The Unfinished History of Usage Rights for Spectrum

Martin Cave[1] and William Webb[2]

PART ONE: INTRODUCTION AND OVERALL APPROACH

Introduction

Until recently there was relatively little study of “usage rights” for radio spectrum. Licence holders were typically given a maximum transmit power up to which they could operate specified apparatus at a specified location for a specified purpose, together with an undertaking that they would not suffer “harmful interference” – a term that was generally ill-defined. As spectrum has become more densely used and as types of usage have changed the shortcomings of such an approach have become increasingly clear, for example in the interference caused between Nextel and public safety users in the US and in other on-going disputes.

This led to a number of different suggestions as to better ways that the rights of licence holders could be defined, including novel approaches in Australia, New Zealand and the UK. For example, the UK has pioneered spectrum usage rights (SURs).[3] Here the rights of a licence holder are defined in terms of the maximum levels of interference they can cause others rather than the maximum power levels that are allowed to transmit. By reciprocity licence holder can also determine the interference that they can expect from their neighbours. While more complex than simple transmit power levels, SURs can be shown to provide a more robust conceptual framework for clearly defining the rights of licence holders and one that allows the market to find the optimal levels of interference between neighbours. SURs lend themselves best to licences with single national licence holders – as the number of licence holders in a band increases it becomes progressively more difficult to divide the interference allowance between them.

Conversely, in the area of unlicensed or licence-exempt spectrum there has been very little innovation in the definition of rights. Unlicensed users were generally assumed to have no explicit protection from interference and hence very little in the way of defined rights. Instead, work in this area has concentrated on means to enable fair and efficient sharing of the resource among those contending for it at any given time. Concepts developed in this space include dividing users into classes depending on the amount of interference they caused and requiring manufacturers to adopt fairness protocols.

In this paper we first outline the role of usage rights - a subset or close cousin of the more general notion of property rights – and consider the implication for an efficient allocation of resources of having a clear system of such rights. We also define the relevant dimensions of such rights. On this basis we are able to review the forms of usage rights which have been deployed to date in apparatus licensing, existing formulation of spectrum usage rights, and, by way of contrast, unlicensed spectrum

[1] Professor, London School of Economics
[2] Professor, Surrey University, and Director of Technology Resources, Ofcom
[3] See http://stakeholders.ofcom.org.uk/binaries/spectrum/spectrum-policy-area/spectrum-management/spectrum-usage-rights/sursguide.pdf

access or commons; we also look at more recent rights established in relation to ultra wide band and white spaces.

Looking ahead, we briefly examine innovative forms of spectrum use which are consistent with alternative formulations of spectrum usage rights. We then identify a set of ‘promising’ reformulations of rights which are consistent with technical developments, and discuss some of the technical problems of implementing them. Finally we consider the division of labour between spectrum regulators and spectrum users in introducing new forms of rights.

The nature and implications of usage rights

Usage rights are a subset of property rights, which have been defined as consisting of four elements[4]:

- the right to use the good or asset
- the right to earn income from it
- the right to transfer it to others
- the right to enforcement of property rights

The spectrum management regime affects all of these. Assignment of spectrum licences confers (usually constrained) rights to use, which in respect of commercial users permit use of the spectrum to generate revenue. This need not entail a right to transfer the usage right to others, by gift, lease or outright sale. Finally, the interference management regime addresses the issue of enforcement of rights. In this article, we are concerned with the set of issues relating to how the usage rights are defined; how they can be protected; and the ability of licensees/holders of usage rights to enforce those rights via an interference management regime.

The link between property and usage rights and efficiency is captured in the Coase theorem.[5] This states that if property rights are properly specified, and if transactions costs are absent, private bargaining among the parties can produce an efficient outcome, even in the presence of externalities. Essentially what we need to do is translate the right to interfere with others or alternatively to be free from interference into a property right; the parties can then buy and sell this property to produce an efficient configuration of use of assets.

The key externality in spectrum use is interference, both within band and out of band. In relation to this externality, the Coase theorem envisages that if there is rivalry between two users of spectrum as a result of interference, then the parties will be able to maximise the proceeds from the spectrum by exchanging use rights between them, either for money or by barter. Provided the process starts with them having clearly specified non-overlapping rights, the bargaining or exchange process goes on until the total revenue product of the spectrum is maximised.

