be possible to resolve the conflict between two parties at a border to find a mutually-agreeable solution. But with more, agreement is very difficult; unfortunately it is likely that there will be many more than two. If the properties are defined on small spaces, the propagation of signals will make any border agreement affect more neighbors than just the two that share the immediate border. The smaller the parcels, the greater the number of agreeing agents that are required. If we go all the way to real-time markets, the parcels are *very* small and therefore the number of players involved is very large. The privatization camp offers a possible solution: if border disputes are hard to resolve, then reduce the number of borders to simplify the process [51]. In other words, we should restrict the number of possible players.

In a commons, the need for coordination is even greater. Here, there are no borders that can be enlarged, so a bargain needs to be struck with all the parties who want to use the spectrum in that area to set protocols that everyone can agree on. In [38], Buck walks through a set of criteria required for a commons to work in the real world, and gives analogs that would be required with radios. Many of the requirements rely on the ability to create local rules and use social pressures to encourage players to follow them. But, when we require pervasive connectivity, locality is no longer a small geographic space. Effectively, we would need “local” rules that include everyone. This would greatly limit how many parties “everyone” could include. Note that WiFi is an example of a network that did evolve in the unlicensed band. But, the problem is not whether networks could evolve, or whether they would provide valuable service. The problem comes in trying to coordinate between different networks, providing different services, with interest only in coexistence and not interoperability. What kind of signaling and local rules would be required then?

3.3.2 Dispute resolution

Most scholars on both sides [11], [52] agree that when coordination breaks down and disputes arise, common courts should be employed to arbitrate them. Further, the dispute resolution should lie within civil law and look like either trespass or nuisance. Trespass occurs when someone interferes with a property owner's right to use their property. Applying this law is a natural outgrowth of the spectrum-as-land metaphor. Interference would correspond to a device interfering with another's ability to use the spectrum. Goodman [25] argues, however, that trespass may not be an appropriate choice because spectrum violations are not a pure harm. There is social value to the service being provided by the devices causing the interference. Therefore, interference maps more naturally to nuisance law which covers cases in which someone is interfering with an owner's right to enjoy their property. Unfortunately, nuisance law is messy and includes a judge weighing the social value of the conflicting objectives to determine whether injunctions should be issued or damages paid. In either case, as with most civil law, the purpose of litigation is to compensate the victim for any harm the offender may have caused.

Note that nuisance law looks much like what the FCC does under the command-and-control model, exchanging a judge for the FCC committee. In [50], [53], Weiser and Hatfield argue that this arbitration should remain with an entity like the FCC regardless of the future direction of regulation. Although courts are practiced at arbitrating civil disputes, spectrum presents a significant technical challenge. Because the courts have never dealt with spectrum disputes, there is no precedent to follow in making future decisions. Therefore judges and juries would be required to appreciate the relevant technical details and implications en route to a decision. When legal scholars and technical experts do not know how to handle spectrum, we should be careful about what precedent we allow the courts to set.

If we are left to the courts, we run into a problem with liability. When technology moves to smaller, mobile, fairly intelligent devices, who is liable for spectrum violations? Is it the spectrum owner for offering the service that interfered? Is it the manufacturer who made a product that could cause a dispute, or the end user for wandering too far from their intended service area or using the device improperly? Ideally, the liability would rest with the end user because if the providers or manufacturers are liable, they would engineer the system to reduce their risk of litigation. This means the flexibility and capability of the end user would be greatly reduced. If the point of adding

flexibility to the regulation is to allow development for greater social value to the end user, we should not design liability such that the manufacturer has to limit freedom to manage the risks.

However, there are lots of end users and the courts do not handle lots of small players well. Judicial law evolved handling disputes between two individuals, or two larger entities. When a person harms his neighbor, the litigation is common and relatively straightforward. When a large entity harms a lot of small agents, we use class action suits to make the small agents look like one large agent, and we revert to the original case. However, when lots of small players harm a large one or many small players harm many other small players, the courts no longer know what to do. Indeed, the only known way to arbitrate disputes between groups of people with conflicting interests is to invoke politics and diplomacy. Congress passes laws governing things like traffic rules so that individuals have a guideline for behavior so that individuals do not have to work out agreeable arrangements. Yet, traffic tickets are still often disputed. But guidelines for groups are not necessarily always available; they would likely not be for boundary disputes. Therefore, this would require consensus by many parties. If this were a simple process where mutually beneficial agreements could be made easily, our history would likely not be marked by so many wars.

3.3.3 How can technology help?

The real problem in coordination and dispute resolution is the need to coordinate large numbers of people. Machines have very little problem performing coordination tasks when they are programmed with a common standard. For example, consider TCP/IP. Billions of end users are trying to access different information. All of these diverse interests are being handled and coordinated in a distributed fashion without much need to resolve dispute. The most likely problem with routing on the Internet is congestion from too many packets. Therefore, TCP/IP was designed with a congestion etiquette. Although there are still bottlenecks, and other problems with the Internet, most traffic goes through without a problem. TCP puts some constraint on what can be done with the Internet, but most people agree that this constraint is reasonable in order to allow the Internet to run smoothly. Can technology help in a similar way with spectrum?

The key is the design of commodity products with a common standard. If we can design something that everyone uses, we can design the interaction to be such that large numbers of the devices can interact without dispute. Even if we cannot get all the way to a common device, there may be common ideals that we want to certify. For TCP/IP, it was congestion control. For cars, there

are traffic laws, license plates, and policemen. For spectrum, it may be particular enforcement mechanisms or signaling strategies so that we do not have to employ politics or the courts very often.

Therefore, the questions here are what would a commodity product in spectrum look like and what should the standards be? Also, if we cannot find a commodity solution, what common components should we include to limit disputes and encourage enforceability?

3.4 Government regulation

Like the feasible number of players, reality will require a closer look at the role of the government in either a privatized or commons regime. This is primarily a question of at what level regulation should occur. In order to effectively manage either regime, what must be enacted in law by Congress, what should be prescribed by the FCC, and what can be left to standards bodies like the IEEE?

Obviously, changing the form of regulation will require an act of Congress, the same way the auctions required Congressional approval before they could begin. But is there a further role for Congress to play? This is a question we will not delve into here, but it should be kept in mind that changes to regulation involve not only the FCC but also Congress.

In future regimes, what is the role of the FCC? Both privatization and commons proponents aim for a light-handed regulation with a simple certification process, leaving most of the decisions to industry or standards bodies because they will be able to more quickly adapt. However, the FCC may have a larger role to play than anticipated.

The classic argument for privatization includes the FCC at the level of defining original spectrum parcels, device certification, and record-keeping. A centralized body will be required to determine the original parcels of spectrum before secondary markets take place. Also, device decisions cannot be completely arbitrary, so although the extent of the certification is not yet clear, some certification will be necessary. The FCC will also be needed as a formal record keeper as without records of the bargains and secondary markets, privatization has the potential to be as opaque as the regime it is intended to replace. If the courts are deemed unable to handle spectrum disputes as Weiser and Hatfield [50] claim, the FCC will have a further role to play in arbitration. Although market driven, privatization will require significant FCC involvement to run smoothly.

The commons will likewise require quite a bit of government involvement. Indeed, Faulhaber [54]

claims that the level of regulation necessary to manage a commons would be so high that it would be as if command-and-control were still being employed. Because the commons would likely not be able to naturally induce either the social pressures necessary to enforce easily or the assumed mesh-network architecture, the need for external regulation has potential to be great. Hazlett [11] even argues that the government regulation would actually cause a second kind of tragedy of the commons in which the regulations are so strict that no one is able to use the resource effectively.

The questions we should be asking here include the correct tradeoff between law, FCC regulation, and standards bodies. Furthermore, if we manage to find technical solutions to the problems with large numbers of players, we will likely need government regulation to determine and enforce the standards. The case of UWB indicated that common, large-scale coordination solutions would likely not evolve on their own. We would therefore want to determine what we want the coordination system to look like and how to elicit this response in the simplest, least expensive way.

3.5 Does technology even allow us a choice

Until now we have assumed that the direction of regulation will be determined by conscious choice between commons and privatization schemes. The introduction of frequency-agile devices and cognitive radios has been treated as an opportunity for commons to coexist with either privatization or command-and-control regimes. But once frequency-agile devices exist, will we actually have a choice concerning what constraints regulation can impose?

Although frequency-agile devices may find their first home in a licensed or unlicensed band, the ability to jump frequencies gives these devices the ability to migrate. If these devices can move from a congested band to one that is underused and not able to bar their entry, they certainly will. As soon as these devices are available, we have the possibility of running into Faulhaber’s “hit and run radios” [55]. Although Faulhaber intended this term to be used for legal cognitive radios which turn on, cause considerable interference, and disappear without a trace,² the threat is even greater when the devices are not officially regulated because then there is *no* way to control their movement.

If they are so problematic, why should we allow frequency-agile devices at all? Banning these

²Note that radios are useful only if they can talk to others. The problem of hit and run radios is not that a jammer will turn on and blast for a few minutes, causing a lot of damage. The worry is that a mobile network, for example one that connects cars, will move past a relatively static network, like an Internet hot spot. The vehicular network may cause considerable interference for a short time as it moves past. It may be desirable for this level of interference to continue, but if it cannot be identified or stopped, the interference will continue along the travel path through the entire lifetime of the device.

devices would be a quick solution, but we would lose the potential for frequency-agile devices to open much more spectrum to utilization. We discussed before that one source of scarcity is the lack of chipsets that can work in many of the bands. If we can reconfigure the frequency tuning without changing the hardware, this limitation suddenly disappears. The second easy way to regulate these devices is to use the common components to open the band and then fix the frequency of the device. But, we do not just want to be able to configure for different frequencies, we want to be able to reconfigure the frequency in the field. Consider international travel: different countries allocate different bands of spectrum to particular companies or uses. There would be demand for a cell phone that could operate in any country the owner visits because it is capable of tuning to the correct country-specific frequency band. If we want this kind of flexibility, we cannot restrict the abilities of frequency-agile devices.

Far better than banning them, then, is to anticipate this coming change by answering the questions we have asked in this chapter. In a hybrid regime, all of the questions raised earlier still apply. Indeed, even more questions are raised. How should the devices be allowed to coexist? Although the FCC has already determined that white space devices can operate if they follow a GPS registry of known television tower locations, the trend will likely be toward devices that can sense their surroundings and act appropriately because this kind of interaction will allow the devices to share public service bands. In this case, how do we enforce sharing rules? How do we stop the “hit and run radios” or address another of Faulhaber’s concerns about devices of different powers respecting each others needs [55]? The regulatory choices ahead are difficult, and will require the joint effort of policy-makers and technology researchers to determine a good path for the future.

3.6 Summary of the critique

In this chapter, we have taken a closer look at the proposed replacements for the command-and-control regulation model. We see that both privatization and commons improve the flexibility of use. However, both face significant technical challenges. Because of these challenges, we cannot yet understand what flexibility we will have to give up in order to enforce. Wireless networks will likely evolve with different standards and different architectures to support different services. Ideally, we want to limit how much regulation we impose in order for these networks, which do not want to interoperate, to coexist. Producing commodity products and extensive standards may allow simple coexistence, but we must balance this against the flexibility of new services.

We have been able to identify technical questions that will help to understand the balance. These are summarized in Table 3.2. Finally, we looked at the opportunity technology presents in frequency-agile devices. Because we want to be able to leverage this technology, we will have to solve significant technical challenges. However, there is a large overlap between the questions we must answer if we adopt these devices and the questions we will inevitably have to answer in order to support a privatization or commons scheme. Therefore, the adoption of these devices seems inevitable: we will have to answer the questions anyway, and these devices bring with them potentially very significant benefit.

Assumption	Technical questions
Scarcity	<p>How many different services can coexist?</p> <p>What kind of exclusions are really necessary?</p> <p>What are the bottlenecks to efficient use and how do they evolve with technology and regulation?</p>
Infrastructure	<p>Property rights:</p> <p>What kinds of exclusion does technology support?</p> <p>How will we catch small mobile devices that do not respect boundaries?</p> <p>How should we define boundaries to be fair and minimize disputes?</p> <p>Can defined rules be easily certified and enforced?</p>
Number of players	<p>Commons:</p> <p>Can we make universal routing rules?</p> <p>How do we deal with greedy or malfunctioning nodes, and how much overhead does this require?</p> <p>What density of deployment is required for this mobile network to maintain reliable connectivity?</p> <p>What kinds of commodity products could we use to help grow the number of feasible players?</p> <p>What should the standards be?</p>
Government interaction	<p>If not commodity products, what common components could be employed to help with enforcement and dispute resolution?</p> <p>What is the correct balance between rules determined by Congress, the FCC, and standards bodies?</p> <p>How will we be able to regulate the coordination between players with the least possible overhead?</p>

Table 3.2. Summary of technical questions identified in this chapter

Chapter 4

Bridging the gap

The last chapters have gone through the spectrum policy debate and indicated that the proposed solutions rest on technical questions that require answers. The questions are mostly concerning enforcement: although the schemes will allow spectrum to be used, neither quite appreciates what overhead will be required to allow dispute resolution or coordination. In this chapter, we focus on enforcement as a technical question.

The chapter begins with a discussion of Lessig's warning in [56] that we must be careful in assigning liability and setting standards because the technical solutions evolving in the private sector are limiting constitutional rights without question or oversight. This has an analogy in the real world with wage slavery and debt bondage. Slavery is illegal. But people are still essentially enslaved by poor wages and debt. We port this warning to spectrum regulation to indicate that there are things in spectrum that we care to protect, such as flexibility of use and services that cannot share well with others. The method of protection is important; as is the related regulation and liability laws because they determine how much flexibility the manufacturers will pass on to the users. We close the chapter with a proposal that regulated cognitive radio will protect the flexibility we want. We then discuss how primary-secondary rules could be enforced to motivate the jail-based model explored mathematically in the next chapter.

