Emerging Technology and Spectrum Policy Reform

Jon M. Peha
Carnegie Mellon University

Abstract

Emerging technology, coupled with effective policy reform, could lead to tremendous gains in spectral efficiency. This would alleviate the spectrum scarcity that many nations have experienced. To be effective, reform in spectrum policy must fit the realistic capabilities of emerging technology, as well as applicable economic theory. This paper discusses three general approaches to reform. The first approach is increased reliance on market-based mechanisms, possibly leading to a spectrum property system. The second approach is the expanded use of commons, or blocks of spectrum that are available to all devices for sharing. It is argued that both approaches have significant merit that regulators should exploit. However, each of these approaches is ineffective when taken to its extreme, where one must make unfounded assumptions about technology. The third approach discussed in this paper is sharing between a primary spectrum user that is licensed, and one or more secondary users, where secondary users may not cause harmful interference to the primary spectrum user. Many people underestimate the potential importance of this third approach in today’s debates over spectrum policy reform. Emerging technology such as cognitive radio, location technology, and secure micropayment schemes will make a variety of primary-secondary sharing schemes ranging from real-time secondary markets to unlicensed opportunistic access more practical. Each of these schemes could be highly beneficial for a different set of applications and circumstances.

1 The author gratefully acknowledges CyLab for support of the security research described in this paper, and Intel for the spectrum-sharing research described in this paper.

2 Jon M. Peha, Associate Director of the Center for Wireless and Broadband Networking, and Professor of Electrical Engineering and Public Policy, Carnegie Mellon University, peha@cmu.edu, www.ece.cmu.edu/~peha

© 2007 Jon Peha
1 Introduction

The spectrum policies used in most nations have their roots in the applications and technologies of almost a century ago. One consequence of these policies is unnecessary spectrum scarcity. Even in populous cities where wireless devices are commonplace, measurements show that much of the premium spectrum sits idle at any given instant [1]. New technologies coupled with new approaches to spectrum management can alleviate this scarcity, and improve the availability and cost of many wireless applications. The question is how best to achieve this.

In some of the current debate on this issue, we hear that we must choose between a “technical” solution and an “economic” solution. This is completely untrue. Any approach that is not based on both the realistic capabilities of current technology and the real properties of economics and human behavior is doomed to failure. Sometimes, we hear that the there are two roads forward, one based on “property rights” and the other based on “spectrum commons.” This is also untrue, and the misconception is counterproductive. In reality, both concepts have value if appropriately defined, and both have their place in an effective spectrum policy. Each concept can be used to justify a set of valuable innovations in spectrum management. Each concept also becomes dangerously ineffective when taken to its extreme, where one must make assumptions that are inconsistent with current wireless technology.

Another danger about the typical debate between property and commons is that it obscures a very different set of reforms that hold tremendous promise: reforms that advance spectrum sharing between primary license-holders and secondary users. Emerging technology has made a wide variety of these arrangements possible and practical, and some or all of these arrangements could dramatically improve the problem of spectrum scarcity.

This paper will discuss the properties of spectrum property and spectrum commons in Sections 2 and 3, respectively, and how these properties relate to practical technology. It will discuss the many benefits of both, and the dangers of taking either approach to an extreme where the concept no longer fits the technology. Section 4 will discuss the technology and policy implications of sharing between primary and secondary users of spectrum. The paper is concluded in Section 5.

2 Spectrum Property

While spectrum resources are typically allocated by a central regulator, nations with market economies allocate other resources such as land by defining property rights, and allowing the free trade of property. Many of the market-based concepts underlying a property system can be used to good effect in spectrum management, but as we will discuss, there are important technical differences between land and spectrum. Because of
these differences, the rights granted to users of spectrum should not be as far-reaching as those granted to users of land.

There is no consensus on exactly what rights befall a property owner when the property is spectrum, and the ambiguity over definitions is a source of confusion [2]. One general definition of property is the right to hold, subdivide, transfer, use, and admit or exclude others from using a given item [3]. This section will argue that some but not all of these rights are appropriate for spectrum.

