Sharing Spectrum through Spectrum Policy Reform and Cognitive Radio

Jon M. Peha
Carnegie Mellon University

Abstract

Traditionally, interference protection is guaranteed through a policy of spectrum licensing, whereby wireless systems get exclusive access to spectrum. This is an effective way to prevent interference, but it leads to highly inefficient use of spectrum. Cognitive radio along with software radio, spectrum sensors, mesh networks, and other emerging technologies can facilitate new forms of spectrum sharing that greatly improve spectral efficiency and alleviate scarcity, if policies are in place that support these forms of sharing. On the other hand, new technology that is inconsistent with spectrum policy will have little impact. This paper discusses policies that can enable or facilitate use of many spectrum-sharing arrangements, where spectrum-sharing arrangements are categorized as being based on coexistence or cooperation, and as sharing among equals or primary-secondary sharing. A shared spectrum band may be managed directly by the regulator, or this responsibility may be delegated in large part to a license-holder. The type of sharing arrangement and the entity that manages it have great impact on which technical approaches are viable and effective. The most efficient and cost-effective form of spectrum sharing will depend on the type of systems involved, where systems under current consideration are as diverse as television broadcasters, cellular carriers, public safety systems, point-to-point links, and personal and local-area networks. In addition, while cognitive radio offers policy-makers the opportunity to improve spectral efficiency, cognitive radio also provides new challenges for policy enforcement. A responsible regulator will not allow a device into the marketplace that might harm other systems. Thus, designers must seek innovative ways to assure regulators that new devices will comply with policy requirements and will not cause harmful interference.

Keywords: spectrum sharing, spectrum policy, cognitive radio, band manager, unlicensed, etiquette, opportunistic access, interruptible access, secondary spectrum market, commons

1 The author gratefully acknowledges the financial support of CyLab, ARO, and Intel Corp.
I. INTRODUCTION

When airplanes were new, pilots flew as they pleased, without risk of midair collision. As more planes shared the skies, the need grew for communications and cooperation among pilots. Today, hundreds of planes can fly just a few meters apart in perfect formation through complex maneuvers, all without collision. This could never occur without the technological advances that made precision flying possible, but technology alone would have accomplished little. Before planes could be so densely packed, it was necessary to invent new rules that govern how pilots fly and interact with each other.

Similarly, as the density of wireless transmitters grew, rudimentary coordination in the form of policy became essential. In most prime spectrum, licensing gave communications systems exclusive access to blocks of spectrum, thereby almost eliminating the danger of harmful interference, but leaving the majority of this spectrum idle when and where the license-holder is not active [1]. A few bands were designated for unlicensed devices with few restrictions, but with power limits that generally kept utilization low enough to limit mutual interference (although utilization and serious interference problems sometimes grow over time). These approaches were reasonably effective as a means of limiting interference problems, and appropriate for the technology of their time. However, overwhelming demand for wireless products and services motivates us to support a greater density of wireless devices through adoption of new technology and policy. Emerging technology, including cognitive radio [2], mesh networks, location technologies, spectrum sensors, and even micropayment schemes will make this possible, if and only if we simultaneously develop spectrum policies that make effective use of the new technology.

Wireless technology that is irreconcilable with important policy concerns, or that was implicitly designed for the wrong policy environment, is of no practical value. With something as potentially revolutionary as cognitive radio, it is easy for designers to make this mistake. For example, some cognitive radio systems considered by researchers are based on the assumption that all devices fully cooperate to reduce interference, even when these devices are owned and operated by countless independent users. While this is not impossible with corresponding policy reform, designers who do not realize that they are implicitly proposing a dramatic change in policy may be disappointed in the limited impact of their technical innovations. Even when new technology is compatible with current spectrum policy, that technology may be prohibited because wise regulators are necessarily cautious about allowing new devices that may have unintended impact on existing systems.

Successful innovation requires consideration of the entire system, where a system includes wireless technology, spectrum policy, and the human beings who interact with both. This includes feasible and cost-effective technical mechanisms through which devices can minimize the interference they cause (and perhaps maximize the interference they tolerate). It also includes the incentive structure that would convince the designers and users of wireless devices to reduce the interference they cause others. Regulators
must define rules that facilitate use of those mechanisms that are effective and mutually compatible, deter use of mechanisms that are ineffective or incompatible with other systems in the given band, and still allow enough flexibility to encourage innovation. At the same time, regulators must ensure the rules that unleash new technologies while controlling interference are legally and technically enforceable. Otherwise, the new technologies are not ready for use.

Increased spectrum sharing is essential to addressing today’s serious scarcity of available spectrum. This paper will discuss a range of spectrum policies that can make novel forms of spectrum sharing possible, and a taxonomy to better understand those forms of spectrum sharing. We will also discuss the implications of these policies and spectrum-sharing arrangements for various technologies, with an emphasis on cognitive radio that befits this special issue of the Proceedings of the IEEE. Indeed, cognitive radio is in part in a race with other technical approaches for ascendance in spectrum sharing. Note that the pros and cons of various categories of spectrum-sharing, as well as the taxonomy proposed in this paper which defines these categories, represent the views of the author. This paper addresses many highly contentious issues for which there is no consensus opinion.

In a given spectrum band, adoption of one form of spectrum-sharing usually precludes many other sharing possibilities. Since policy must change slowly to protect legacy systems, a policy decision in favor of one form typically precludes the alternatives for many years. For example, allowing unlicensed devices to operate in the “white spaces” of TV spectrum [3], i.e. in a TV channel but far from any TV tower operating in that channel, precludes the possibility of offering licenses to those who would share with TV broadcasters in a different manner [4]. Thus, the policy and technical communities must make hard choices. The good news is that they can (and should) make different choices in different bands.