4.1 Code as Law

In [56], Lessig describes a movement occurring with new technologies and the evolution of the Internet to move real-world laws directly into the code that defines how the new technology can be used. However, this movement is proceeding without the oversight of the checks and balances intentionally placed in our Constitutional system of government. So, it is proceeding without an eye toward the values and freedoms it is limiting or even removing completely. He issues a warning that without care, private companies will be able to restrict our freedom in ways that Congress cannot, without us even knowing it.

The ability to write law into code is most obvious in the virtual worlds developing on the Internet. What occupants of these worlds can do is quite literally determined by what the developers will let them do. If the developers want to write even the laws of nature differently, online players would be able to “fly” or “live underwater” indefinitely without oxygen. This could be very good and lead to an online experience that is vastly different from what we experience in normal life. However, the ability to write laws also presents an opportunity to limit your real world rights. For example, the developers could decide that doors are locked if there are more than ten players already in a room, thus eliminating your right to gather. It could also be made impossible to send messages to groups of players or post public bulletins, limiting your right to publicly criticize. It would be simple to even eliminate the anonymity that the Internet as a whole currently supports. This would limit your freedom to speak anonymously. These rules may even be made so discreetly that most online players never even realize that their rights are being limited.

While obvious and relatively benign in online virtual worlds, the ability to limit freedom by writing laws into code extends much farther into much more serious territory. Consider, for example, copyright for books. Books cannot be copied in general, but personal use and academic critique are exceptions. Before copy machines and printers, people would copy books by hand for personal use. With copiers came the ability to copy portions of books for personal use. Most legal electronic books, however, have restrictions that do not allow them to be printed. Although we still have the ability to copy the book by hand, our freedom to use our rights while taking advantage of new technology is limited. Effectively, limiting our ability to invoke rights on the new technology is raising barriers as it is much easier to simply no longer invoke the right at all.

The real problem is that rights are being limited by private manufacturers because limiting freedoms in code is an easy way to constrain illegal use and reduce their risk of being sued. However,

these restrictions are being placed without the oversight of our elected Congressional delegates, and without any obvious means of complaint through the courts. The question, then, becomes whether we can protect the rights and freedoms we care about while still allowing simple, technical solutions to enforcement problems?

4.2 Porting to spectrum policy

Although Lessig’s original warning concerns Constitutional values, it extends to how we think about spectrum policy. Spectrum agility is an analog to the copyright situation, although the freedoms we are trying to protect are not as obvious. With the introduction of frequency-agile devices, there is the ability for one device to be used in several bands. This presents an obvious problem in that devices can then migrate into other frequencies where they are not legal. It would be simple to ban frequency-agile devices, but then we would not have devices that could travel internationally or the ability to change spectrum regulation on a local level. It would be easy, too, to limit freedom by simply making the device manufacturers liable for the use of their products. The manufactures would respond by severely limiting what the devices are capable of doing to limit their risk, but without any oversight or question of the freedoms they are denying.

So what should we do? With the future of spectrum regulation as yet undecided, we have the opportunity to question what we value and design our spectrum usage rules and enforcement mechanisms accordingly. No method of enforcement is completely free—we will likely need to give up some flexibility in order to enforce any regulation scheme we choose, like the identity we give up with license plates in order to catch those driving unsafely. However, if we understand the implications of the regulation choices we make, we will know what we are giving up. Therefore, we can choose the right balance between flexibility and rules we can enforce.

Exclusion in property rights was created as an easy solution to the problem of interference, but the last chapter argued that the technical reality means that exclusion is actually very difficult. But is exclusion required to protect certain services? In [57], Etkin et al propose a game-theoretic model that explores the ability of spectrum devices to self-enforce. The result is that self-enforcement is possible, but only if the devices are capable of causing each other equal amounts of harm. Unfortunately, services are not homogeneous, and many services are not capable of causing other devices significant amounts of interference. For example, radio and television require broadcast to essentially dumb receivers; radio astronomy requires silence; and emergency services require unanticipated, im-

mediate access to spectrum. None of these services would be able to effectively coordinate with other users or provide threats of interference that could protect their use. If we value these services, we are left with an option like that proposed by Werbach in [52] where the rules of spectrum use are determined only by what you are not allowed to do (transmit in radio astronomy bands), instead of what you are (hold a piece of spectrum all to yourself). Note that these kinds of rules are in the same spirit as laws concerning crime: define what is illegal, not an exhaustive list of what is legal.

But what do we mean by “protected services”? The traditional property-rights model gives these protected services a swath of spectrum for their personal use. But do these services actually require their own separate bands, or can we simply define and enforce rules that allow these services to operate while sharing the band? Cognitive radio may provide a solution with the correct tradeoff.¹ The protected services get primary access to the band, but cognitive radios fill in the holes left at the interstices of the primaries. However, the technical requirements for protecting services are not simple [58]. So, is it possible to define rules that protect but do not overly constrain the freedom of agile devices? More importantly, can we design technical means of enforcement that are easy to certify and allow the flexibility we want to be retained?

4.3 Enforcement for cognitive radio

In cognitive radio, we face an interesting enforcement problem. This is because we are trying to make services coexist without them having to interoperate. So, we cannot simply offer utility without guarantee like the Internet because there are no interoperability incentives driving manufacturers to follow protocols. For example, radio astronomy requires a completely free band so that it can simply listen to the background noise. There is no incentive for others to cooperate with such a system. Therefore, the rules we are trying to enforce are not naturally occurring. We need the players that can cause harm to respect those they can hurt, even though they have no stake in doing so.

If we cannot rely on interoperability, how about reputation? Manufacturers rely on reputation to retain customers. However, in this case, reputation is not directly tied to quality of service. If a manufacturer’s reputation is that their products are poorly designed, they will lose customers. If however, the bad business practices of the manufacturer are causing the bad reputation, this reputation may not have enough of a detrimental effect to incentivize change. The example here is

¹Note that any definition of cognitive radio, whether it is a real-time market or primary-secondary model, will allow us to protect the desired services. The different definitions will also both require the same kind of mechanisms to enforce the rules.

sweatshop labor: even though customers generally know which companies use sweatshops, it does not greatly impact buying behavior because it is not central to the quality of the product. Therefore, we will need external regulations to impose any cognitive sharing rules.

In enforcing any kind of rules, there are two approaches: make the rules the only way to exist, or define more flexible enforcement strategies. The first is like Lessig's coded laws in virtual worlds: the code determines exactly what you are allowed to do. Unfortunately, outside the defined confines of a virtual world, natural uncertainty makes guarantees impossible. The second enforcement method is more reminiscent of the real world where incentive mechanisms like police and jails deter criminal activity. In this section, we discuss what guarantees and flexible enforcement mechanisms would mean in a cognitive radio setting.

4.3.1 Big Brother

If we know that our goal is to protect certain services, it seems that coding sharing rules directly into the device software would be a good way to guarantee protection. In fact, much of the current research on enforcement in cognitive radios has gone into trying to develop etiquettes that can be coded directly into devices. For example, [59] is a DARPA project defining policy-compliant codes. Software solutions provide a radio equivalent of an Orwellian society where the government watches every movement for anyone who is not following the law. However, the analogy also indicates how difficult it is to design guarantees because good standards are hard to design, certification is likely to be complicated if not impossible, and the process is not robust to mistakes.

The first difficulty is in defining correct standards. For wireless devices, someone, most likely the government, would have to design sharing rules. For example, a proposed standard for Cognitive Radio is the -116 rule [60], [61]. This rule says that cognitive devices should be able to detect the presence of primary signals of strength -116dBm. However, as indicated in [44], in the television bands, this rule will allow cognitive devices to use only 22% of the television band when 56% of it is actually available. Another standard likely to be required is a power limit. However, this power limit must be applicable to devices operating both in urban areas where high congestion would dictate small limits and in rural areas where power limits may be higher. A good standard is very difficult to define, especially if it must be applicable to different scenarios and be robust to technical advancement.

Once the standard is developed, devices must be certified to be compliant with it. This can be

done in two ways: lab testing and source code analysis. Unfortunately, to guarantee a particular functionality in wireless devices, both of these methods are very hard. The difficulty comes from the natural uncertainty of the wireless medium. Uncertainty, even about the characteristics of the noise, create fundamental limits on what a device can detect [62], and therefore limits on how well a device can protect a primary service. The uncertainty can be reduced, thereby making the detection more reliable, if we introduce cooperation. Indeed, it seems that for cognitive radios to adequately sense a protected user, cooperation will be required [45]. Unfortunately, cooperation makes certification very difficult. For one device, certification requires lab testing: a device simply needs to show that it can hear and respect a primary signal at a particular signal strength. But testing a network of cooperating devices because lab testing cannot possibly cover all of the situations likely to occur in the field.

If lab testing is not available, the only other option is to certify the software itself. Unfortunately, certifying code is even harder. Again, there are two ways to do this: test the code with an algorithm, or provide a proof. An algorithmic test is likely unreliable for the same reason a lab test of a network of devices is unreliable: determining the possible operating scenarios is hard, and simulating the effects in a comprehensive test that will not cause halting is even harder. This leaves only the possibility of requiring the code to carry a proof that it will respect sharing rules. Determining what such a proof would even look like is very hard and Gödel's incompleteness theorem [63] suggests such proofs may even be impossible.

Finally, for guaranteed functionality to be effective, the certification process must be robust to mistakes. Unfortunately, mistakes will always happen, and with this approach to enforcement, mistakes could be very costly. If a non-compliant unit makes it through the certification process or a unit malfunctions in the field, a coded solution will not be able to respond. The devices will operate, causing potentially serious amounts of harmful interference, perhaps for the entire lifetime of the device. If want to be able to cover the contingencies, some amount of policing will be required. Directly coding the desired response into cognitive radio devices will simply not be effective as a method of enforcement.

4.3.2 More flexible approaches

If policing will be required regardless to catch mistakes, perhaps a better solution is to push the enforcement predominantly onto a run-time scheme with minimal certification. This presents

the possibility for making certification very simple and entail only those elements required to make the run-time policing more effective. Notice that this seems very much akin to Hazlett's argument about border definitions [51]. You specify only what you need to, and the rest will either fall into place or be dealt with later.

But what would a policing solution look like? Policy experts often try to apply civil law to spectrum enforcement [25]. The idea in nuisance law is to weigh the social value of the conflicting actions, and force a party to pay damages to compensate for the harm they have caused. Unfortunately, as a technical solution, this seems very hard. Radios are not good at making subjective decisions, so they simply cannot weigh social values. Even if they could, and fairly assign monetary damages, such a scheme would require an extensive billing system that would incur significant unnecessary overhead.

Another possibility is to build a system based on deterrence instead of compensation. This would be the radio equivalent of criminal law and jails. The concept is that punitive damages are enough to incentivize the radios to follow the rules by making the legal path the most appealing option. Such a system would likely require some change to the current law and it is unclear whether it could sit inside civil law, but this question will not be discussed here. Of importance for our purposes is whether such a system could be built, what would it entail, and would it be effective?

The first question, then, is how to punish a radio? This is reminiscent of situations that occur in science fiction. For example, Asimov's robot novels [64] encounter the question of whether robots have rights and therefore whether they should be proved guilty and punished appropriately or just shut down upon suspicion of criminal behavior.

Even outside of fiction, legal scholars have begun work in determining whether a robot could be considered a person [65], and how to legally assign responsibility [66], but have not come to a conclusion on whether punishments are appropriate for robots. With other independent beings, like animals, the owner is liable but can assume the animal can learn that it should stop doing to offending activity. Robots, however, cannot feel pain, cannot learn from punishment, and cannot rationally decide to reform. This moots the point of a punitive action. But radios are not exactly robots—punitive action may actually have the appropriate effect by guiding the manufacturers to design radios that avoid punishment (ie, follow the rules). If our aim is for designers to design better radios, punishment does not even necessarily require a billing system like the compensatory approach. Punishment can be in kind, by interfering with the radio's ability to access the spectrum and therefore

interfering with the radio’s ability to deliver service. This would harm the manufacturer’s reputation for quality products, which would presumably affect the manufacturer’s ability to earn money.

One possibility exists in [67], [68], where Satapathy and Peha describe a system for deterring greed in a commons. It is based on a listen-before-talk scenario. It is important for radios to not be greedy in a listen-before-talk environment because if they all take up the radio spectrum for too long, then once they stop talking it will be very hard to regain use of the resource. To deter greed, the radios are required to wait before talking again for a length of time proportional to the amount of time they have just spent talking. Therefore, using the spectrum too much imposes a penalty. However, this system potentially has a large overhead because it does not change based on congestion: if there is no other traffic, a radio is still punished for being greedy.

Here, we propose a system that explicitly removes privileges the same way they are removed in the real world: by implementing a “spectrum jail.” This way, the punishment system can respond not to those activities that might not be allowed, but to those activities that actually cause harm. Also, this kind of approach leverages functionality that will likely already be required: malfunctioning nodes are potentially very harmful, and we cannot wait for a recall to discover them and shut them down. Therefore, a kill-switch is likely to be required [69]. If the functionality is already in place for a radio to respect a shut down command, little extra effort is required to make a radio respect a “go to jail” command which would shut the radio down for only a short amount of time.

Note that this scheme does involve the trade of some freedoms for the ability to enforce. Beyond the overhead incurred by having radios sit in jail, which we will discuss quantitatively in the next chapter, we require identity in order to catch offending radios. This means that users have to give up their right to transmit anonymously. As with license plates, we believe that this is a fairly small price to pay for enforceable sharing rules. In [70], [71], Atia et al describe a method of giving radios identity by implanting an identity code into the transmitted signal. They also discuss the overhead required for this system to allow radios to be identified.

But how well would a jail system actually work? Economics scholars have investigated the effectiveness of real-world jails as a deterrent for many years, beginning with the seminal work by Becker [72]. He creates a game-theoretic model that allows a rational criminal to weigh the value of committing the crime against the threat of punishment. He also models the social value of the crime and the social cost of the deterrence to determine the amount of policing and punishment that should be applied against criminal behavior. Along these same lines, in the next chapter we

will develop a simple game-theoretic model that captures the reactions of a cognitive system to a primary user and jail sentences. We would like to determine the amount of punitive action required to optimize social value by deterring cheating but not making the band unusable. We would also like to evaluate the overhead incurred when these parameters are optimized.