In a market economy, land typically goes to those who value it most and are willing to pay for it. This occurs because land-owners can subdivide their land in any way that increases its value, and keep, sell, or rent every part of that land. A similar phenomenon could occur with spectrum, where spectrum users assemble licenses that cover the geography, frequency range, and time period that they value most, and no more. In addition, property owners have incentive to use their property efficiently, because they derive all the benefit. The same can apply when spectrum users have exclusive access. Thus, there is an argument that license-holders should be able to subdivide, transfer, use, and admit or exclude others from using a block of spectrum. Similarly, they can use market-based mechanisms to get spectrum licenses from regulators, i.e. open spectrum auctions.

Some believe that, by definition, property rights also include the flexibility to use spectrum in any way the license-holder (or property-owner) wishes, without interference from a regulator, and that by definition, property rights can never expire. These two characteristics are related; if property rights never expire, then the regulator cannot wait for the expiration of a license to change how a given block of spectrum can be used. Thus, it is necessary to give the license-holder complete flexibility to make that decision. The value of flexibility is ambiguous. On one hand, flexibility can greatly enhance the value derived from spectrum [5]. It allows a license-holder to use spectrum for the most valued application, or if other market mechanisms are in place, to transfer the spectrum to someone who will do so. Flexibility also allows a given frequency range to be used for different purposes in different parts of the country, which is important because the needs of rural and urban areas can differ greatly. Flexibility in property rights for land makes land more valuable, and expanding the flexibility associated with some spectrum licenses could similarly make that spectrum more valuable.

However, despite the appeal of the analogy between spectrum and land, there are significant differences, some of which relate to flexibility. In some cases, flexibility in spectrum use can come at a significant cost. There is value to regulator-imposed uniformity [5]. If all television stations use the same technical standard and operate in the same frequency band, then consumers can move anywhere and their televisions will still work. If television had emerged in an era where complete flexibility prevailed, it is possible that incompatible standards or frequency ranges would have emerged in different regions. This problem is well predicted by economic theory. Following such a standard yields a positive externality, and this can lead to market failure in a property

---

3 For example, see the seminal work of De Vany et al [4].
scheme unless a regulator has the authority to act. Another disadvantage of flexibility, if it is taken too far, is that interference levels must be sufficiently low that they do not interfere with anything a license-holder is allowed to do - regardless of what the license-holder is actually doing. Thus, increasing the license-holder’s flexibility also decreases the discretion of a regulator to adapt to new needs and new technologies. For example, the US Federal Communications Commission decided to allow ultrawideband devices to operate between 3.1 and 10.6 GHz [6]. This useful step would not have been possible if any incumbent license-holder in that range had complete flexibility and property rights that never expire, allowing this license-holder to successfully argue that ultrawideband would infringe on its rights. The problem is that people defining property rights for an earlier generation of technology could not have foreseen the emergence of ultrawideband. In some sense, technical evolution makes it necessary to periodically redefine how to draw boundaries around spectrum “property,” which is one of the ways that spectrum and land differ. For these reasons, regulators should not consider making spectrum rights permanent. Licenses must expire, so that regulators have the opportunity to introduce change.

Another potential cause of market failure in a property scheme is the difficulty and transaction costs associated with assembling many licenses into a large block of spectrum that covers a region [7]. Despite the technical progress we have made to date in software radios, spectrum is not fungible as we might wish. In practice, there are applications that will only be cost-effective with a block of spectrum that is contiguous across a large frequency range, and/or contiguous across a large geographic region (such as a nation). If even one license-holder in the target region refuses to sell, progress may be impossible. In such instances, a regulator is needed. Again, a regulator cannot play this important role if property rights are indefinite, so licenses never expire.