Section 2 will describe the basic features by which spectrum-sharing schemes can best be defined. Section 3 uses these features to present and categorize a wide range of new spectrum-sharing paradigms, and their implications for both policy and technology. Section 4 addresses the important question of who governs spectrum policy within a given band, which also has implications for technical design. Issues of policy enforcement are discussed in Section 5, and the paper is concluded in Section 6.

II. DEFINING FEATURES OF SPECTRUM-SHARING PARADIGMS

We argue that a sharing arrangement can be well characterized and categorized by two defining features [5, 6]. The first is whether sharing is based on cooperation or coexistence. The distinction has profound implications for both policy and technology. In a model based on cooperation, even devices under different administrative control must communicate and cooperate with each other to avoid mutual interference. Among
other things, this means a common protocol must be defined, and must be supported by all systems in the band. With a coexistence model, devices try to avoid interference without explicit signaling. At most, devices sense each other’s presence as interference. As can be seen throughout this special issue, cognitive radio is a powerful tool for sharing based on coexistence; the ability to reconfigure a device based on the sensed interference levels from neighbors is valuable when avoiding mutual interference [7, 8, 9]. With a cooperative model, devices may react to what they are specifically told, and this ability to independently sense the environment may or may not be important, depending on the scheme.

Both approaches have their place. Where devices can exchange information that would otherwise be unavailable, we would expect these devices to make more efficient use of spectrum while better avoiding mutual interference [10]. However, cooperation has its drawbacks. Thanks to hidden terminals and other pragmatic concerns, communications may not always be possible. The need for interaction can also impose transaction costs, including device complexity, delay, and communications overhead, perhaps even to the extent that such costs outweigh the benefits of cooperation [11]. Moreover, the need for a common protocol in cooperative sharing forces a degree of homogeneity among devices that can inhibit innovation. This is also a problem when attempting to share spectrum with legacy equipment, which was not designed or deployed with new sharing arrangements in mind. Forcing systems with no trust relationship to cooperate also poses complex security challenges. Such tradeoffs will be further apparent in the next section, when we discuss specific sharing models. Further research will determine when coexistence is preferable and when cooperation is preferable. In many cases, extensive use of one within a given band makes the other impractical, if not impossible.

The second defining feature of a spectrum-sharing arrangement in the proposed taxonomy is whether the arrangement comprises sharing among equals or primary-secondary sharing [5, 6]. In the latter case, some systems have the right to operate as a primary spectrum-user, and policy mandates that secondary devices are not allowed to cause harmful interference to a primary system. In the former case, all devices have equal rights, and typically more flexibility about how to behave in the presence of peers. As a result, an important question wherever sharing among equals takes places is whether all systems have incentive to limit the interference they cause to others. Cognitive radio is most often discussed as a tool for primary-secondary sharing, but if designed in accordance with an appropriate policy, cognitive radios may prove useful for sharing among equals as well.

Probably the most commonly used descriptive feature for a class of wireless devices is whether they are licensed or unlicensed. Note that this applies to a set of systems within a spectrum-sharing arrangement, and not to the arrangement itself. As we will see in Section 3, primary devices can be licensed or unlicensed, and secondary devices can be licensed or unlicensed. A licensed system must get permission from the regulator to operate within a given frequency band. The licensing process is an opportunity for the regulator to ensure exclusive access to a block of spectrum if it wishes,
which is strong protection from the problems of interference and congestion. In contrast, unlicensed devices need no permission from the regulator to deploy a device. Devices can typically be deployed anywhere, which means there is no limit to the number of devices that might be operating in a given location.

III. MODELS FOR SPECTRUM SHARING

This section presents the policies underlying a wide range of spectrum-sharing models. Section 3.1 addresses the “commons” models, in which many devices are allowed to share spectrum as primary users. Section 3.2 discusses how secondary devices can share spectrum with a primary license-holder. Building on Sections 3.1 and 3.2, the methods by which these secondary systems share with each other are discussed in Section 3.3. Section 3.4 discusses sharing among equals for a limited number of regional providers, rather than a potentially unlimited number of localized systems. The taxonomy is summarized in Section 3.5

A. Sharing Among Equal Primary Devices

Some of the most heated debates over spectrum policy in the last few years have been over the creation of new spectrum “commons.” In any commons model, spectrum is shared, and no one is given clear priority. As discussed in Section 2, policy can prescribe devices that cooperate or devices that merely coexist. While both possibilities are sometimes lumped together under the ambiguous heading of “commons,” the two are fundamentally different. (Unfortunately, this confusion over definitions has made much of the debate about spectrum commons meaningless, as advocates and opponents based their arguments on incompatible definitions.) [5, 6].

The coexistence model exists today in many (but not all) nations, and has spurred tremendous innovation and productivity, as is demonstrated by the popularity of WiFi and cordless phones [12]. Many of these devices were born because of the creation of new unlicensed spectrum bands, or to be more precise, bands where devices are “licensed by rule.” In other words, any device that complies with the regulator’s rules can be deployed without explicit permission. (Equivalently, a regulator can explicitly grant an unlimited number of nonexclusive licenses with some rules associated with these licenses, such as the requirement to use contention-based algorithms in the 3650-3700 MHz band within the US [13].) Unlicensed spectrum has many advantages [14, 15]. It requires spectrum sharing, which can potentially lead to greater spectral efficiency than exclusive access, where spectrum often sits idle because the license-holder is not transmitting. Unlicensed spectrum is necessary to support entire systems that are mobile or portable, such as a group of laptops that form an ad hoc wireless local-area network wherever they

---

2 This category of mobile or portable systems does not include systems that consist of fixed infrastructure, and mobile devices that work only when in the area served by fixed infrastructure, e.g. cellular systems.
happen to be. It is useful for inexpensive low-power consumer products such as cordless phones, where the cost of coordination and licensing would probably far exceed the cost of the equipment, and the interference impact on neighbors is small.