Note that such a system would be very easy to certify. Unlike complicated sensing, listening to a “go to jail” command is relatively simple. The only element that would require certification is a hardware switch and the code that activates it upon hearing a particular message. This kind of functionality could even be done with pre-made software because it would not affect the rest of the system and therefore would not limit flexibility.

4.4 Summary so far

When frequency-agile radios are available, we are faced with the problem of bringing spectrum usage to the masses. This is desirable for both free speech and to encourage innovation, but dealing with large numbers of independent players is very difficult. In fact, dispute resolution on this scale would normally *require* politics. Although commodity designs and standards may help to alleviate the courts’ need to resolve disputes, what should we actually standardize? If the intent is to allow flexibility in designing and operating the network, perhaps the only elements we should standardize are simple rules to allow coexistence in an easily enforced way.

But what does it mean for enforcement to be simple? We could enforce by limiting behavior, but this would be very difficult to certify given the uncertainty of the wireless medium. Therefore, it is likely to be simpler to design a scheme based on *deterrence*, which would make following coexistence rules in the best interest of all parties involved. Therefore, we need first a way of catching those who are misbehaving, which requires a system for determining identity. We then need a method of punishing the offending devices. If we take the traditional route of civil disputes and punish with fines, we run into two possibilities: we could require an extensive billing infrastructure that would allow individuals to pay fines. Otherwise, we could address fines to the manufacturers who would likely respond by severely limiting the abilities of end users to avoid having to pay. We take a different route here by having the punishment be in kind by degrading the radio’s quality of service when it does not follow the rules. This is the radio equivalent of a very painful jail.

Chapter 5

Deterrence perspective

In this chapter, we develop a simple game-theoretic model¹ to evaluate the possibility of using a spectrum jail to deter cheating by cognitive users. From the real world, we can infer that a jail system will likely be able to actually deter cheating. However, it is not clear how much we will have to pay for this functionality. Therefore, we develop this model in order to evaluate the overhead involved in such a scheme. The first kind of overhead is from the jail itself: a cognitive user cannot actually reclaim unused spectrum if it is sitting in jail and is not allowed to transmit.

But the model can be used to evaluate other kinds of overhead as well. We are trying to find correct enforcement parameters to deter cheating. This involves not only a jail sentence, but also the ability to catch the offending party and the ability to catch the *correct* offending party. By building these parameters into the model, we can get an idea of how good our policing must be, and how this relates to the required jail sentence. Intuitively, the jail sentence must be longer if our catching ability is worse in order to maintain the same level of threat. But, being able to catch an offender incurs its own overhead. Indeed, the more reliable and more accurate this system needs to be, the more overhead the scheme will incur. See, for example, the overheads required by the code-based identity scheme discussed in [70], [71]. If we can determine what these enforcement parameters must

¹In this game, we will model the primary as an independent, non-reactive player. The cognitive user will be reacting to the enforcement parameters and the transmission of the primary. The regulator will be setting the enforcement parameters. Although we formulate the utilities of the players and discuss how the parameters interact, we do not look at equilibrium for this game. The idea is to give some sense of the relative importance of the parameters, and the design goals for the regulation. However, as this is a very simple model, the equilibria may not have physical significance. Also, it is not clear what the regulator really wants to maximize—are they looking to maximize social value or to protect incumbent interests? So, we leave the utility function of the regulator undefined. We want to understand the interaction, not the outcome of this very simplified game.

be and which are of particular importance, we can get an idea of how much overhead will be required in policing.

In order to maintain simplicity while understanding the tradeoffs between different enforcement parameters, we have made modeling assumptions. The first set of assumptions concerns the primary and the interaction between the primary and cognitive users. As we have indicated earlier, there will likely be a need to protect certain services. How we decide which services to protect can be decided with any criteria we like, including maintaining legacy systems or marketing off different levels of protection in the package of property rights. We will call these protected services the primaries. We will assume that the sharing rule is that the cognitive user must not cause interference. Therefore, it must not transmit when the primary is transmitting.

We will further assume that the primary and cognitive users are sharing a spectrum band in time. This means that the cognitive user must wait until the primary shuts off before it is allowed to transmit, and then it must watch for the primary to return. This is a realistic assumption because sharing in time can be thought of as sharing in space with movement. For example, a cognitive user may be traveling in and out of the service area of a television tower. We will simplify the situation by further assuming that time is slotted and synchronized, so although a cognitive user may fail to detect correctly, the detection itself produces no lag.

Finally, we will make assumptions about the cognitive user that represent a simpler, worst case for the primary or the regulator who is actually performing the policing. First, we will assume that the primary is very sensitive to noise, so if it and the cognitive user are transmitting at the same time, the primary will experience interference. We will also assume that the cognitive user is not sensitive to noise, so if it transmits along with the primary, it will experience the same quality of channel use as when the primary is not also present. The cognitive user will also be assumed to be the worst kind of player: one with no concerns for delay and always something to transmit. This means that the cognitive user will try to maximize its expected amount of channel use, regardless of the patterns of delay with respect to its quality of service.

In the first part of this chapter, we will introduce a Markov-chain model to capture the interaction between a single primary and a single cognitive user in one band. We will use this simple case to understand how the enforcement parameters affect the cognitive user's choice to cheat. We will also look at the overall utility of the band, which will give an idea of how much of the unused time we are able to reclaim with cognitive use. The second part of this chapter will extend the model to a

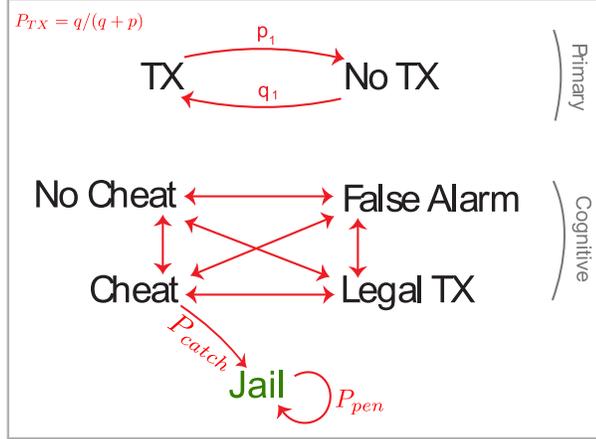


Figure 5.1. Markov model for enforcement in a single band: the primary user follows a transmission pattern characterized by the two-state chain at the top. In response to this usage, the cognitive user chooses between the actions in the bottom chain. When the cognitive user is cheating (i.e. using the band when the primary is also transmitting), it has a probability of being sent to jail. While in jail, the cognitive user is unable to use the band until a timer (determined by P_{pen}) runs out.

multiband scenario, where each band houses an independent primary user. There will continue to be a single cognitive player. We will use this case to explore the parameters further and understand how many bands a cognitive user will want to use to maximize its time transmitting. Finally, we will use this case to understand the overhead incurred by the jail system and the relative importance of enforcement parameters in deterring cheating.

5.1 Single Band

5.1.1 The model and basic regulatory conditions

In order to capture the conditions we have outlined above, the interaction between the primary and cognitive users is modeled as the set of Markov chains depicted in Fig. 5.1. Because time is slotted and synchronized, we can model the interactions with a discrete chain. The primary is modeled by the top chain, simply turning off and on with probabilities p and q respectively. We will also make use of the overall probability that the primary is transmitting $P_{TX} = q/(q+p)$.

In response to the movement of the primary, the cognitive user traverses the bottom chain in Fig 5.1: if the primary user is not active, the cognitive user may either be legally transmitting, or seeing a False Alarm, with corresponding probabilities of $1 - P_{FA}$ and P_{FA} . When the primary user is transmitting, the cognitive user can choose to cheat with probability P_{cheat} . It also has a

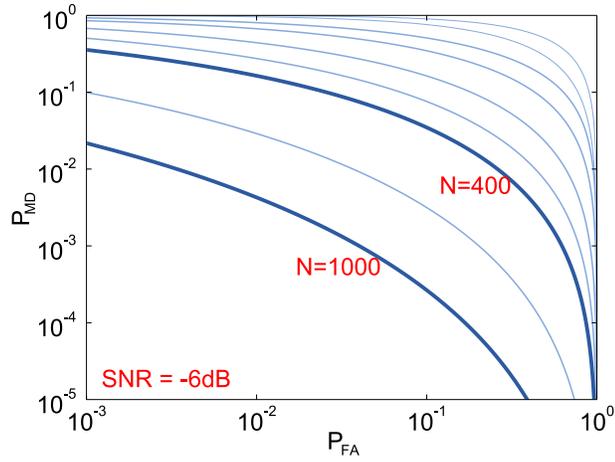


Figure 5.2. P_{FA} vs P_{MD} tradeoff curves for an energy detector. Bold lines are at $\text{SNR} = -6\text{dB}$ with number of samples $N = 400$ and $N = 1000$

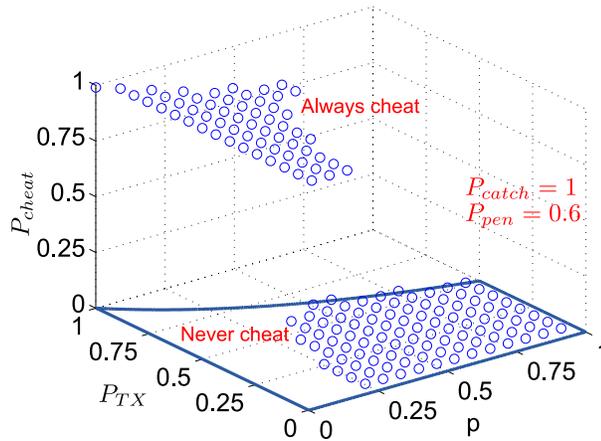


Figure 5.3. Typical cheating behavior for a cognitive user. If it is worthwhile to cheat, it is worthwhile to always cheat. The choice of whether to cheat is a function of the probability of being caught, the duration of jail sentences, and the transmission characteristics of the primary. If the primary is always transmitting, jail as a time-out is not an effective deterrent against cheating because an honest cognitive user would not get to transmit anyway.

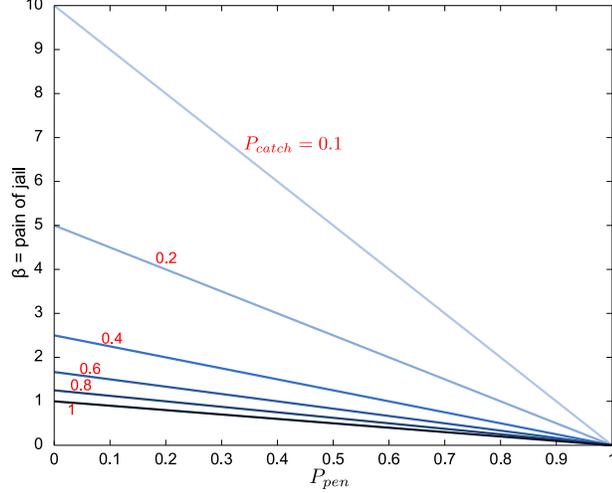


Figure 5.4. These lines describe how painful jail must be in order to deter cheating. If β is above these line corresponding to the other enforcement parameters, there is no incentive to cheat, regardless of the primary transmission pattern.

probability of missed detection P_{MD} , which affects its entrance into the cheating state. So, the cognitive user will cheat with an overall probability of $P_{MD} + (1 - P_{MD})P_{cheat}$. Note that P_{FA} and P_{MD} are not independent. Indeed, we can assume that they lie on a curve determined by the cognitive user's sensing mechanism. For all simulations here, the detector is assumed to be an energy detector which has the approximate test statistic [62]

$$T(Y)|H_0 \sim N\left(\sigma^2, \frac{1}{N}2\sigma^4\right) \quad (5.1)$$

$$T(Y)|H_1 \sim N\left(P + \sigma^2, \frac{1}{N}2(P + \sigma^2)^2\right) \quad (5.2)$$

where σ^2 is the variance of the noise, and P is the power of the primary. This produces the family of curves shown in Fig. 5.2 where the shape of the curve is determined by the energy detector itself and the slope of the curve is determined by the SNR and number of samples N . The middle bold curve is the nominal one used for most of the simulations. When a better curve is required for comparison, the other bold curve is used.

If it is in the cheating state, the cognitive user will go to jail on the next step with probability

P_{catch} ² (the probability of being caught). Once in jail, the cognitive user must wait there for an expected amount of time $1/(1 - P_{pen})$ before it is allowed to rejoin the game.³

The cognitive user adjusts its probability of cheating and its probability of missed detection to maximize its utility, defined as the average amount of time it is transmitting.⁴ So, its objective is:

$$\max_{P_{cheat}, P_{MD}} U_{cog} = \max_{P_{cheat}, P_{MD}} \pi_{legal} + \pi_{cheat}, \quad (5.3)$$

where U_{cog} is the utility gained; π_{legal} and π_{cheat} are the stationary probabilities of legally transmitting and cheating, respectively.

On the other hand, the primary (or its proxy the regulator), wants to minimize the time the cognitive user is cheating using its adjustable parameters, P_{catch} and P_{pen} . Note that the regulator must use these parameters in slightly different ways: P_{catch} is constrained by how simple it is to catch cognitive users and the deployment density of “policemen”. As conditions change during run-time, this parameter can be adjusted. The effectiveness of P_{pen} , however, relies on cognitive users respecting the jail-time. Therefore, it must be set beforehand with a particular P_{catch} in mind; devices must be certified to respect particular jail sentences. Note that, as opposed to sensing, jail time is a simple function to certify: you can simply put the device in a box and check whether it responds to messages or sends traffic after a “go to jail” command.

Fig. 5.3 shows the typical behavior of the cognitive user over different primary transmit characteristics when $P_{catch} = 1$ and $P_{pen} = 0.6$. When the primary is rarely active and switches quickly from transmitting to not transmitting (low $P_{TX} = q/(q + p)$ and high p), the cognitive user has no incentive to cheat. Therefore, $P_{cheat} = 0$. However, if the primary user is nearly always active or switches slowly, the cognitive user will be more tempted to cheat. In fact, if the primary user is always transmitting, the cognitive user will always be tempted regardless of the values of P_{pen} and

² P_{catch} captures two distinct effects. The first is the primary user’s imperfect catching mechanism, which will miss some cheating cognitive users. It also captures the case when the signal between primary and cognitive users is faded. So even though the cheater is caught, it may not hear the “go to jail” command and thus it could continue to transmit.