### 3 Spectrum Commons

In any commons model, spectrum is shared, and no one is given special priority. Sharing can come in different forms. Devices might cooperate or they might merely coexist. While both possibilities are sometimes lumped together under the ambiguous heading of “commons,” the two are entirely different [2]. (Unfortunately, this confusion over definitions has made much of the debate about spectrum commons meaningless, as advocates and opponents implicitly use incompatible definitions.) The coexistence model exists today in many nations, and has spawned successful products such as Wifi and cordless phones. When systems merely coexist, explicit communications is pointless; a cordless phone and a Wifi card do not decode each other’s transmissions (although one might try to sense when the other is transmitting for simple collision avoidance). In contrast, with cooperative sharing, devices must communicate with a common protocol, and work together. For example, all devices could self-organize to form one ad hoc network. This section will address the two flavors of commons separately.
Recent allocations of unlicensed spectrum using the coexistence model have spurred tremendous innovation and productivity, as best demonstrated by the rise of wireless LANs. For example, in 1993, Carnegie Mellon University (CMU) began development of an experimental wireless system [9] designed to blanket campus with broadband coverage, both indoors and out. CMU used a precursor of what later became the IEEE 802.11 standard [8]. This system has become an important part of campus infrastructure, and organizations around the world now have similar systems. In practice, CMU could never have developed this system if unlicensed spectrum was not available. Perhaps CMU could have obtained a license, giving us exclusive access to a block of spectrum throughout the neighborhood, but this would have been exorbitantly expensive, and deservedly so. Exclusive access would be an incredibly inefficient use of spectrum. Computer communications are highly bursty, and some collisions are tolerable, so it makes sense for CMU, the University of Pittsburgh, several large hospitals, and many small business and individuals in our neighborhood to share the same block of spectrum. Alternatively, CMU could have tried to get highly localized site licenses for all transmitters, and coordinated their locations with our neighbors and/or with the US Federal Communications Commission. However, this might require the university to contact the regulator every time one of our 800 transmitters is deployed or moved. The transaction costs of explicit coordination could exceed the value of the system. Finally, CMU could have called a licensed wireless service provider, who might have offered us a carrier-based 3G-like service that was more expensive, less flexible, less useful, and less spectrally efficient for our particular application. This is just one example of a system that can flourish in unlicensed spectrum, but would probably fail if a license were required, with or without market-based mechanisms.

As demonstrated by the example above, unlicensed spectrum has many advantages. It requires spectrum sharing, which can lead to vastly superior spectral efficiency than exclusive access, where spectrum often sits idle because the license-holder is not transmitting. Unlicensed spectrum is necessary to support mobile systems, such as a group of laptops that form an ad hoc wireless local-area network wherever they happen to be. It is useful for inexpensive low-power consumer products such as cordless phones, where the cost of coordination and licensing would unnecessarily dominate system cost and the interference impact on neighbors is small.

When releasing unlicensed spectrum, regulators must guard against two related sources of inefficiency. One is that unlicensed spectrum will attract applications that would operate more effectively and efficiently in licensed spectrum. The other is that engineers will design “greedy” devices, i.e. those that transmit with greater power, duration, or bandwidth than necessary, because they have little incentive to conserve spectrum that is shared. In the extreme, greedy devices can lead to a tragedy of the commons, where many devices are greedy, and all devices in the band experience inadequate performance as a result [10]. Both of these dangers can be addressed by establishing appropriate technical rules to govern the unlicensed bands, possibly (but not necessarily) influencing power levels, modulation, back-off schemes, etc. At minimum, these rules will prohibit a device from transmitting at high power for extended periods without interruption. 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 even at high utilization levels into the technical rules of operation, i.e. the etiquette [10, 11].

As with the introduction of the market principles discussed in Section 2, there is much to be gained through a commons based on coexistence, but the approach should not be taken to its extreme. Unlicensed spectrum is not a replacement for licensed spectrum, any more than public parks are a replacement for private homes. Unlicensed bands are more efficient and appropriate for some applications, such as those discussed above. Licensed bands are more appropriate for other applications, such as broadcast TV or public safety communications, for which quality of service should be guaranteed.

The characteristics of a commons based on cooperation are quite different [12, 13]. In this commons, all devices cooperate, even though they serve different owners. 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. 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 may not fear oversubscription the way users of a commons based on coexistence do. The potential advantages of a system with cooperative gain are enormous, and these systems deserve serious consideration. However, compared to the commons based on coexistence, this is a relatively immature technology. There are significant challenges ahead.

At Carnegie Mellon University, we are conducting research on issues associated with security and selfishness in a cooperative commons. When devices 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. A selfish node may not cooperate fully, and this can lead to problems. Even worse, a malicious node may take deliberate steps to disrupt the network. We have found that many current protocols do not meet the unique security needs of cooperative networks. We are seeking potential solutions. However, for now, there are still open research issues, and regulators must consider this.