Of course, a commons based on coexistence necessarily comes with the risk of mutual interference. In today’s unlicensed bands, there is no limit to the number of devices in a given location, and no limit to the potential congestion. This is one reason such bands are typically limited to low-power devices. Cost-effective market-based methods have been suggested [16] but not adopted that would allow the regulator to limit the number of devices deployed in a nation if interference proves problematic, without imposing additional constraints beyond the traditional unlicensed paradigm, but even that does not prevent all of the devices in a nation from appearing in the same place. Consequently, only applications and devices that do not require quality of service guarantees, such as wireless local-area networks, should be deployed in these unlicensed bands.

Use of devices that detect each other’s presence and change behavior accordingly, such as those using cognitive radio, could reduce the problems of mutual interference. But would such techniques ever be deployed in a coexistent commons? The success of this approach probably depends on the policies that prevail in the band. As described in Section 2, it is not sufficient with sharing among equals for devices to have the technical ability to reduce the interference they cause to others. Users and designers of these devices must also have the right incentives. Some engineers may design “greedy” devices, i.e. devices that transmit with greater power, duration, or bandwidth than necessary [15]. When choosing between improving the performance, cost, or battery life of my device, versus conserving our spectrum and reducing interference inflicted on some stranger’s device, there is a natural tendency to choose the former. In the extreme, the presence of greedy devices can lead to a tragedy of the commons, where many devices consume too much of the shared resource, and all devices in the band experience inadequate performance as a result [17, 15, 18]. For example, in the 1990s, the US Federal Communications Commission established an unlicensed band in which all devices must comply with a “listen-before-talk” (LBT) requirement. We have shown using game theory that under some circumstances, devices in this band could improve their own performance by transmitting for longer durations than necessary. Moreover, if one device chooses to be greedy in this manner, a neighboring device would optimize its own performance by being considerably more greedy than the first, potentially leading to an escalation of overconsumption that severely degrades the performance of all [15, 18]. Similar problems can occur when greedy devices transmit at greater power than necessary, as shown in [19], or when they use more bandwidth than necessary.

The dangers of greed can be addressed in two ways. One option is to keep spectrum utilization sufficiently low that performance is good, and there is little incentive for greed. This might be done through power limits, deployment fees, wideband allocations, or other means. The other option is to build incentives to conserve spectrum and reduce interference to neighbors into the technical rules of operation defined by the regulator, which as shown in [15, 20, 19], has the effect of changing the game’s incentive
structure. For example, a device that has been transmitting for a long time or at high power might be required to defer to other devices by delaying its next transmission until it senses the channel to be idle throughout a fairly long monitoring period. These regulator-imposed rules, which are known as an *etiquette*, do not define how devices must be designed, but they impose some constraints on design. A well-designed etiquette could unleash the power of cognitive radio to improve spectral efficiency in unlicensed bands. On the other hand, a poorly designed etiquette may lead to ineffective designs, or slow technical innovation.

The characteristics of a commons based on cooperation are quite different [5, 6, 21, 22]. In this commons, policy dictates that all devices will actively communicate and cooperate, even though they serve different owners. At minimum, devices can negotiate with each other for the right to transmit without collision. For greater benefits, devices might autoconfigure into a mesh network, and carry each other’s traffic. It has been shown theoretically that cooperation can lead to *cooperative gain*, i.e. the capacity in the system can actually increase with the number of active devices [21, 23, 24]. As more devices are added, the mean distance between devices decreases, allowing devices to transmit at lower power, thereby conserving spectrum. Thus, users of a commons based on cooperation need not fear oversubscription as much as users of a commons based on coexistence. Another advantage of this approach is that when devices from many owners combine to form a single infrastructure for all, this fundamentally changes the economics of wide-area coverage. Thus, for example, a cooperative approach might make it cost-effective to cover a large region in which today’s WiFi-based microcellular systems would not be financially sustainable, as discussed in [25].

The potential benefits of a cooperative commons are great, and so are the challenges, which transcend technology and policy. Once again, with sharing among equals, we must consider the incentives of the users and designers of devices. When devices delay their own transmissions for another, or carry each other’s traffic, some altruism is required, e.g. one device might increase delays for its own traffic and drain its own battery by transmitting a stranger’s packet. For example, we have shown that devices can manipulate today’s routing protocols to avoid carrying traffic for other devices, thereby improving their own performance, and it is only possible to detect such behavior with non-standard protocol modifications [26]. Even worse, a malicious node may take deliberate steps to disrupt the network, and in some proposed cooperative schemes, the impact of malicious behavior could be great. Unless adequate technical and policy safeguards are devised to address these problems, the cooperative commons will have limited application.

Another inherent challenge that is unique to the cooperative commons is that all devices must share a detailed communications protocol. A device should not be deployed in the band unless it can communicate using this protocol. Traditional voluntary standards will not suffice. Who will specify this protocol, and make changes to it over time as technology evolves? This difficult challenge will be discussed further in Section 5 which addresses who must be in charge, and Section 6 which addresses enforcement.
Note that when defining a sharing arrangement in a given band, one must choose between cooperation and coexistence. The technologies underlying these policies are quite different, and each has its own set of technical challenges. Those who champion cognitive radio will probably find greater use for their ideas in a coexistence regime, although there can be applications with cooperation as well [2].

**B. Sharing Between Primary and Secondary**

Primary-secondary sharing is probably most useful when the primary system has been given exclusive rights through licensing, as there are generally times and/or locations where other devices could transmit in this spectrum without causing harmful interference. Once again, sharing can take one of two forms: cooperation or coexistence [5, 6].

When sharing is based on coexistence, secondary devices are essentially invisible to the primary. Thus, all of the complexity of sharing is borne by the secondary. No changes to the primary system are needed, which is especially good for legacy systems that are difficult to change. Before any secondary devices can be deployed, rules must be established to protect the primary. Secondary devices can either be allowed to transmit at such low power that they never cause harmful interference to the primary, as with ultrawideband [27], or they can be allowed to transmit opportunistically when and only when they determine that transmissions will not cause harmful interference [7]. Opportunistic access is probably inadequate for applications that require guaranteed quality of service, and therefore guaranteed prompt access to spectrum. However, many applications can benefit from opportunistic access.