³In the real world, we generally think of jail sentences as being for a specific, deterministic duration. Here, time in jail is considered as a probabilistic quantity with the correct mean to simplify the analysis. Upon running empirical simulations, we note that deterministic and probabilistic sentences with the same mean produce the same effect on the probability of cheating. So in practice, the jail sentence can be deterministic.

⁴This utility function relies on the assumption above that the cognitive user is not hurt by primary transmission. This could happen in reality if the path from the primary transmitter to the cognitive user is badly shadowed, but the path from cognitive user to the primary receiver is not. So, the primary user can cause little damage to the cognitive user, but the cognitive user can cause major damage to the primary user. This is the case most worrisome to a regulator.

P_{catch} because it costs the same amount to sit in jail as it does to wait for the primary user to turn off.

Because the regulator must set the requisite P_{pen} at certification time, the chosen parameters must work for any p and q . So, the regulator should set them with respect to the worst case: when the primary user is always transmitting. As noted above, P_{catch} and P_{pen} are insufficient to deter cheating in this case because the cognitive user gains utility by simply bouncing in and out of jail. Therefore, we introduce a factor β which is the cost of sitting in jail. In the real world, β would correspond to an extra punishment, such as a fine or the emotional/physical hardship of being in jail. We will consider a physical meaning of β later. The cognitive user's objective is now

$$\max_{P_{cheat}, P_{MD}} U_{cog} = \max_{P_{cheat}, P_{MD}} \pi_{legal} + \pi_{cheat} - \beta \pi_{jail}. \quad (5.4)$$

where π_{jail} is the stationary probability of the cognitive user being in jail. Using the extra β factor, the primary can set P_{pen} and P_{cheat} to account for the worst case. If the primary is always transmitting, the cognitive user will always be in the left half of the bottom chain in Fig 5.1. Therefore, we can model this case with a simple two-state Markov chain with states "jail" and "not in jail". The cognitive user will go to jail with probability $P_{cheat}P_{catch}$ and will leave with probability $1 - P_{pen}$. The regulator should then set the enforcement parameters such that

$$P_{cheat}\pi_{not_jail} - \beta\pi_{jail} < 0, \quad (5.5)$$

or the utility lost by being in jail is greater than that gained by cheating for any value of P_{cheat} . This leads to a condition

$$\beta > \frac{1 - P_{pen}}{P_{catch}} \quad (5.6)$$

to dissuade cheating regardless of transmission pattern. Note that this condition does not actually depend on the detector employed by the cognitive user. The boundary for this condition is shown in Fig. 5.4. As long as the regulator sets β above the line corresponding to the other enforcement parameters, the cognitive user will have no incentive to cheat. Notice here that the β requirement can be arbitrarily low if we set P_{pen} closer to 1. This means that as the jail sentence gets longer, the utility gained by cheating cannot balance out the very small β paid for a very long amount of time. When $P_{pen} = 1$, the required $\beta = 0$ is intuitively pleasing because this penalty is a essentially

a kill switch. When one instance of cheating could bar a user from legally using the band forever, this user will have no incentive to cheat.

But how should the regulator actually set β , and is it bad to set β too high? Although the regulator now has a guiding equation, how should it balance the parameters? That depends on how do they interact to deter cognitive cheating. In order to answer these questions, we need to understand how the cognitive user responds to different parameters beyond simply having an incentive to cheat or not.

5.1.2 A deeper look into the effect of the enforcement parameters

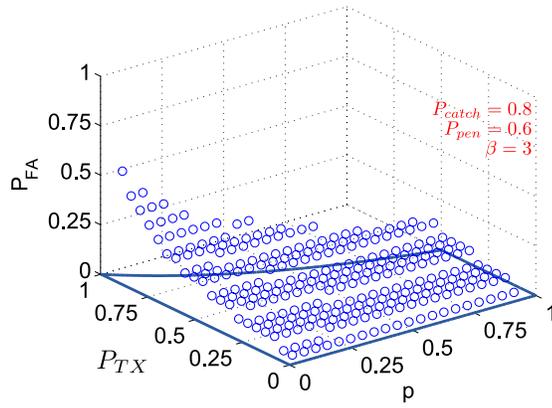
Overall channel utilization

We begin by trying to understand the effect of different parameters on the overall utilization of the channel. Ideally, the regulator will want the cognitive user to be able to recover as much of the white space as possible without causing too much disruption to the primary service. Although (5.6) will make it in the best interest of the cognitive user to not cheat, what happens if the penalty is raised? Can the penalty be too high?

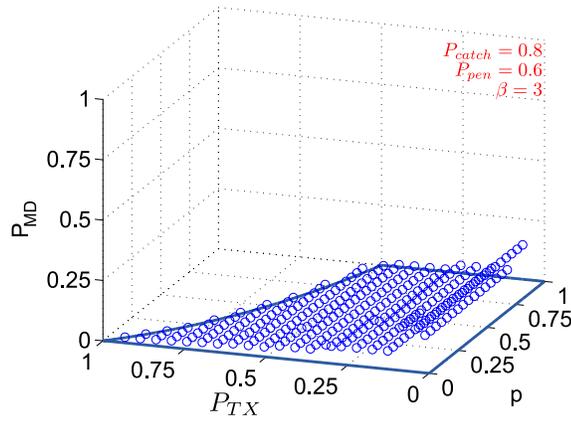
We have already seen that if the penalty is too low, the cognitive user will cheat with probability 0 or 1 depending on the primary transmission characteristics. If the penalty is too high, however, the cognitive user's reaction will look like that in Fig. 5.5: the cognitive user will begin to raise its P_{FA} in order to reduce its P_{MD} so it can avoid jail time. This reaction has a detrimental effect on the overall utilization of the channel.

All the cases that follow can be compared to Fig. 5.6, which shows the utilization when only the primary is present. It is equivalent to P_{TX} , the steady state probability of the primary transmitting. For all other utilization figures, two quantities are examined: The first is total channel utilization, defined as

$$U_{total} = \pi_{legal} + P_{TX} \quad (5.7)$$



(a) P_{FA}



(b) P_{MD}

Figure 5.5. P_{FA} and P_{MD} against channel use when $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 3$. When the penalty is high enough, the cognitive user begins to avoid it by raising its own P_{FA} .

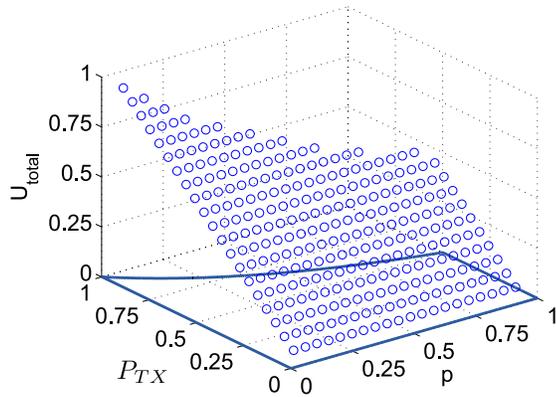
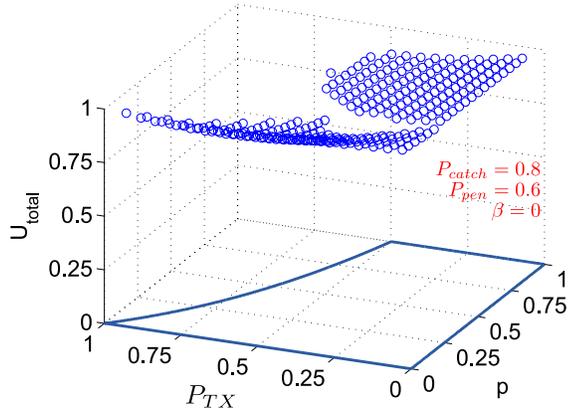
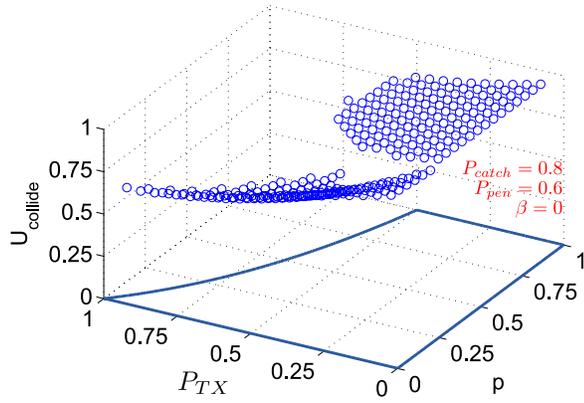


Figure 5.6. Channel utilization if only the primary is present.

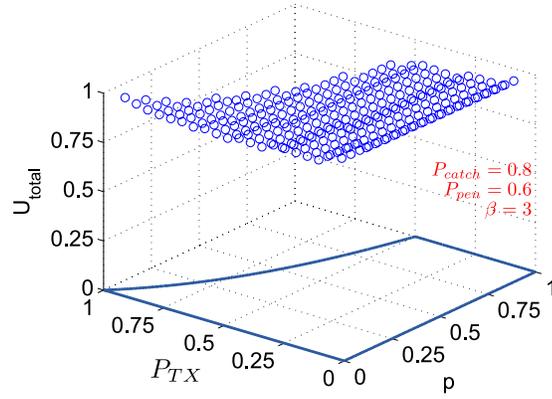


(a) Overall Utility

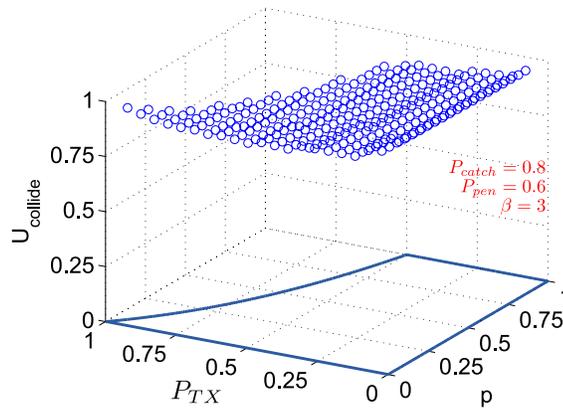


(b) Utility with full collisions

Figure 5.7. Overall utility (percentage of time either the primary or cognitive user is using the channel, without collision) and utility with full collision (if both primary and cognitive users are transmitting, not counted toward utility) in a low penalty regime with $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 0$.

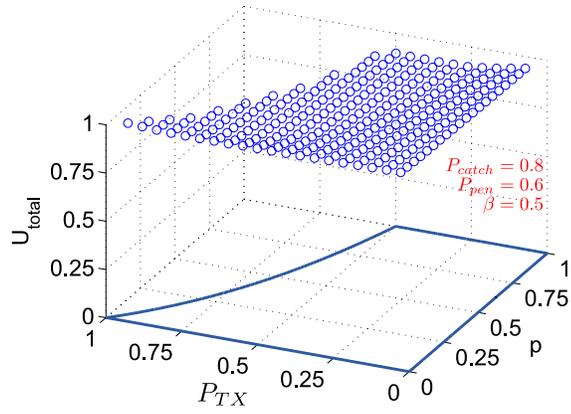


(a) Overall Utility

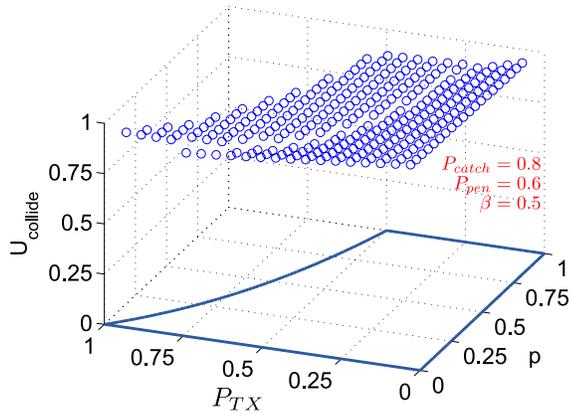


(b) Utility with full collisions

Figure 5.8. Effect of conservative P_{FA} on utility: $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 3$

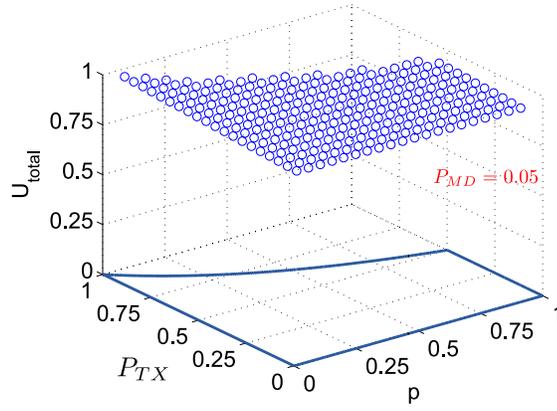


(a) Overall Utility

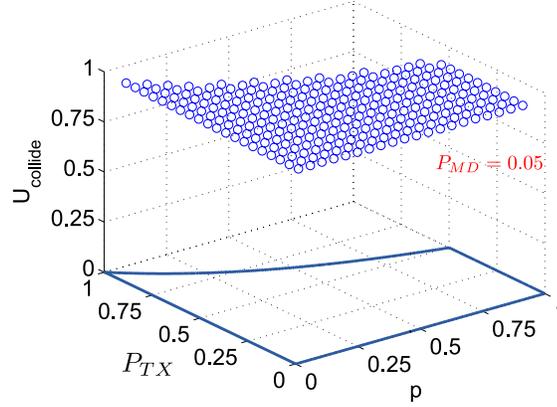


(b) Utility with full collision

Figure 5.9. Utilization with no cheating or extra false alarms $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 0.5$



(a) Overall Utility



(b) Utility with full collision

Figure 5.10. Utilization when P_{MD} is fixed at 0.05. Overall utilization is not much better than with a purely incentivized scheme.

which is the percentage of time anyone is using the channel. The second quantity, utility with collisions, is defined as

$$U_{collide} = \pi_{legal} + \pi_{no_cheat} \quad (5.8)$$

which does not count interference as utilization.