Another challenge that is unique to the cooperative commons is that all devices must share a detailed communications protocol. To foster cooperation, devices should not be deployed in the band unless they can communicate using this protocol. Who will specify this protocol, and make changes to it over time as technology evolves? If the cooperative commons resides in an unlicensed band, then this responsibility falls to the regulator. There is certainly precedent for regulatory control over a standard, but standards imposed by a regulator today tend to change very slowly, such as the standard for FM radio. With a cooperative commons, regulators may have to move at a pace that is more typical of open standards organizations such as the Internet Engineering Task
Force [14] or the IEEE 802.11 Working Group [8], which would be a challenge for most regulatory bodies.

With either coexistence or cooperation, a “spectrum commons” could be created by a license-holder instead of the regulator. Rather than using unlicensed spectrum, a private entity might obtain a license, establish its own operating rules, and allow devices to operate in its spectrum [5]. The latter approach is particularly appropriate for a cooperative system, because it eliminates the latter challenge described above. For example, devices might operate in a band licensed to an equipment manufacturer, and this manufacturer would make sure all its devices share an effective protocol for cooperation. Still, the regulator has an important role to play. This approach probably requires a single nationwide license of appropriate bandwidth, and sufficient technical flexibility. As discussed in Section 2, such a band may never emerge without deliberate assistance from the regulator.

4 Primary-Secondary Sharing

Applications that need guaranteed quality of service are given exclusive access to spectrum through some form of licensing. These exclusive allocations also insure that spectrum will not be fully utilized, as there are generally times and/or locations where other devices could transmit in a given block of spectrum without causing harmful interference. Thus, spectral efficiency can be greatly improved through primary-secondary sharing, where one system has primary usage rights that make quality of service guarantees possible, and one or more secondary systems operate without causing harmful interference to the primary. Emerging technology offers many new ways to do this.

As with a commons, primary-secondary sharing can take one of two forms: cooperation and coexistence [2]. Cooperation means there is explicit communications and coordination between primary and secondary systems, and coexistence means there is none.

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. A spectrum-user may get permission to operate as a secondary from the regulator, in which case the regulator must establish rules that prevent harmful interference to the primary. For example, unlicensed devices might be allowed to operate in a licensed band provided that they obey these rules. Alternatively, permission may come from the license-holder, in accordance with rules established by the license-holder. To protect the primary, secondary devices can either transmit at such low power that they never cause harmful interference to the primary, as with ultrawideband [6], or they must transmit opportunistically when and only when they
sense the environment and determine that transmissions will not cause harmful interference [15].

Thanks to cognitive radio, global positioning systems (GPS), sensor networks, and other emerging technologies, opportunistic access is becoming more practical, but significant research challenges remain. The extent of the challenges depends on the nature of the applications. The first practical deployments will probably occur where the primary application has fixed rather than mobile transmitters, and the modulation schemes are well known, as with broadcast television or radio. Moreover, because quality of service guarantees are generally not possible for the secondary device, the role of secondary spectrum user under this arrangement works well for some applications and not for others.

When sharing is instead based on cooperation, the primary and secondary 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 a distinct 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 [16]. There has been considerable discussion about creating a secondary market where these blocks of spectrum can be “rented” out for months or years [17]. It is important to note that more dynamic forms of sharing, where spectrum is given out for minutes or milliseconds, are also possible, and should not be precluded by regulation.

As with coexistence, the practicality of this approach depends on the applications involved, and other factors. For example, the primary system needs a component to act as gatekeeper, which is much easier when the primary system is a cellular system rather than a broadcaster. We have analyzed scenarios where extensive communications among secondaries is possible with little impact on the primary [16]. This was facilitated by a variety of technologies, including location technology which was used to enhance frequency reuse, and secure payment technology so that primary systems can receive payments from previously unknown secondary devices.

It should be noted that a secondary spectrum user could also be licensed. Both licensed and unlicensed secondaries are precluded from causing harmful interference to the primary. The 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. Once again, this arrangement can work with or without cooperation. With cooperation, the secondary system operates much like any other licensed system when the primary is not active. For example, in an interruptible system, the primary signals for the secondary to cease all transmissions when the primary needs the spectrum. This might be useful for a public safety communications system, which could increase capacity by claiming the shared spectrum during a serious emergency, but would otherwise leave the spectrum to secondary users, as discussed in [18]. Without cooperation, a licensed secondary might
operate in white spaces, i.e. geographic regions where the primary systems are not operating, in guard bands, or opportunistically. We have analyzed scenarios where extensive communications among secondaries is possible with little impact on the primary, sometimes because the primary’s needs are sporadic as with public safety communications, and sometimes because there is white space between broadcast towers to exploit.