The challenges of opportunistic access are well suited to cognitive radio combined with spectrum sensors. Alternatively, it is possible to ensure that a secondary transmitter is sufficiently far from primary receivers using a location technology such as the global positioning system (GPS) in combination with a database containing the locations of primary devices. The secondary device also needs a method of retrieving the inevitable changes to this database, and the device must stop all transmissions if its database is too far out of date. This database might be extended to a full radio environment map [28], which also includes information on spectrum policy, available services, geography, and even path loss.

Proponents of opportunistic access must overcome several technical challenges; the first is proving that primary systems will be protected. At present, this battle is being fought over opportunistic access to the “white space” in television spectrum. The US Federal Communications Commission (FCC) has tentatively ruled in favor of this form of sharing after the 2009 transition to digital television [3], although many of the details remain undecided, and for better or worse, there is still time for a change in course. Two prototype devices have already been submitted to the FCC for evaluation [29].
In policy circles, the debate has often focused on the technical question of whether secondary devices do or do not interfere with television, which is clearly the wrong question. All transmissions cause interference. The appropriate question is whether this interference rises to the level of harmful interference, which is not a purely technical question [30]. Interference is harmful if it is great enough to cause a significant disruption in service. This definition of harmful interference is much easier to apply when the primary system operates effectively as long as the signal to noise ratio (SNR) exceeds an established threshold, and one can demonstrate that this threshold is greatly exceeded. Unfortunately, broadcast technologies like television do not work that way. There is always an edge to coverage area, where SNR is at the minimum acceptable level. The slightest interference is harmful for a receiver at that exact location. The question then becomes how much coverage area can be lost before interference is considered harmful? [4] 10%? 1%? 0.1%? This is a non-technical and largely subjective question. Moreover, the problem is compounded by the notorious difficulty of predicting the exact contours of coverage at any given SNR threshold through predictive modeling [31].

When sharing is instead based on cooperation, the primary and secondary systems interact. For example, a secondary device may ask the primary for permission to use spectrum before transmitting. This exchange provides an opportunity for the primary to guarantee quality of service for the secondary, which is an important advantage of cooperation over coexistence for the secondary device. This is also an opportunity for the license-holder to demand payment, which is an advantage of cooperation for the primary spectrum-user. If payment is demanded, this is a form of secondary spectrum market, but one that operates in real time [32]. There has been discussion about creating a secondary market where these blocks of spectrum can be “rented” out for months or years [33]. It is important to note that more dynamic forms of sharing, where spectrum is given out for minutes or milliseconds, are also possible, and in some countries, legal [34]. An important variation on this model is that of band manager [5, 34]. The regulation is similar; one entity is given a license to control a block of spectrum, and the rights to grant temporary access for a fee through a secondary market. Again, access might be allowed for a short period, as in [35], or for a very long period. The difference in models is that unlike the cellular provider discussed above, the band manager has no need for spectrum other than to rent it out.

Whether the primary has its own need for spectrum or not, access to spectrum may be granted for free, for a fixed fee, or through some form of auction. Auctions allocate resources to those who value those resources the most, which is one reason many countries now assign long-term spectrum licenses through auctions [14]. It has been suggested that a policy that allows a real-time auction for spectrum could yield similar benefits [36], and many papers have been written recently to address related technical issues. Indeed, sufficiently dynamic real-time auctions can play a role similar to technical congestion control mechanisms, and under some circumstances, are almost (but not quite) as effective at relieving congestion [37]. While there are certainly benefits to auctions, there are greater challenges when auctions are used for real-time allocation, especially if one is auctioning the right to transmit for a small time period or a small geographic area. First, auctions are effective when multiple parties want the same thing,
and granting access to one bidder means denying access to the others. Unless all devices are collocated, this is not necessarily the case with spectrum, and this greatly complicates the problem. Second, multiple parties can bid in an auction when their need for spectrum occurs at the same time, but not when requests for spectrum are sporadic and unsynchronized, as might be the case. Finally, in contrast to a simple commons model, auctions require an auctioneer, and associated infrastructure [15]. The problem gets easier if one increases the granularity of the spectrum being auctioned in both time and geographic area, but this comes at the cost of spectral inefficiency.

A cooperative model is also useful for systems that need guaranteed quality of service, but use the spectrum infrequently. For example, an infrastructure built exclusively for public safety agencies (e.g. firefighters, police, providers of emergency medical services) might be given primary rights to a block of spectrum. This spectrum could be used the vast majority of the time by commercial service providers, but they would only be granted interruptible access to this spectrum, so public safety agencies have ample spectrum during major emergencies [38].

The US Federal Communications Commission recently adopted a sharing policy that takes advantage of these characteristics, although it is the sharing of infrastructure and spectrum rather than the sharing of spectrum alone that is the focus [39]. Under this FCC plan, it is envisioned that a commercial license-holder and a public safety organization would each be allocated adjacent spectrum bands nationwide. The commercial provider would build out a common broadband infrastructure, such that both the spectrum and the infrastructure will be used by both public safety agencies and the general public. In the initial auction, no commercial company was interested in making a bid beyond the required minimum. It remains to be seen what will happen in this band.

As with primary-secondary coexistence, the practicality of the cooperative approach depends on the applications involved, and other factors. For example, the primary system needs devices throughout the service area that can provide admission control and possibly toll collection, which is much easier when the primary system is a cellular system rather than a broadcaster. Indeed, this approach may someday prove commercially attractive to cellular carriers. We have analyzed scenarios where extensive communications among secondaries is possible with little impact on the primary, for a case where the primary license-holder is a commercial cellular system [32]. Effective spectrum-sharing was facilitated by a variety of technologies, including location technology which was used to enhance frequency reuse, and secure micropayment technology so that primary systems can receive payments from previously unknown secondary devices [40].