Fig. 5.7 shows the utilization for the cheating behavior observed in Fig. 5.3. Note in this figure that when the cognitive user is always cheating, the utilization follows more closely the curve in Fig. 5.6 because the cognitive user is spending significant amounts of time in jail. When the cognitive user is not cheating, it more effectively fills spectrum holes because it does not spend its legal transmit time in jail. The detriment of cheating to channel utilization is magnified when considering collisions; here when the primary is usually transmitting, the cognitive user is always cheating. The resulting collisions lower the channel utilization significantly.

Fig. 5.8 shows the overall utilization when the cognitive user is avoiding punishment, with P_{FA} as shown in Fig. 5.5. When the primary is always transmitting, it dominates utilization even when discounting collisions. When the primary is rarely talking, the utilization is good because the cognitive user has a low P_{FA} in this region. The middle, however, shows relatively poor utilization as the cognitive user is still actively avoiding getting caught even though the primary is not fully using the band.

Fig. 5.9 shows the utilization in between the last two cases: the cognitive user is not cheating, but it is not actively avoiding the penalty either. Notice that qualitatively this curve looks much like that in Fig. 5.8, but it gets better utilization in the middle, where before the cognitive user had an inflated P_{FA} . This is the sweet spot for regulation because it is the point of maximum utilization.

Now we fix the P_{MD} , as is done in current standards like the -116dBm rule [60], [61]. If $P_{MD} = 0.05$, the resulting channel utilization appears in Fig. 5.10. The channel utilization is approximately equal to our scheme. Therefore, little is lost by employing a light-handed, incentive-based regulation instead of one with tightly controlled standards. The difference then is in simplicity of enforcement. We noted in the last chapter that certifying guaranteed standards is nearly impossible; tweaking enforcement parameters, however, is much simpler.

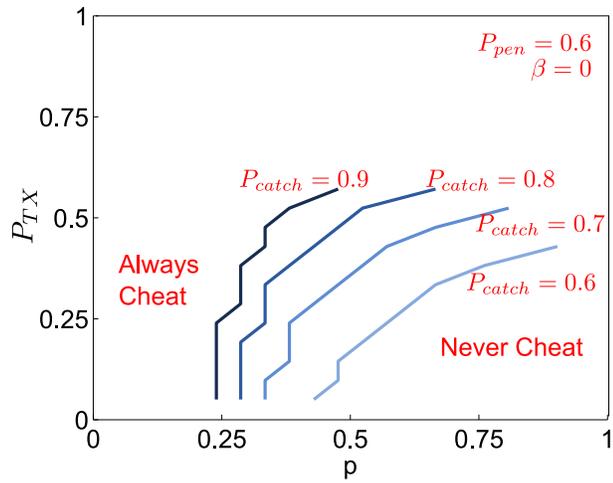


Figure 5.11. While keeping $P_{pen} = 0.6$ and $\beta = 0$ fixed, the boundary between always cheating and not cheating migrates with different values of P_{catch} .

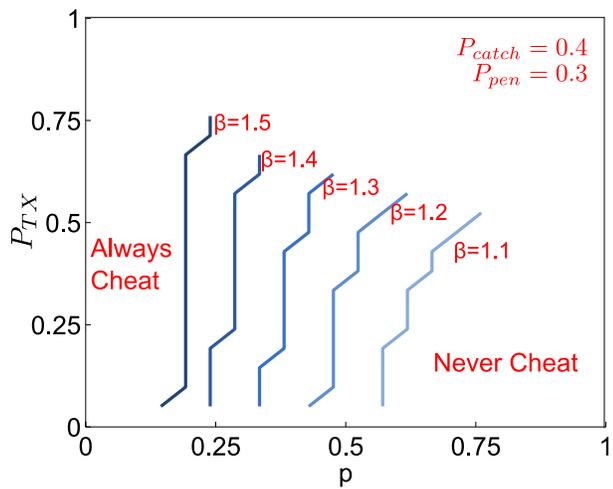


Figure 5.12. While keeping $P_{catch} = 0.4$ and $P_{pen} = 0.3$ fixed, the boundary between always cheating and not cheating migrates with different values of β .

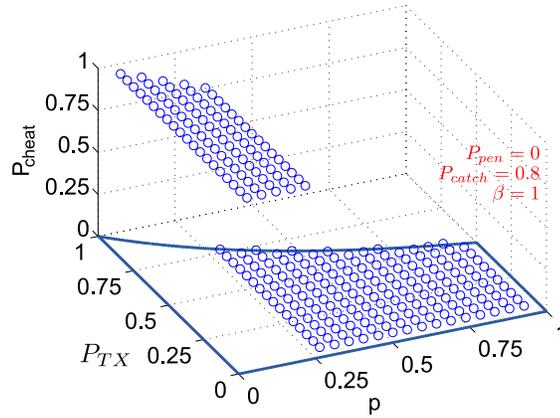


Figure 5.13. When P_{pen} is very low ($P_{pen} = 0$ here), with $P_{catch} = 0.8$ and $\beta = 1$, the boundary between cheating and not cheating is determined only by the chance of legally transmitting in the next time slot, p .

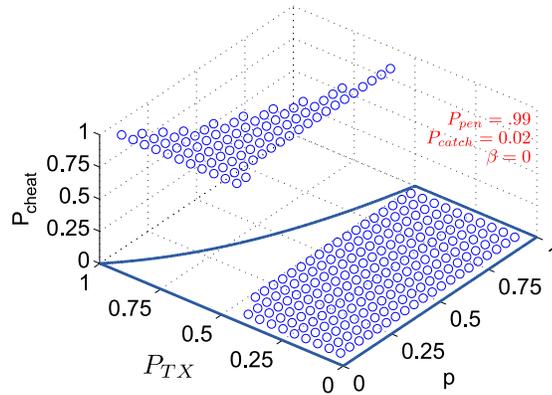


Figure 5.14. When P_{pen} is very high ($P_{pen} = 0.99$ here), with $P_{catch} = 0.02$ and $\beta = 0$, the choice to cheat or not cheat depends on the long term average channel availability, so it is determined almost entirely by the steady state probability of the primary transmitting

The effect of individual parameters

Since it is important to tune the parameters carefully, we need to get a sense of what effect each has on the cognitive user's cheating behavior. So, we will isolate the effects of different parameters in turn.

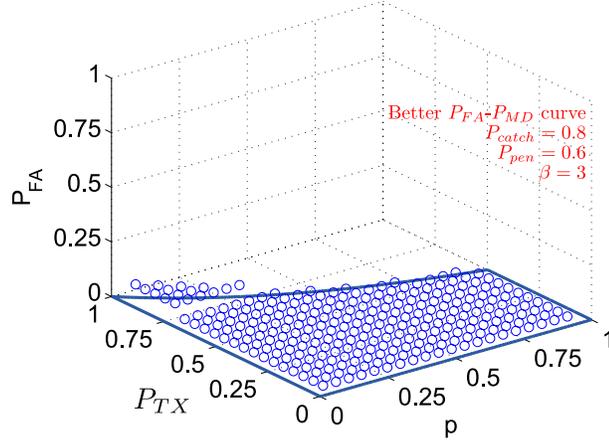
We can evaluate the effect of P_{catch} by tracing how the boundary between always and never cheating moves when P_{pen} and β are held fixed (Fig. 5.11). P_{catch} determines the location of the boundary with respect to the probability of the primary transmitting. β affects the probability of cheating in much the same way as P_{catch} ; the boundary between cheating and not is shown in Fig. 5.12 for a fixed value of P_{catch} and P_{pen} .

But how hard are these to change? P_{catch} is fairly difficult to improve because it would require either improving the method of identifying culprits or deploying more monitoring devices. The ease of changing β depends on what kind of extra punishment β actually is. If β is a fine, then once the billing system is in place, it is very easy to simply specify a different amount of money paid during a jail sentence. The caveat is β has units of cognitive utility, and the mapping from any kind of fine to utility may not be as clean as expected. We will return to this problem later.

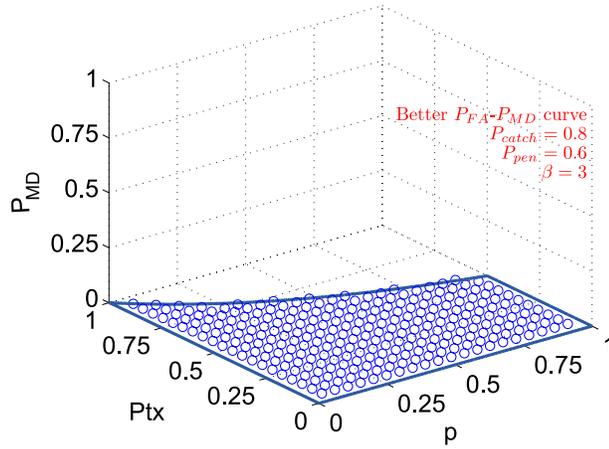
P_{pen} determines the length of stay in the penalty box, effectively determining how the cheating probability depends on P_{TX} and p . It is also the parameter that is easiest to set arbitrarily, but hardest to change. P_{pen} is just a jail sentence, so on the one hand, we can set this to anything without having to improve the infrastructure. However, it also needs to be set at certification time, so once a system is deployed, the jail sentence will likely be hard to change.

All examples thus far have P_{pen} at a moderate value. To get a better sense of its effect on behavior, consider the extremes. When $P_{pen} = 0$ as in Fig. 5.13, the cognitive user will stay in the penalty box for only one time step. So, whether it cheats is based on two things: the cost of cheating and whether the next time step will be an opportunity for legal transmission. The first is determined by P_{catch} and β as before, but the second is determined solely by the transition probability p and does not depend on long term channel usage. Therefore, the boundary between cheating and not in this case is a function only of p .

When P_{pen} is very high, as in Fig. 5.14, the cognitive user will have to spend a long time in the penalty box; the time to return to legal transmission therefore does not matter as much as a long



(a) P_{FA}



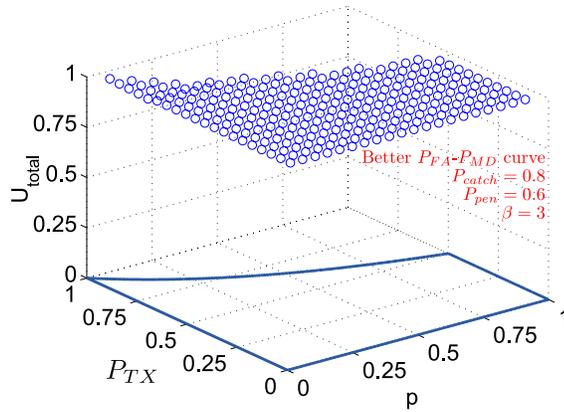
(b) P_{MD}

Figure 5.15. With $P_{catch} = 0.8$, $P_{pen} = 0.6$, and $\beta = 3$, a better P_{FA} - P_{MD} curve means that the cognitive user does not have to decrease its P_{MD} (thereby increasing P_{FA}) much to avoid penalty.

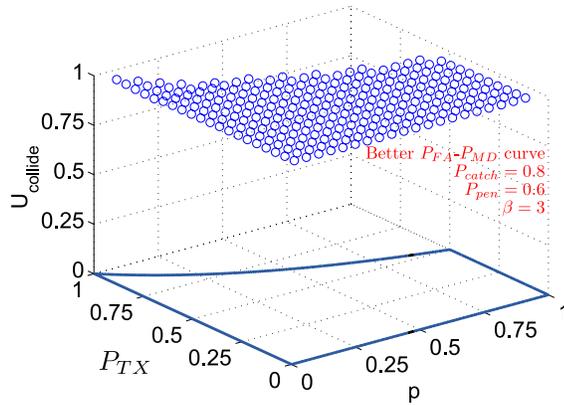
term average of channel availability. In this case, the dependence is almost entirely on whether the primary is transmitting and not on p .

Dependence on cognitive technology

As discussed in the last chapters, a good regulation scheme will encourage technical innovation. How well does a jail-based incentive mechanism do in this respect? We can model the effect of better cognitive detection technology in these simulations by simply replacing the P_{FA} - P_{MD} curve with a better one from Fig. 5.2. Under the same conditions that produced the levels of P_{FA} and P_{MD} in Fig. 5.5, the better tradeoff curve produces Fig. 5.15. Notice that while the original tradeoff curve requires the cognitive user to actively avoid the penalty, this one does not because the natural P_{MD}



(a) Total Utility



(b) Utility with full collisions

Figure 5.16. With $P_{catch} = 0.8$, $P_{pen} = 0.6$, $\beta = 3$, and a better P_{FA} - P_{MD} curve, the overall utility is increased.

is already low enough to not incur the penalty too often. With the improvement in P_{FA} even comes a corresponding improvement in overall utility, shown in Fig. 5.16.

Note that another tradeoff is suppressed here. The better P_{FA} - P_{MD} tradeoff can also be achieved by simply sensing the channel longer. However, the better curve achieved by sensing longer must be traded off against the opportunity to use the available channel and may not always be worthwhile. So, it is possible that the detector curve should be a function of the channel usage as well, which is not yet captured here.

It may be of interest to intentionally exceed the threshold defined by (5.6): better P_{FA} means better utility for the cognitive user, so by simply setting the penalty very high, we can encourage better cognitive-sensing technology. When the cognitive-sensing technology is perfect ($P_{MD} = P_{FA} = 0$), the cognitive user will never have to actively avoid punishment, and will get the best possible utility. Therefore, regulation does not have to set a particular standard P_{FA} and P_{MD} that all users must meet; it is possible to simply incentivize better technology.