As shown above, there are a variety of primary-secondary arrangements that are becoming more practical with emerging technology, but practicality depends on circumstances, and on applications. No model is best in all cases. Figure 1 shows examples of different kinds of primary-secondary schemes that we have considered.

<table>
<thead>
<tr>
<th>No coordination between primary and secondary</th>
<th>Secondary is unlicensed</th>
<th>Secondary is licensed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unlicensed underlay. e.g.</td>
<td>Licensed secondary with exclusive access in white space or guard bands, e.g.</td>
<td></td>
</tr>
<tr>
<td>Primary system: Broadcasters with site licenses.</td>
<td>Primary system: Broadcasters</td>
<td></td>
</tr>
<tr>
<td>Secondary systems: Opportunistic devices without quality of service guarantees</td>
<td>Secondary system: Microcellular or cellular network</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coordination between primary and secondary</th>
<th>Real-time secondary market, e.g.</th>
<th>Secondary with exclusive access but interruptible access, e.g.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary system: Cellular</td>
<td>Primary system: Public safety</td>
<td></td>
</tr>
<tr>
<td>Secondary systems: Devices with temporary quality of service guarantees</td>
<td>Secondary system: Cellular network</td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Examples of primary-secondary models.

5 Conclusions

Emerging technology has provided an opportunity to significantly alleviate spectrum scarcity, but only if spectrum policies are reformed in ways that fully exploit the new technology. These spectrum policies must also be built on sound economic principles, as well as a realistic assessment of what the technology can do.

Advocates of “property” rights are correct that market-based mechanisms can improve both technical and economic efficiency of spectrum. By employing spectrum auctions, and by making it easier to transfer licenses, or to subdivide licensed spectrum
and make a portion of it available to others, regulators can make spectrum available to those who value it the most, in the amount they value the most. In some cases, granting license-holders more flexibility will further allow spectrum to be used for the application in greatest demand, although flexibility is not always beneficial. Nevertheless, there are technical differences between spectrum and other goods that are exchanged in an open market, and as a result, there are sound reasons not to take the property approach too far. We will need regulators who can change the way spectrum is used as technology changes. For example, in response to new technology, a regulator might establish new rules for spectrum-sharing, or make sure that spectrum is available in large contiguous blocks. Thus, spectrum licenses must be temporary, so the regulator can change rules after a license expires.

There is also merit in the arguments for shared spectrum commons. There are actually two kinds of commons with vastly different properties: those based on coexistence and those based on cooperation. We have seen great commercial successes in the commons based on coexistence over the last decade, and there is reason to hope for more successes in the future. This approach to spectrum management can be quite effective for applications that were not well served under traditional licensing. This includes cases where there are large numbers of low-powered devices, or where entire wireless systems are portable, such as a wireless PBX that is moved from one site to another, or where best-effort service is adequate. However, this commons approach also cannot be taken to its extreme, as there are applications that require guaranteed quality of service, and are better served through licensing.

A very different kind of commons is based on cooperation rather than coexistence. Cooperation makes it possible to achieve much greater efficiencies. While research continues, and some pilots are underway, this is a comparatively new approach. There are significant technical challenges ahead, notably in the area of security.

Some of the most compelling new opportunities for improving spectral efficiency involve sharing between a primary license-holder, and one or more secondary systems that are not allowed to cause harmful interference to the primary user. Many such models are possible. Secondary systems may be licensed or unlicensed. They may get permission to operate from the regulator, or from the license-holder in exchange for payment. They may cooperate with the primary, or coexist in a manner that is invisible to the primary. Once again, different spectrum-sharing models are more effective for different kinds of applications and circumstances, so all models have their uses.

Overall, a wide variety of models for spectrum use are becoming more practical. Different models are more appropriate for different applications. Rather than try to find the “best” approach, regulators should provide a variety of options to those who design and use wireless devices.
6 References


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