C. Sharing Among Equal Secondaries Devices

It is sometimes assumed that secondary devices must be unlicensed. Actually, like primary users of spectrum, secondary users can be licensed or unlicensed. Both licensed and unlicensed secondaries are prohibited from causing harmful interference to
the primary. One difference is that a licensed secondary system need not worry about interference from other secondaries. Thus, quality of service can be guaranteed for the secondary when (and only when) activities of the primary do not get in the way.

Moreover, because there is only one entity causing interference to the primary, it may be easier to ensure that interference never reaches a level that would be harmful. This may also make it possible to safely use much higher-power secondary devices. On the other hand, if access is unlicensed, many more systems may make use of the spectrum.

This policy debate is currently most relevant to the TV white space. As discussed in Section 3.2, the FCC has tentatively concluded that devices may operate in the white space, but one of the issues left undecided is whether those devices should be licensed or unlicensed [3]. Should white space be licensed for use by commercial cellular companies? (With 3-4 TV channels, a carrier could cover the nation [4, 41].) Or should it be unlicensed devices that use cognitive radio, such as the prototypes that Philips and Microsoft have recently unveiled [42, 43, 29]? In the end, the decision may depend on what kinds of products and applications are in greatest demand by the public, and views on this could easily differ from country to country.

If multiple systems are allowed to operate as secondary users, they can (as usual) coexist or cooperate with each other. Coexistence is technically simpler in many ways. Early successes such as [44, 45] of using cognitive radio to avoid interference with a primary system are coming from devices that determine on their own whether it is safe to transmit. Of course, that does not solve the problem of mutual interference among secondary devices, for whom all of the dangers of a simple commons discussed in Section 3.1 remain, including harmful interference, overconsumption of spectrum due to greedy designs, and even the tragedy of the commons. Clearly, imposing a primary system on top of what would otherwise be a commons cannot make the usual problems of a commons any easier.

When there is only coexistence between primaries and secondaries, there are good technical reasons to want cooperation among secondary devices. Having more devices capable of sensing the primary and exchanging measurements is a good defense against hidden terminals, exposed terminals, and imprecise sensor measurements. Cooperation among secondaries could ensure that multiple devices will not combine to cause interference that is harmful to the primary. Moreover, cooperation may allow secondary devices to more effectively avoid interfering with each other. For example, after coordinating on a designated signaling channel, frequency-agile devices might switch to bands that are less congested. Thus, the IEEE 802.22 standards group [46, 47] and a number of researchers (e.g. [48, 43, 49, 50]) have considered cooperative mechanisms for secondary devices.

However, the success of some cooperative approaches may depend on significant policy changes. It is easy to envision cooperation among homogenous devices in the same administrative domain, but will heterogeneous devices with different owners cooperate effectively, and will all systems still work efficiently and effectively if cooperation breaks down? A technical “solution” that works only if devices all cooperate,
without the technical and policy means to ensure cooperation, is useless at its best and
dangerous at its worst; such a system may convince policy-makers to allow something
that will ultimately lead to harmful interference. There may be good reasons for
cooperation even if there is no way to force all devices to cooperate, but in this case,
protection against harmful interference must be guaranteed even in the absence of
cooperation. There are policy reforms that can require full cooperation of all devices, but
this comes at a cost, and we must assess whether those policy reforms are worthwhile.

As with the cooperative commons discussed in Section 3.1, the problem is
compounded by the fact that devices may be amply rewarded for providing
misinformation to neighbors [26], perhaps convincing those neighbors to switch to
another band by falsely claiming that the primary system is active. If it is necessary that
all devices must follow the same cooperative algorithms, then someone must define those
algorithms, enforce them on all secondary devices deployed in the band, and handle
violations. Once again, voluntary standards will not suffice, and this profoundly affects
which technical designs are viable. The question of who is in charge of defining
algorithms is discussed in Section 4, and challenges of enforcement are discussed in
Section 5.

D. Sharing Among Equal Regional Infrastructures

Another form of spectrum sharing among equals occurs when a handful of
regional service-providers share spectrum. Each service-provider is seeking cost-
effective ways to deploy infrastructure and exploit spectrum resources so as to cover its
entire service area. This raises many of the same challenges as the sharing-among-equal
models discussed in Section 3.1, but because of the smaller number of players and the
large investment in wide-area infrastructure, this scenario deserves separate consideration.

As Hatfield and Tenhula have pointed out, one of the first great success stories for
dynamic spectrum sharing comes from the land mobile radio (LMR) systems used for
public safety [51]. In the US, many thousands of public safety agencies have their own
communications infrastructure [38], even when coverage areas substantially overlap.
(This is one reason the number of transmission towers used by public safety agencies in a
US county depends more on the number of municipal governments in that county than on
area or population [52].) Giving each public safety agency its own dedicated
communications channel leads to tremendous spectral inefficiencies. To avoid a
spectrum shortage, trunking was introduced, whereby LMR systems share channels
dynamically, either in a centralized or decentralized manner. The latter can be seen as a
forerunner of cognitive radio, in that systems dynamically access spectrum after sensing
its availability [51]. With trunked LMR, there have been relatively few complaints about
the usual problem with sharing among equals - that some users claim more of the shared
resource than needed. Indeed, trunking may have reduced this very problem, since it
reduced the incentive for public safety agencies to permanently horde spectrum that they
do not (yet) need. Such selfless behavior is encouraging, although there is reason to be
more hopeful in this scenario than most. It is always the same small number of public
safety agencies that share a given spectrum pool, i.e. a “repeated game,” and each public safety agency has a genuine interest in the success of the agencies with whom it shares spectrum.