5.1.3 The meaning of β

Throughout this section, we have assumed that it is possible to have an extra punishment β for sitting in jail. However, this punishment may be problematic: although it is relatively easy to implement and certify a “shut down” command, fines would require extra overhead in terms of a government-certified billing system. They also have other, less obvious effects such as enabling Hatfield’s “spectrum trolls” [24]. An alternative is to have β payable in kind so that it requires no additional infrastructure and does not encourage destructive behavior. Consequently, we propose considering cognitive radio as a bandwidth expander in which each user has a dedicated home band of value β and may expand into other bands by staking its home band against unlawful use. This interpretation lends itself naturally to expansion into several bands. Therefore, we now consider the multiband case.

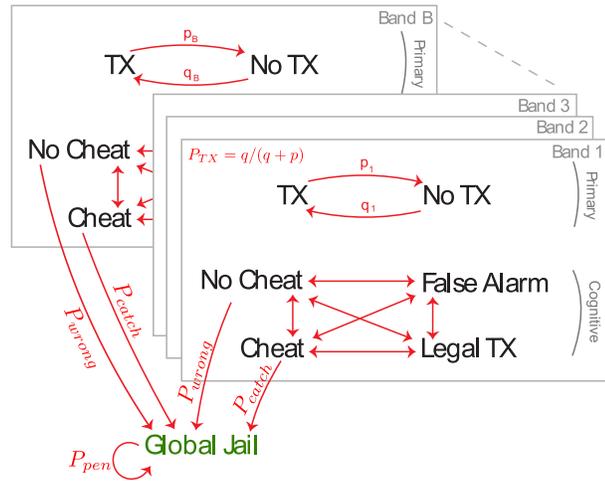


Figure 5.17. Depiction of the Markov chain for multiple bands. Here, each band has an independent primary user with an independent transmission pattern. The cognitive user may choose how to cheat on each band independently, but if it is caught cheating in any of the bands, it is sent to a Global Jail. While in jail, the cognitive user cannot transmit in its home band or in any of the other bands.

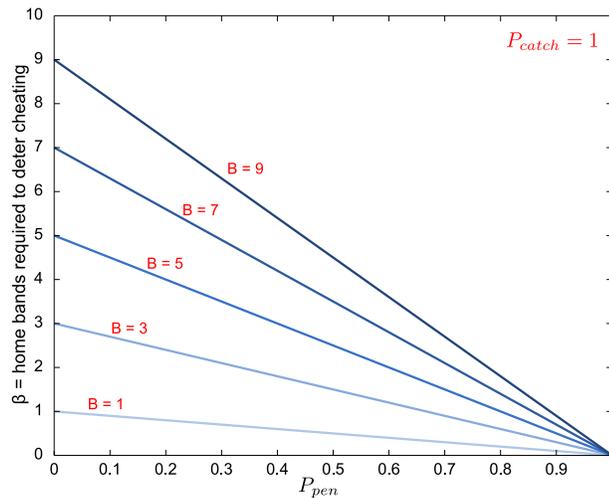


Figure 5.18. Requirement on the number of home bands to deter cheating regardless of primary transmission pattern when multiple bands are considered. Each new band introduces one more possible unit of utility, and thus more temptation. The required β is now B times that required for a single band.

5.2 Multiple Bands with Global Jail

5.2.1 Perfect Justice

The multiple-band case can be modeled with the Markov chain depicted in Fig. 5.17. Each band has an independent primary with separate characteristics. The cognitive user can choose whether to cheat separately in each band, but if it is caught cheating in even one band, it goes to a Global Jail, where it is not allowed to use either the expansion bands or its home band. Note that the Global Jail is a necessary consideration with multiple bands because if each band had a separate jail the multiple-band case would devolve into several single-band problems.

For now, we will continue the assumption implicitly adopted in the last section: the regulator has no uncertainty and so will catch the correct offending party with probability 1. In this multiband case, we will define P_{catch} as the probability that the cognitive user hears the “go to jail” command despite a faded wireless channel. This means that P_{catch} will remain constant regardless of the number of bands into which the cognitive user expands. Perfect detection of cheaters also implies that the probability of being wrongfully convicted (denoted P_{wrong} in the Markov chain) is zero. To simplify the analysis, we assume that the bands are identical: the primary transition probabilities p and q , although independent, are the same for all bands. We also assume that the cognitive user has a perfect detector (so $P_{FA} = P_{MD} = 0$).

In this case, we can derive a bound on the necessary number of home bands required to dissuade cheating by considering the worst case of all the primaries always transmitting. Solving for the β required for the cost of jail to be greater than the utility gained by cheating, we find the condition

$$\beta > B \frac{1 - P_{pen}}{P_{catch}} \quad (5.9)$$

where B is the number of bands the cognitive user is capable of expanding into. The boundary of this function is shown in Fig. 5.18. This equation is intuitively pleasing as the temptation to cheat should scale with the number of opportunities to cheat. This can alternatively be thought of as a condition on P_{pen} :

$$P_{pen} > 1 - \frac{\beta}{B} P_{catch} \quad (5.10)$$

because the number of home bands is presumably fixed before the enforcement parameters are set.

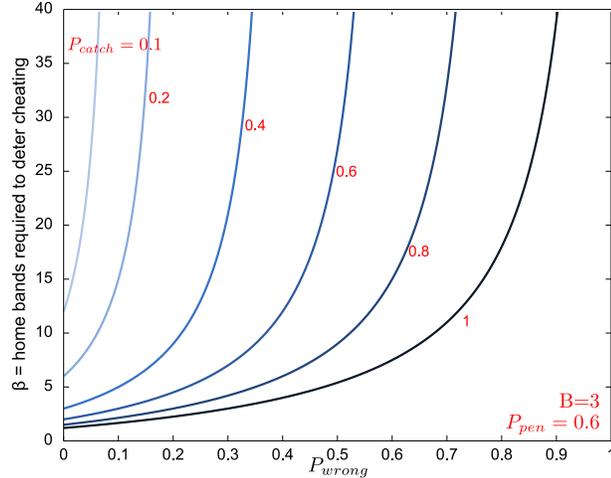


Figure 5.19. The required β to deter cheating rises dramatically when $P_{wrong} \neq 0$. Of particular note is the behavior when P_{wrong} is very close to P_{catch} . If you will be sent to jail with the same probability anyway, you might as well cheat and get some utility for it.

Note that P_{pen} is getting closer to 1 as the number of expansion bands grows. However, the honest cognitive user will never be sent to jail in this model, so the rising P_{pen} does not hamper further expansion in any way. In reality, there is a chance that innocent cognitive users are sent jail. We now consider the effect of this uncertainty.

5.2.2 Imperfect Justice

We model regulator uncertainty with the parameter P_{wrong} in Fig. 5.17. This parameter captures two effects: if the cognitive user misses detecting the primary and uses the band, it will get sent to jail even though it is not intentionally cheating. With this interpretation, the effective P_{wrong} could scale to 1 as the number of bands increases. P_{wrong} also captures uncertainty with the regulator if it mis-identifies the cheating user, or collectively punishes many cognitive users for one misbehaving node. In either of these cases, the cognitive user will be sent to jail innocently with some probability that does not scale with the number of bands. We have already assumed that the cognitive user has a perfect detector, so here we consider the case where P_{wrong} does not scale with the number of expansion bands. As before, the regulator wants to set the enforcement parameters so that there is no incentive to cheat even if the primary users are all transmitting all the time. Again, if the primaries are all always transmitting, the cognitive user is stuck on the left-hand side of the Markov chains in each band. Therefore, we can simplify the problem to a two-state Markov chain with states “jail” and “not jail”. The cognitive user goes to jail if it is caught, regardless of innocence, which

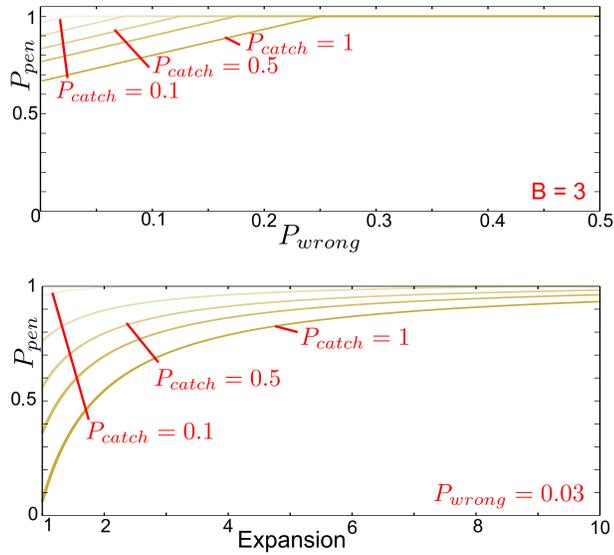


Figure 5.20. The regulator sets P_{pen} such that for a given value of home band, β and a particular expansion B , there is no incentive to cheat.

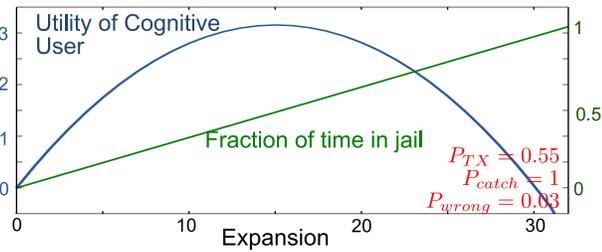


Figure 5.21. As P_{pen} grows, so does the amount of time spent in jail for wrongful convictions. This influences the utility an honest cognitive user will gain from extra bands. Shown here is the cognitive user's utility as a function of the number of expansion bands, along with the percentage of time the user spends in jail. The utility peaks around the time the user spends 50% of its time in jail.

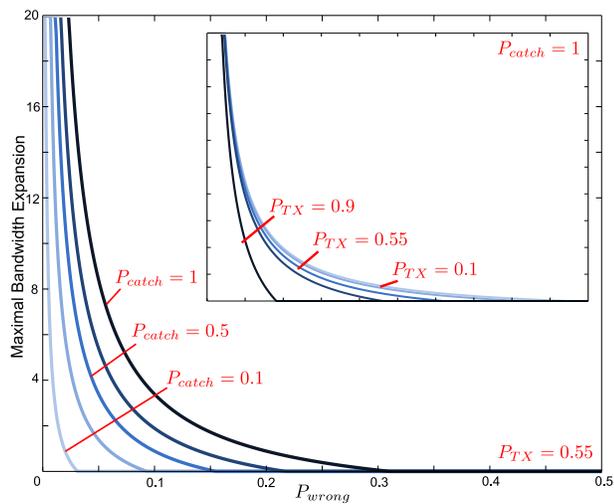


Figure 5.22. Assuming P_{pen} is as low as possible to satisfy the primary, the cognitive user can choose how many bands to expand into to maximize its overall utility. Notice that for expansion to be large, P_{wrong} must be very small.

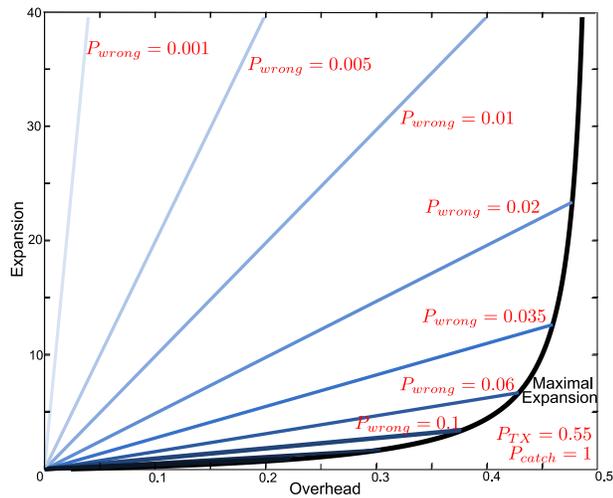


Figure 5.23. We define overhead as the proportion of utility available that the cognitive user cannot gain due to spending time in jail. When plotted against the B which maximizes utility, we see that the cognitive user will never have an overhead greater than 50%.

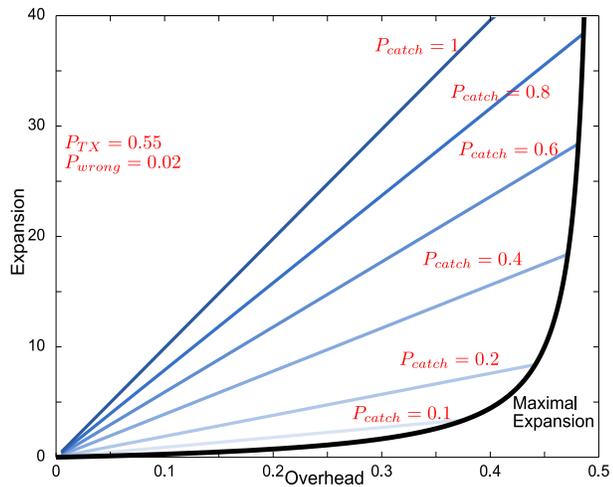


Figure 5.24. The overhead incurred while varying P_{catch} . $P_{wrong} = 0.02$, and $P_{TX} = 0.55$ Notice that the maximum amount of expansion for any given amount of overhead is determined by P_{wrong} . P_{catch} can only make the possible expansion worse.

happens with probability $P_{cheat}P_{catch} + (1 - P_{cheat})P_{wrong}$. The probability of leaving jail remains $1 - P_{pen}$. The regulator must set enforcement parameters so that

$$BP_{cheat}\pi_{not-jail} - \beta\pi_{jail} < 0, \quad (5.11)$$

or the utility gained by cheating is less than the utility gained by not. This produces the condition:

$$\beta > B \frac{1 - P_{pen} + P_{wrong}}{P_{catch} - P_{wrong}}. \quad (5.12)$$

The boundary of this condition is shown in Fig. 5.19 for $B = 3$, $P_{pen} = 0.6$, and different values of P_{catch} . The interesting thing to note here is that as P_{wrong} approaches P_{catch} , the necessary β goes to infinity. No amount of extra punishment can deter cheating if you will be sent to jail with the same probability anyway. Also notice that whereas before we could have β arbitrarily low if we increased P_{pen} , with wrongful conviction this is no longer the case. If the penalty acts as a kill-switch, P_{wrong} determines the amount of time a cognitive user will be able to act before it is taken offline. Cheating may shrink the amount of time it has to operate, but it may be able to get more utility by cheating with a shorter lifetime than being honest with a longer one.