The benefits of sharing are also great for commercial cellular systems, but the incentives are different. Consider the case where several cellular carriers have collocated transmitters to serve the same cell. At a given time, one provider may be at peak capacity, and therefore forced to block any additional calls. Meanwhile, other providers may have idle spectrum at that moment that could be used for the new call. We have shown that if these providers shared spectrum, they could reduce blocking probability for all providers, or equivalently, they could achieve the same blocking probabilities with much less spectrum [53]. Moreover, although it is less obvious, simulations show that the benefits of sharing are as good or better when towers are not collocated. So why has this not occurred? The problem is that these carriers would not have incentive to share spectrum in a selfless manner. Some may simply transmit when transmission is unnecessary, to deprive their rivals of spectrum. Even if this can be avoided, the bigger problem may occur in infrastructure deployment. Any cellular carrier can improve spectral efficiency by deploying more towers and thereby reducing cell size, but this significantly increases costs. A carrier has little incentive to accept the full burden of this cost when the resulting benefits of spectral efficiency are shared with its rivals; if each carrier pursues its own self-interest independently, this can lead to poor performance for all [53].

A combination of technology and spectrum policy reform is needed to solve this problem. One policy solution is to grant each of the carriers the presumption of control over a portion of the spectrum, and the right to negotiate for highly dynamic access to each other’s spectrum, as proposed for the Dynamic Radio for IP-Services in Vehicular Environments (DRIVE) system [54, 55]. Such negotiation might allow each carrier comparable incursions into the other’s spectrum, or the carrier making greater use of its neighbor’s spectrum might pay for the privilege. Another viable policy is to use a variant of the band manager described in Section 3.2, wherein the band manager would give one of several cellular carriers temporary access to a shared band dynamically for a fee [14, 56, 57]. This approach has commercial promise, and is already legal in a few countries [34].

E. Sharing Among Equal Regional Infrastructures

In Section 2, we defined spectrum sharing arrangements as sharing among equals or primary-secondary sharing, and as based on cooperation or coexistence, yielding four possibilities. Figure 1 presents examples of each of these four possibilities (along with the section above which discusses the model underlying this example in greater detail). Also, as noted in Section 2, there is sometimes confusion between the spectrum-sharing arrangement and the determination of a set of devices as either licensed or unlicensed. These policy decisions are certainly related, but it is important to separate them. The more effective choice depends in part on the applications involved. To demonstrate this, Figure 2 shows examples of primary-
secondary sharing we have considered in research where secondaries are licensed, and examples where secondaries are unlicensed.

<table>
<thead>
<tr>
<th>Coexistence</th>
<th>Sharing Among Equals</th>
<th>Primary-Secondary Sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All devices share a block of unlicensed spectrum. (Section 3.1)</td>
<td>Secondary devices use cognitive radio to opportunistically share spectrum with primary spectrum users. (Section 3.2)</td>
</tr>
<tr>
<td></td>
<td>Unlicensed secondary devices share spectrum with each other when and where it is not used by primary users. (Section 3.3)</td>
<td>Secondary devices use GPS and a database of transmitter locations to access spectrum where primary uses do not operate. (Section 3.2)</td>
</tr>
<tr>
<td></td>
<td>LMR public safety communications systems share spectrum through distributed trunking. (Section 3.4)</td>
<td>Secondary devices use ultrawideband technology to share spectrum with primary users. (Section 3.2)</td>
</tr>
<tr>
<td>Cooperation</td>
<td>Unlicensed devices all use prescribed common protocols and carry each other’s traffic in a cooperative commons managed by a regulator or license-holder. (Section 3.1)</td>
<td>Secondary devices explicitly request permission from a license-holder whenever they wish to transmit in a real-time secondary market. (Section 3.2)</td>
</tr>
<tr>
<td></td>
<td>Unlicensed secondary devices all communicate and cooperate to prevent interference to primary spectrum users and each other. (Section 3.3)</td>
<td>An interruptible system has exclusive rights to spectrum until or unless a primary user (such as public safety) temporarily preempts this system. (Section 3.2)</td>
</tr>
<tr>
<td></td>
<td>LMR public safety communications systems share spectrum through centralized trunking (Section 3.4)</td>
<td>One cellular carrier experiencing excessive call volume coordinates with another to briefly use the latter’s spectrum for a fee. (Section 3.4)</td>
</tr>
</tbody>
</table>

Figure 1: Examples of spectrum-sharing models of each type.
Coexistence between primary and secondary

<table>
<thead>
<tr>
<th>Secondary is unlicensed</th>
<th>Secondary is licensed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary system: Licensed TV broadcasters. Secondary systems: Opportunistic devices with no quality of service guarantees</td>
<td></td>
</tr>
<tr>
<td>Primary system: Licensed TV broadcasters Secondary system: Microcellular or cellular network which defers to primary, but does not share with other secondaries.</td>
<td></td>
</tr>
</tbody>
</table>

Cooperation between primary and secondary

<table>
<thead>
<tr>
<th>Secondary is unlicensed</th>
<th>Secondary is licensed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary system: Cellular Secondary systems: Devices that get temporary quality of service guarantees in a real-time secondary market</td>
<td></td>
</tr>
<tr>
<td>Primary system: Public safety Secondary system: Cellular network with exclusive but interruptible access to spectrum</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Examples of licensed and unlicensed secondary systems.

IV. REGULATOR OR LICENSE-HOLDER IN CONTROL

All of the sharing schemes described in Section 3 fail unless there is some entity that establishes the sharing rules. For example, how much interference are secondary devices allowed to impose on the primary system? Must secondary devices cooperate when sensing the primary, and if so, using what mechanisms? What etiquette must devices follow in a commons based on coexistence? Is each device required to sense other devices in the band and respond in some way, or is there simply a limit on maximum transmit power that does not depend on activity in the band? Traditional voluntary standards organizations have important roles to play, but they do not and should not have the authority to mandate requirements that are not voluntary; they were not constructed to be effective for that very different purpose.

Such decisions for a given band will be made either by the regulator, or by a license-holder to whom the regulator has granted sufficient authority. Traditionally, this has been the sole responsibility of the regulator. Some countries have now given license-holders sufficient latitude in some bands so that they too could play this role [34], but many countries have not. In return for managing the spectrum, the license-holder might charge a fee to all devices that use the spectrum. This could be a one-time fee when the device is first deployed, or the fee might be usage-based. Alternatively, the license-holder might be a manufacturer that will only allow its own devices to operate, thereby ensuring a degree of technical homogeneity.

Whether a regulator or license-holder is in charge has a significant impact on technology, especially for sharing arrangements that require highly detailed technical
rules of interaction. Such rules must be changed over time as technology evolves. Regulatory bodies are designed to make decisions in a slow and methodical process that is transparent to everyone, and to promote compromise among competing interests. Such processes are great at avoiding corruption and favoritism, but not well suited to reacting quickly to exploit a new technology, or to repair a newly visible interference problem or security vulnerability. Thus, where a regulator is in charge, all else being equal, there are more reasons to select a commons model or a primary-secondary sharing model based on coexistence.

A band where the license-holder is in charge is more conducive to complex cooperative approaches where homogeneity is important, such as a commons based on cooperation, or primary-secondary sharing based on cooperation. For example, a cellular carrier with a spectrum license that wants to allow secondary users (for a fee) can more easily choose between a coexistence model such as [58] or a cooperation model such as [32]. Similarly, the manufacturer of devices that could constitute the building blocks of a cooperative commons may get a nationwide license for its own devices, and those it deems compatible. This may limit opportunities for competitors, but it also reduces the dangers of incompatibility.

Many nations are making the transition from analog to digital television, and freeing up prime spectrum in the process. Perhaps we will see manufacturers or new kinds of service providers obtain a license for some of this spectrum, and move some of the more complex spectrum-sharing approaches from theory to practice.

Another difference between a regulator and a private license-holder is that the regulator has the ability to influence all spectrum, rather than just one small block. This too has technical implications. Some technologies under consideration are only viable if they can be deployed nationwide (or throughout multiple nations). For example, mobile consumer devices cannot easily be restricted to one portion of the country where spectrum is available. A regulator can more easily clear a spectrum band nationwide and ensure consistent rules throughout the band. A private entity can also provide consistent rules once it has a nationwide license, but getting a nationwide license for this purpose may be difficult without involvement from the regulator [59].

The most obvious difference between the regulator and license-holder is in objectives; regulators are supposed to make decisions to advance the long-term public interest, although they may not always succeed, whereas a private license-holder is likely to maximize its own profit. If the license-holder is the only entity that can provide spectrum for certain applications or devices, e.g. there is one and only one band where a license-holder has the authority to act as band manager or to provide a high-power fixed-wireless broadband service, then it may use its monopoly power in ways that harm the public interest, such as imposing excessive fees on certain applications [60]. However, this can be avoided if the license-holder faces sufficient competition, or at least the realistic fear that sufficient competition will emerge if it exploits its monopoly position. In this case, a profit-seeking license-holder will be
rewarded for granting access to as many potential spectrum users (customers) as possible, and meeting their technical requirements to the maximum extent possible. Thus, the license-holder should act to improve spectral efficiency wherever possible, at least over the time horizons considered by profit-seeking companies, which unfortunately may be shorter than the time horizons considered by regulators [14].

V. POLICY ENFORCEMENT

Emerging technology is making enforcement of spectrum policy more and more challenging. Any technology that assumes compliance with an unenforceable policy will never be viable. Thus, it may be necessary to alter designs to ensure that devices will comply with established rules and to demonstrate that compliance to regulators, even when doing so increases the complexity and cost of devices.

A number of the sharing models discussed in Section 3 involve unlicensed devices. After deployment, it is technically difficult to determine when some unlicensed device is violating the rules of the band in which it operates, and even more difficult to determine which is the offending device. Consequently, policies are enforced primarily by preventing the deployment of devices that are capable of violating these rules rather than detecting and punishing violators afterwards. To do this, regulators test and approve products before they can be widely distributed.

The technology behind software radios and some software-defined radios greatly complicates the problem of ensuring a priori that a device will not cause harmful interference to its neighbors. Although such concerns are further in the future, nanotechnology has the potential to create a similar challenge. The problem is that after a device has been certified as safe, its functionality can be changed with relative ease. A promising approach to address this problem is to create a limiter in hardware (e.g. [61]) or in the operating system (e.g. [62]) that cannot be altered simply by downloading new software. The regulator can then authorize a product based on the effectiveness of its limiter with little concern about the rest of the device. One can even make a limiter that imposes different policies depending on current location and frequency band. Moreover, it is possible to accommodate policies that change over time by requiring that the limiter be updated periodically [63]. If the limiter has not been updated sufficiently recently, it must prevent all transmissions until it can be updated.

The technology behind cognitive radios creates another set of complications, even for cognitive radios developed entirely in hardware. The more a device alters its behavior in accordance with what it senses from neighboring devices, or its current location, or other dynamic factors, the harder it is to exhaustively test that behavior. Thus, the risk of undetected faults is greater. Devices designed for a cooperative model will also be hard to test. For example, consider the possibility of secondary devices that cooperatively sense whether the primary system is using the spectrum. How can the regulator or license-holder be sure that the devices cannot cause harmful
interference if these devices fully cooperate? How can one be sure that such devices will fully cooperate?

It will be incumbent on the creators of novel systems that share spectrum to find innovative ways to make testing simple for the regulator or license-holder that must grant permission before a new device can use a given spectrum band. (Of course, this problem is easier if the designer of the device is also the entity that must give permission, i.e. the license-holder.)