As before, we should be considering this condition as the P_{pen} required for a given number of home and desired number of expansion bands:

$$P_{pen} > 1 + P_{wrong} - \frac{\beta}{B}(P_{catch} - P_{wrong}). \quad (5.13)$$

P_{pen} is again rising with B . With wrongful conviction, however, the cognitive user cares how high P_{pen} becomes because it will be sent to jail with some probability despite honest use.

Consider the following: for a given number of home bands, β , and a particular P_{catch} , the primary sets P_{pen} in (5.13) as low as possible as a function of the bandwidth expansion B and P_{wrong} .⁵ The result is shown in Fig. 5.20. As P_{wrong} grows, P_{pen} goes to 1. Likewise, for a given P_{wrong} , as the cognitive user chooses to expand into more bands, the P_{pen} required to assuage the primary goes up. Spending more time in jail affects the cognitive user's utility as shown in Fig. 5.21. The utility peaks around the point where the cognitive user spends half its time in jail, and then falls, becoming negative at some point.

⁵We saw before that if the cognitive user's detector was not perfect, the regulator would want to set the enforcement to deter cheating everywhere, but no more to maximize overall utility while assuaging the primary. Therefore, setting P_{pen} as low as possible is the best response for the regulator.

This utility function determines the game that the cognitive user is willing to play. If the cognitive user can choose only between not playing or playing with a particular number of expansion bands, the zero-crossing of the utility function is important. The cognitive user will play if the number of bands results in positive extra utility, and not otherwise. If the cognitive user can choose how many bands it will expand into, it will choose the number of bands that maximizes its utility. So, P_{wrong} dictates that even a greedy cognitive user will choose to expand only so far into other bands.

To make this utility maximization more precise, consider the cognitive user operating on a two-state Markov chain with states “jail” and “not in jail” with a probability of going to jail P_{jail} . The probability of leaving jail is $1 - P_{pen}$. The cognitive user wants to solve the following optimization problem:

$$\max_B \left[\frac{1 - P_{pen}}{P_{jail} + 1 - P_{pen}} U_{noTX} - \beta \frac{P_{jail}}{P_{jail} + 1 - P_{pen}} \right] \quad (5.14)$$

where P_{pen} is set to have equality in (5.13), and U_{noTX} is the utility gained when the primary is not transmitting.

Assume that at run-time, the cognitive user estimates the average probability that the primary users are transmitting ($P_{TX} = q/(q + p)$). It is then allowed to adjust its number of expansion bands, and corresponding P_{pen} , to maximize its utility. However, the time spent in jail does not depend just on the average probability of primary transmission; it depends on how quickly the primary users transition between transmitting and not. If the primary users switch very quickly, the cognitive user will spend more time in jail relative to slowly switching primary users. So, the worst case is when the primary users bounce back and forth at every time step. In maximizing its utility, the cognitive user should account for the worst case. However in real situations, primary users are more likely to stick in particular states rather than simply bouncing back and forth. Therefore, it is sufficient to calculate the maximum utility for the case when the primary usage is iid at each time step ($p + q = 1$).

With the iid primary usage assumption, $U_{noTX} = B(1 - P_{TX})$ and $P_{jail} = P_{wrong}(1 - (1 - P_{TX})^B)$. Since we are worried about when the primary user is mostly active, $(1 - (1 - P_{TX})^B) \approx 1$

for most cases of interest. So we will approximate P_{jail} as P_{wrong} . The optimal bandwidth expansion B/β is then

$$\frac{B}{\beta} = \frac{P_{catch}(1 - P_{TX}) + P_{wrong}(P_{TX} - 2)}{2P_{wrong}(1 - P_{TX})}. \quad (5.15)$$

This function is plotted in Fig. 5.22, varying P_{catch} in the larger plot and varying P_{TX} in the cut-out. Note that for large bandwidth expansions, P_{wrong} must be very small, and small changes produce large variations in the maximal expansion. Therefore, the cognitive user has an incentive to create the best detector possible to keep P_{wrong} as low as possible.

We would also like to get a sense of the overhead incurred by this punishment scheme. The overhead is the percent of available bandwidth that the cognitive user is not able to use because of jail time. Therefore it is defined as

$$Overhead = \frac{\beta + BP_{TX} - (U + \beta)}{\beta + BP_{TX}}, \quad (5.16)$$

where U is the utility being maximized in (5.14). The maximal expansion vs. overhead is plotted in Fig. 5.23. This figure gives a guide for the expansion possible for a given amount of allowed overhead and a particular P_{wrong} . Notice that the amount of overhead never exceeds 0.5 for desirable amounts of expansion. Also note the dependence on P_{wrong} : if we want to be able to have lots of expansion with a reasonable amount of overhead, P_{wrong} must be very small. Indeed, P_{wrong} determines how high the expansion can actually be. Fig. 5.24 shows the overhead varying P_{catch} for $P_{wrong} = 0.02$. Comparing this to Fig. 5.23, we see that P_{wrong} determines the maximum amount of expansion we can get for a particular amount of overhead. P_{catch} can degrade this value, but it is less important for P_{catch} to be large than it is for P_{wrong} to be very small. Therefore, we see that even radios follow the intuition from the real world: it is better to let many guilty men go free than to convict one innocent one.

Chapter 6

Concluding remarks

In this thesis, we have explored the interaction between technology and spectrum policy. We find that the current method of regulation, still dominated by command-and-control, strikes a particular balance between flexibility of use and enforceability. Although this made sense and was simple when radio regulation was first developed, we are no longer in the same situation. Wireless networks are no longer defined only by broadcast services, and interference is no longer something we need to avoid at all costs. This suggests that we might be able to strike a better balance in which we retain more flexibility while still being able to deter improper use. As history progressed, theorists began to realize that the balance of command-and-control was not even encouraging technical innovation or protecting the values of First Amendment rights or fair competition. In response to these realizations, two camps have emerged to propose new methods of regulation.

With new kinds of regulation comes new balances between what flexibility we have and what we can enforce. Unfortunately, although both privatization and commons camps have valuable insight into the problem, it is not clear where the balance would lie in either scheme. This is because key technical questions have not yet been answered concerning the infrastructure the schemes require and the number of players each scheme is actually able to support. Technology may be able to overcome any difficulties we face in implementing the schemes, but it is not clear how much will have to be regulated in order to keep the system running smoothly and keep the number of litigated disputes to a minimum.

However, we cannot just focus on the basic questions of how to get the system to run. Although we may be able to find technical solutions to some of the more tricky coordination or enforcement

problems, we have to be careful how these solutions are designed. This is because regardless of what we choose the enforcement to look like, we will have to give up something in return for the functionality. For example, we could just decide that the manufacturers and service providers can do whatever they want, but mandate that they are liable for whatever interference their users cause. The natural response of the manufacturers would be to severely limit the freedom of the users in order to make sure they do not get sued. Alternatively, we could design an enforcement mechanism that requires the ability to identify particular radios so that only the offending radios could be punished for any interference. The users must give up their right to transmit anonymously and may incur efficiency overhead, but in this scenario presumably the manufacturers would be able to grant the users far more freedom without fear of litigation. Both methods control interference; only one of these methods makes a reasonable trade of freedoms for enforceability.

Therefore, this thesis begins the process of understanding the tradeoffs involved in being able to enforce new schemes of regulation. We do this by building a simple game-theoretic model to capture some of the important effects of a particular scheme of enforcement relying on spectrum jails. We pick the scenario of a protected service and a cognitive user because this is realistically what we are likely to encounter in the future. Some services require more protection from interference, and if we value them while still desiring the flexibility of frequency-agile devices, we will have to impose some sort of primary-secondary sharing rules. We then consider the worst case of a cognitive user who is protected against interference from the primary, always has something to say, and is not delay-sensitive. If we can enforce sharing rules in this case, it is likely that a light-handed, incentive-based scheme will work for cognitive users with a more demanding quality of service constraint.

Our model indicates that there are simple tradeoffs between the enforcement parameters and the probability that a cognitive user will cheat. We also notice that the cognitive user has an incentive to improve their technology, so although the enforcement parameters do not have to be set to accommodate particular sensing technology, such an incentive scheme encourages cognitive innovation. We finally note that the intuitions gained from the real world transition well into a radio deterrence system: the single most important enforcement parameter is the chance of punishing an innocent cognitive radio. It is far better to lower the probability of catching the culprit than it is to catch the culprit reliably but take the innocent bystanders to jail too.

But this represents only the beginning of an investigation. This model relies of fairly substantial assumptions in order to maintain simplicity while still evoking important results. We note that the assumptions made have been in the direction of the worst possible case for any regulator. However,

these results should be extended to include different kinds of cognitive users with different required kinds of quality of service. If the cognitive user is very delay sensitive, the probability of being caught may not be as important as how long they would have to sit in jail before transmitting again. Likewise, if the cognitive user does not care about delay, but needs to maintain a consistent, long duration link, it may risk jail more often in order to not interrupt the service when the primary turns back on.

In arriving at our results, we also assume β can represent the utility of a clean band to stake against cognitive use in other bands. This means that the cognitive devices would be protected users in some bands, and secondary users in others. However, this is not the only model of interest. If the cognitive users have the choice between operating around protected users or in an unlicensed band, the situation would become more difficult because the cognitive users no longer have a clean band to stake. Evaluating necessary punishments in this case would likely depend on the level of congestion in the unlicensed band. If the unlicensed band is very congested, the cognitive user would not be harmed much if forced to give it up. Therefore, the punishments will likely have to be much higher with congestion.

Taking a broader view, a jail system is not the only way to deter cheating. There are other methods of punishment in the real world, like socially ostracizing or corporal punishment. These may have radio analogs that are effective with reasonable amounts of overhead. Investigation would be required to determine if other methods present a better tradeoff between ease of enforcement, flexibility of use, and overhead.

Broader still, the questions raised in this thesis are not limited to technical enforcement problems. There are many questions to answer to determine feasibility of the proposed changes to regulation and whether they will be as open to freedom and innovation as we would like. We must also answer many questions on the policy and law sides to determine what we can delegate to technology and what the law must accommodate. On the policy side, we need to be concerned with the difficulty of regulation and how incumbent interests and government politics play into this scenario. If we are to determine the most effective direction for future regulation, it will require the joint effort of experts in technology, policy, economics, and law. No one field can possibly solve the regulation problem alone.

References

- [1] NTIA, "U.S. Frequency Allocations," available online: <http://www.ntia.doc.gov/osmhome/allochrt.pdf>.
- [2] "Spectrum policy task force report," Federal Communications Commission, Tech. Rep. 02-135, Nov. 2002. [Online]. Available: http://hraunfoss.fcc.gov/edocs_public/attachmatch/DOC-228542A1.pdf
- [3] M. A. McHenry and K. Steadman, "Spectrum occupancy measurements, location 1 of 6: Riverbend park, Great Falls, Virginia," Shared Spectrum Company, Tech. Rep., 2005. [Online]. Available: http://www.sharespectrum.com/inc/content/measurements/nsf/1_NSF_Riverbend_Park_Report.pdf
- [4] R. W. Broderon, A. Wolisz, D. Cabric, S. M. Mishra, and D. Willkomm, "White paper: CORVUS: A Cognitive Radio Approach for Usage of Virtual Unlicensed Spectrum," UC Berkeley, Tech. Rep., 2004. [Online]. Available: http://bwrc.eecs.berkeley.edu/Research/MCMA/CR_White_paper_final1.pdf
- [5] R. H. Coase, "The Federal Communications Commission," *Journal of Law and Economics*, vol. 2, pp. 1–40, Oct. 1959.
- [6] T. W. Hazlett, "The Rationality of U.S. Regulation of the Broadcast Spectrum," *Journal of Law and Economics*, vol. 33, no. 1, pp. 133–175, Apr 1990.
- [7] A. S. de Vany, R. D. Eckert, C. J. Meyers, D. J. O'Hara, and R. C. Scott, "A Property System for Market Allocation of the Electromagnetic Spectrum: A Legal-Economic-Engineering Study," *Stanford Law Review*, vol. 21, no. 6, pp. 1499–1561, Jun. 1969.
- [8] E. Noam, "Taking the Next Step Beyond Spectrum Auctions – Open Spectrum Access," *IEEE Communications Magazine*, vol. 33, no. 12, pp. 66–73, Dec 1995.
- [9] Y. Benkler, "Overcoming Agoraphobia: Building the Commons of the Digitally Networked Environment," *Harvard Journal of Law and Technology*, vol. 11, pp. 287–400, Winter 1998.
- [10] J. Mitola, "Cognitive Radio: an integrated agent architecture for software defined radio," Ph.D. Thesis, KTH Royal Inst. of Tech., Stockholm, Sweden, 2000.
- [11] T. W. Hazlett, "Optimal Abolition of FCC Spectrum Allocation," *Journal of Economic Perspectives*, vol. 22, no. 1, pp. 103–128, 2008.
- [12] "An Act to regulate radio communication," Public Law No. 264, August 13, 1912, 37 Stat. 302.
- [13] "FCC Adopts Rules for Unlicensed Use of Television White Spaces." http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-08-260A1.pdf.
- [14] H. A. Shelanski and P. W. Huber, "Administrative Creation of Property Rights to Radio Spectrum," *Journal of Law and Economics*, vol. 41, no. 2, pp. 581–607, Oct 1998.
- [15] "United States v. Zenith Radio Corporation et al." <http://www.fcc.gov/mb/audio/decdoc/misc/US.v.Zenith-Radio.Corporation.html>, 1926.
- [16] C. Shannon, "A Mathematical Theory of Communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, 623–656, Jul. 1948.
- [17] L. Herzog, "'Public Interest' and the Market in Color Television Regulation," *University of Chicago Law Review*, vol. 802, 1951.
- [18] "Omnibus Budget Reconciliation Act of 1993 (OBRA 93)," U.S. Public Law 103-66. August, 2003, 103rd Congress, 1st session.
- [19] G. Gilder, "Auctioning the Airwaves," *Forbes ASAP*, Apr. 1994.
- [20] "The FCC Makes Spectrum Available for New Unlicensed Equipment," http://www.fcc.gov/Bureaus/Engineering_Technology/News_Releases/1997/nret7002.txt, 1997.
- [21] "New public safety applications and broadband internet access among uses envisioned by FCC authorization of ultra-wideband technology," http://www.fcc.gov/Bureaus/Engineering_Technology/News_Releases/2002/nret0203.txt, 2002.
- [22] "An Act for the regulation of radio communications, and for other purposes," Public Law No. 632, February 23, 1927, 69th Congress.
- [23] "Communications act of 1934," Public Law No. 416, June 19, 1934, ch. 652, 48 Stat. 1064, 73rd Congress.
- [24] D. Hatfield and P. Weiser, "Toward Property Rights in Spectrum: The Difficult Policy Choices Ahead," *CATO Institute*, no. 575, Aug. 2006.
- [25] E. Goodman, "Spectrum Rights in the Telecosm to Come," *San Diego Law Review*, vol. 41, no. 269, 2004.
- [26] R. J. Matheson, "The electrospace model as a frequency management tool," *Proceedings of the ISART Conference*, Mar 2003.
- [27] J. McMillan, "Selling Spectrum Rights," *Journal of Economic Perspectives*, vol. 8, no. 3, pp. 145–162, Summer 1994.
- [28] "U. s. spectrum management policy: Agenda for the future," U.S. Department of Commerce, Tech. Rep. NTIA Special Publication 91-23, Feb. 1991.
- [29] J. McMillan, "Why Auction Spectrum," *Telecommunications Policy*, vol. 19, no. 3, pp. 191–199, Apr 1995.
- [30] E. R. Kwerel and J. R. Williams, "Changing Channels: Voluntary Reallocation of UHF Television Spectrum," *FCC Office of Strategic Planning Working Paper Series*, no. 27, Nov 1992.
- [31] E. Kwerel and J. Williams, "A Proposal for a Rapid Transition to Market Allocation of Spectrum," *FCC Office of Strategic Planning Working Paper Series*, no. 38, Nov 2002.