The choice of sharing model and licensing regime also have indirect implications for enforcement. A licensed device with exclusive or primary rights to a block of spectrum has greater incentive to take action when neighboring devices violate spectrum policies than a device that shares spectrum with peers. With sharing among equals, the benefits of such action fall to everyone who is using the band appropriately, but none may want to assume the entire burden of reporting violations. This also demonstrates that this burden should be kept as small as possible.

VI. CONCLUSION

Scarcity of available spectrum is limiting the growth of wireless products and services. This scarcity comes largely from our use of outdated spectrum policies and wireless technologies. Traditional spectrum licensing prevents harmful interference at the cost of spectral efficiency. Unlicensed bands for low-power devices allow more sharing, but often rely on low utilization to limit the effects of mutual interference. New technology could make more forms of spectrum sharing practical, thereby substantially alleviating today’s spectrum scarcity, if and only if spectrum policy is reformed to match the technology. Conversely, new technology should be developed in part to address policy concerns if it is to have any impact.

Overly simplistic spectrum policy debates sometimes pit the perspective of operators of wide-area wireless systems (e.g. television, radio, commercial cellular, or public safety LMR) against the perspective of designers of localized wireless devices (e.g. local-area networks, point-to-point links, cordless phones). In an ideal (and dangerously unrealistic) world, the latter would like to deploy any technology they choose with minimal constraints. Thus, they generally prefer getting access to spectrum directly from a regulator (who charges no fees and imposes few constraints) rather than a license-holder. They envision either a completely uncongested unlicensed band where their device is primary, or a primary-secondary arrangement where their secondary device must contend with the primary system but no other secondary devices. However, the ability to deploy any technology you want comes with either the necessity to coexist with any technology others choose to deploy, or a mandate from the regulator that everyone will use your chosen technology, either of which tends to destroy the impracticable allure of the vision. There is also the problem that uncongested bands with easy entry can grow congested over time, especially if devices have incentive to adopt greedy designs. Thus,
it is not surprising to see an interest group fight for easy access to a block of spectrum initially, and then later fight to keep newcomers out.

On the other side, wide-area service providers enjoy the quality of service guarantees that come from being a primary license-holder. They would prefer to have unlimited flexibility within the bands they use, and no danger of interference. They have little interest in the creation of commons, and if secondary devices are allowed at all, they believe the owners of those devices should explicitly negotiate for access with the license-holder rather than the regulator. This arrangement works well for some applications, and poorly for others. (It is unworkable in TV white spaces, since by definition, there is no license-holder to grant access in the white space.) Moreover, this approach may fail to exploit cognitive radio’s ability to make use of spectrum that would otherwise sit idle.

A much wider range of policies and sharing arrangements are possible, and deserve serious consideration. It is pointless to look for the “best” form of sharing, as this is highly dependent on the types of applications involved. Thus, we should hope that different forms of sharing will emerge in different bands. Nevertheless, within a given band, adoption of one form of sharing often precludes other forms. Given that opportunities to redefine the policies of a given band occur infrequently, policy-makers must make difficult decisions.

This paper has argued that most sharing arrangements can be well described and categorized as based on coexistence or cooperation, and as sharing among equals or primary-secondary sharing. Within a given band, there may be multiple sharing arrangements of different kinds. For example, there may be primary and secondary systems that merely coexist with each other, while the secondary systems cooperate amongst themselves. For primary-secondary sharing, we must seek technical assurance that the primary system will never experience harmful interference, despite inevitable uncertainties over path loss, and hidden terminals. For sharing among equals, we must address whether technology gives designers and users the means to avoid serious interference and congestion problems, and whether policy combined with technology give them the motivation to do so.

Many factors must be considered when deciding to adopt a particular spectrum-sharing arrangement, including potential gains in spectral efficiency, and the ability to meet the needs of the most likely applications with respect to interference protection, congestion, support for mobility, and more. A less obvious factor that might preclude use of an otherwise attractive technology is the ability to ensure that spectrum policies can be enforced. It is therefore incumbent on developers of novel technology, especially software radio and cognitive radio, to include features that facilitate conclusive testing.

The importance of cognitive radio depends somewhat on which policies prevail, and conversely, the pace at which cognitive radio develops may influence policy decisions. Cognitive radio is already seen as a valuable tool for primary-secondary sharing based on coexistence. It may also be a valuable tool for sharing among equals.
based on coexistence, e.g. a coexistent commons, if policies are adopted that would motivate use of cognitive radio. Where sharing is based on cooperation, cognitive radio may or may not have as large a role to play.

In many countries, the transition to digital television will make prime spectrum available for other purposes, potentially under a novel spectrum policy that allows new forms of sharing. There may even be the possibility of spectrum bands where private license-holders set the rules rather than regulators. For example, a new band manager or a consortium of wireless device manufacturers could obtain a license, which might facilitate use of more complex rules of interaction.

VII. REFERENCES


www.ece.cmu.edu/~peha/policy.html


ABOUT THE AUTHOR

Jon M. Peha received the B.Sc. degree from Brown University, Providence, RI, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA. He is Chief Technologist of the U.S. Federal Communications Commission and a full Professor in the Department of Engineering and Public Policy and the Department of Electrical and Computer Engineering, Carnegie–Mellon University, Pittsburgh, PA, where he has served as Associate Director of the university’s Center for Wireless and Broadband Networking. He has been Chief Technical Officer of three high-tech startups and a member of Technical Staff with SRI International, AT&T Bell Laboratories, and Microsoft. He has also addressed telecom and e-commerce issues on legislative staff in the U.S. Congress and helped launch a U.S. government interagency program to assist developing countries with information infrastructure. He consults for industry and government agencies around the world. His research spans technical and policy issues of communications networks, including spectrum management, broadband Internet, wireless networks, video and voice over IP, communications for emergency responders, universal service, secure Internet payment systems, dissemination of copyrighted material, e-commerce, and network security. He is an *IEEE Fellow* and an *AAAS Fellow*. 