- [32] “Comments of 37 Concerned Economists. Promoting Efficient use of Spectrum Through Elimination of Barriers to the Development of Secondary Markets,” *Before the Federal Communications Commission*, no. WT Docket 00-230, Feb. 2001.
- [33] E. Noam, “Spectrum Auctions: Yesterday’s Heresy, Today’s Orthodoxy, Tomorrow’s Anachronism. Taking the Next Step to Open Spectrum Access,” *Journal of Law and Economics*, vol. 41, no. 2, pp. 765–790, 1998.
- [34] Y. Benkler and L. Lessig, “Net Gains: Will Technology Make CBS Unconstitutional?” *The New Republic*, pp. 15–17, Dec. 1998.
- [35] T. M. Cover and J. A. Thomas, *Elements of Information Theory 2nd Edition*. Wiley-Interscience, 2006.
- [36] I. Csiszar, “Joint source-channel error exponent,” *Problems of Control and Information Theory*, vol. 9, no. 5, pp. 315–328, 1980.
- [37] G. Hardin, “The Tragedy of the Commons,” *Science*, vol. 162, pp. 1243–1248, 1968.
- [38] S. Buck, “Replacing Spectrum Auctions with a Spectrum Commons,” *Stanford Technology Law Review*, 2002.
- [39] C. Bazelon, “Licensed or Unlicensed: The Economic Considerations in Incremental Spectrum Allocations,” in *Proceedings of the Third IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Chicago, IL, Oct. 2008.
- [40] C. Fowler, J. Entzinger, and J. Corum, “Assessment of ultra-wideband (UWB) technology,” *IEEE Aerospace and Electronic Systems Magazine*, vol. 5, no. 11, pp. 45–49, Nov. 1990.
- [41] D. Geer, “UWB standardization effort ends in controversy,” *Computer*, vol. 39, pp. 13–16, Jul. 2006.
- [42] “News releases,” <http://www.wimedia.org>.
- [43] R. Tandra, S. M. Mishra, and A. Sahai, “What is a spectrum hole and what does it take to recognize one?” *Proceedings of the IEEE*, Jan. 2009.
- [44] A. Sahai, S. M. Mishra, R. Tandra, and K. A. Woyach, “Cognitive radios for spectrum sharing,” *IEEE Signal Processing Magazine*, Jan. 2009.
- [45] S. M. Mishra, A. Sahai, and R. W. Broderson, “Cooperative sensing among Cognitive Radios,” *ICC*, June 6-10 2006.
- [46] “Wireless Innovation Alliance.”
- [47] G. M. Rebeiz, *RF MEMS, Theory, Design and Technology*. John Wiley & Sons, 2003.
- [48] B. Lennett, “The Lobby that Cried Wolf: NAB Campaign Against Using TV White Space Follows a Familiar Script,” *New America Foundation Wireless Future Program Issue Brief*, no. 23, Oct 2008.
- [49] J. P. de Vries, “De-situating spectrum: Rethinking radio policy using non-spatial metaphors,” *IEEE Symposia on New Frontiers in Dynamic Spectrum Access Networks*, Oct 2008.
- [50] P. J. Weiser and D. Hatfield, “Spectrum Policy Reform and the Next Frontier of Property Rights,” *George Mason Law Review*, vol. 60, no. 3, 2008.
- [51] T. W. Hazlett, “A Law and Economics Approach to Spectrum Property Rights: A Response to Weiser and Hatfield,” *George Mason Law Review*, vol. 15, no. 4, Summer 2008.
- [52] K. Werbach, “Supercommons: Toward a Unified Theory of Wireless Communication,” *Texas Law Review*, vol. 82, no. 4, pp. 863–973, March 2004.
- [53] P. J. Weiser and D. N. Hatfield, “Policing the Spectrum Commons,” *Fordham Law Review*, vol. 74, no. 2, pp. 663–694, 2005.
- [54] G. R. Faulhaber, “Wireless telecommunications: Spectrum as a critical resource,” *Southern California Law Review*, vol. 79, no. 3, Mar. 2006.
- [55] —, “Deploying Cognitive Radio: Economic, Legal, and Policy Issues,” *International Journal of Communication*, vol. 2, pp. 1114–1124, 2008.
- [56] L. Lessig, *Code v2.0*. Basic Books, 2006.
- [57] R. Etkin, A. Parekh, and D. Tse, “Spectrum sharing for unlicensed bands,” in *First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Baltimore, MD, Nov. 2005.
- [58] A. Sahai, N. Hoven, and R. Tandra, “Some fundamental limits on cognitive radio,” *Forty-second Allerton Conference on Communication, Control, and Computing*, Oct. 2004.
- [59] F. Perich, “Policy-based network management for NeXt generation spectrum access control,” in *Proceedings of the Second IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Dublin, Ireland, Apr. 2007.
- [60] “In the matter of unlicensed operation in the tv broadcast bands: Second report and order and memorandum opinion and order,” Federal Communications Commission, Tech. Rep. 08-260, Nov. 2008.
- [61] C. R. Stevenson, C. Cordeiro, E. Sofer, and G. Chouinard, “Functional requirements for the ieee 802.22 wran standard,” IEEE P802.22, Tech. Rep. IEEE 802.22-05/0007r48, Sep. 2005.
- [62] R. Tandra and A. Sahai, “SNR walls for signal detection,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 1, pp. 4–17, Feb. 2008.
- [63] K. Gödel, “Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme,” *I. Monatshefte für Mathematik und Physik*, vol. 38, pp. 173–198, 1931.
- [64] I. Asimov, *Robots of Dawn*. Doubleday, 1983.
- [65] L. B. Solum, “Legal Personhood for Artificial Intelligences,” *North Carolina Law Review*, pp. 1231–1287, Apr. 1992.

- [66] P. M. Asaro, "Robots and Responsibility from a Legal Perspective," *Proceedings of the IEEE International Conference on Robotics and Automation*, Apr. 2007.
- [67] D. P. Satapathy and J. M. Peha, "Spectrum Sharing Without Licenses: Opportunities and Dangers," *Interconnection and the Internet: Selected Papers From the 1996 Telecommunications Policy Research Conference.*, 1996.
- [68] —, "Etiquette Modification for Unlicensed Spectrum: Approach and Impact," *Vehicular Technology Conference*, vol. 1, pp. 272–276, May 1998.
- [69] W. Xu, P. Kamat, and W. Trappe, "TRIESTE: A trusted radio infrastructure for enforcing spectrum etiquettes," *First IEEE Workshop on Networking Technologies for Software Defined Radio Networks*, Sep. 2006.
- [70] G. Atia, A. Sahai, and V. Saligrama, "Spectrum enforcement and liability assignment in cognitive radio systems," in *Proceedings of the Third IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*, Chicago, IL, Oct. 2008.
- [71] G. Atia, V. Saligrama, and A. Sahai, "Spectrum enforcement and liability assignment in cognitive radio systems," in *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, Nov. 2008.
- [72] G. S. Becker, "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, vol. 76, 1968.

Appendix A

Markov Chain for Simulations

The Markov chain depicted in Fig. A.1 is useful for conceptually understanding how the model works. However, for simulation purposes, it is not simple to calculate transition probabilities for this chain. Therefore, in this appendix, we give a description of the actual chain used for all simulations and calculations.

Because we have for simplicity assumed that all of the bands are identical in utility with primary users who have identical characteristics, we can keep track of how many bands are available, instead of the cognitive user's activity in each individual band. We note also that because the action of the primaries is memoryless, we do not care which primaries are active at any given time, just that a certain number of them are. Therefore, we can model the situation as the Markov chain in Fig. A.2. The states distinguish only whether the cognitive user is in or out of jail and how many primary users are active. The utility functions and transitions, then, capture the cheating and false alarm behavior.

The transitions horizontally along the chain are dependent only on the primary users. To move to another state horizontally, we need to keep track of how many primaries are turning off and on. Therefore, the transitions can be captured by the horizontal movement from i bands being available to j bands being available. The number of primaries currently transmitting is $B - i$, where B is the total number of possible cognitive bands, and the number currently not transmitting is i . Let $k = |i - j|$ or the total number of bands that will change states. The transition probabilities in the horizontal direction are:

$$P_{horiz}(i, j) = \begin{cases} \sum_{m=0}^{\min(B-i-k, i)} \binom{B-i}{m+k} \binom{i}{m} p^{m+k} (1-p)^{B-i-m-k} q^m (1-q)^{i-m}, & i < j \\ \sum_{m=0}^{\min(B-i, i)} \binom{B-i}{m} \binom{i}{m} p^m (1-p)^{B-i-m} q^m (1-q)^{i-m}, & i = j \\ \sum_{m=0}^{\min(B-i, i-k)} \binom{B-i}{m} \binom{i}{m+k} p^m (1-p)^{B-i-m} q^{m+k} (1-q)^{i-m-k}, & i > j \end{cases} \quad (\text{A.1})$$

The movement vertically in the chain is determined by the cheating behavior of the cognitive user, and the enforcement parameters of the regulator. The transition down captures the movement of the cognitive user into jail. We know that the cognitive user will cheat on all of the bands available if it is worthwhile to cheat. We will assume that there is a unique probability of cheating associated

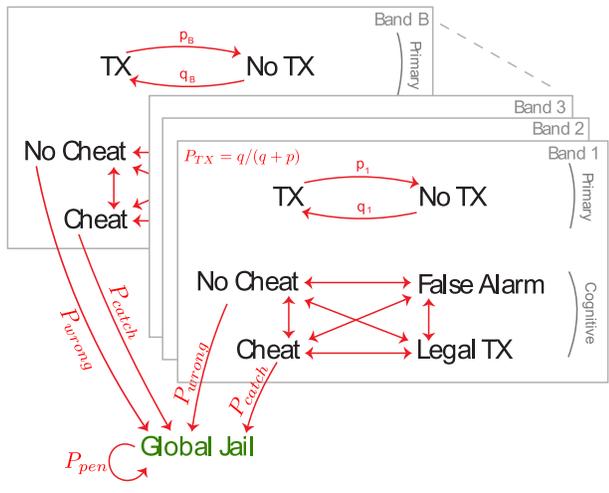


Figure A.1. The Markov chain used for illustration.

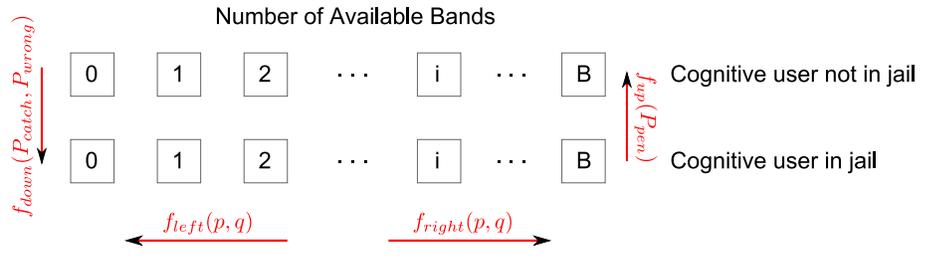


Figure A.2. The Markov chain actually used for simulations.

with horizontal state. Therefore, If i is the current horizontal state and $B - i$ primary users are currently active, the transition into jail happens with probability:

$$P_{i,vert}(no_jail, jail) = 1 - (1 - P_{catch})^{iP_{cheat,i}}(1 - P_{wrong})^{i-iP_{cheat,i}} \quad (A.2)$$

Because the jail is global and does not care in how many bands the cognitive user was caught cheating, the transition out of jail remains, as in the body of the thesis, $1 - P_{pen}$.

The false alarm behavior in reality only affects the utility of the cognitive user. Therefore, in this simulation model, it appears only in the utility function of the cognitive user:

$$U_{cog} = \sum_{i=0}^B (1 - P_{FA})\pi_{i,no_jail} - \beta\pi_{jail}, \quad (A.3)$$

where π_{i,no_jail} is the stationary probability of having i bands available while not in jail and π_{jail} is the stationary probability of being in jail.