Thus if the conditions of the theorem were exactly satisfied, the spectrum management problem would disappear. Parties would exchange usage rights with frictionless efficiency in a wholly decentralised fashion, provide the process started

[4] On property rights in general, see D Lueck and TJ Miceli, ‘Property law’, In A Mitchell Polinsky and Steven Shavell, Handbook of Law and Economics, 2007,pp 183-258.
[5] R H Coase, ‘The problem of social cost’ Journal of Law and Economics, 1960, pp 1-44. The publication of this ‘generalising’ article followed hot on the heels of Coase’s famous examination of the prospects for creating a spectrum market: R H Coase, ‘The Federal Communications Commission’, Journal of Law and Economics, 1959, pp. 1-40.

with an unambiguous specification of rights, and no ambiguity crept in afterwards. Even in the world of apparatus licensing, some bargains of this kind take place at the edge. For example, two broadcasters might agree on small adjustments to transmitter power, to be approved by the spectrum regulator, if doing so would improve the financial position of each.

The point about this example is that a successful outcome is relatively easy to accomplish. The adjustment is marginal and only two beneficiaries are involved. The latter factor bears on the degree of transactions costs – the absence of which is a condition of the theorem. Such transactions costs might arise for several reasons:

- if the value of the putative transaction were small, the managerial costs of decision and the legal costs of transacting might rule out an agreement; whereas, conceivably, a regulated outcome might be cheaper;

- if the parties involved are numerous, the problem of co-ordinating the bargaining process would be difficult to solve; also, if payments were required, some beneficiaries would try to avoid paying, and become so-called 'free-riders'. Imagine trying to achieve by wholly decentralised bargaining, unmediated by national spectrum regulators, the allocation and assignment of DTT broadcasting rights determined for ITU Region 1 at the Geneva conference of 2006. The problem is exacerbated by the complex interference effects associated with spectrum use, which mean that some parties do not know they will be affected until transmission begins.

- if the 'gains from trade' are large, and the number of parties small, each party will devote considerable effort to gaining as large a share as possible for itself; this may lead to a failure to agree, in which case each party makes no gain. Experienced negotiators should be able to avoid this outcome.

One can, of course, try to structure the initial allocation of rights to be as close as possible to an estimated optimum. One can also err in the direction of assigning rights *against* any party which seems best placed either to prevent the interference in the first place or to pay others to control it if it is desirable so to do[6]. However, the extraordinary pace of spectrum-using technologies militates against the achievement of long-run solutions in this fashion. Such developments also raise new issues and new dimensions of spectrum use. A good example is the determination of which body has usage rights in relation to spectrum used for mobile communications from an aeroplane in flight over a country. Ultra wideband and white spaces (discussed below) create new issues.

In practice, spectrum regulators have historically taken a highly activist approach to the management of usage rights. Trading has been confined to relatively few countries. Other forms of decentralised bargaining have not played a large role. Thus it is spectrum regulators which typically set the size of guard bands, rather than negotiation.

Is this inevitable? Liability rules rather than regulation might be more efficient in a static sense in coping with interference, and might also be able to react more quickly to new technological options. The task of, say, organising a database to permit shared use of spectrum may better be done co-operatively by users or by a band

---

[6] The former of these two characters is known in the literature as the 'least cost avoider'; the latter as the 'best briber'. See G Calabresi and A D Melamed, 'Property rules, liability rules and inalienability: one view of the cathedral', Harvard Law Review, 1972, pp. 1089-1129.

manager. We return to this question below, having set out a general model of usage rights in the context of sharing and interference.

A generalised model for usage rights

Any given spectrum band might have multiple users and as technology gets better at sharing access we might expect the number of bands with multiple users to increase[7]. Shared bands could include a number of licensed users and a number of unlicensed users.

- *Licensed users*: If there is more than one licensed user then they will have a defined way to share access such as by geography, time, polarisation, directionality or other. There may be some hierarchy of access, for example one user may have no restrictions on their deployment whereas other licensed users may have to tailor their deployment to avoid interfering with the primary user. If there are no licensed users then the band is unlicensed.
- *Unlicensed users*: Unlicensed users will have some restrictions designed to prevent interference with the licensed user. These may be power levels (eg UWB), geography (eg cognitive access via a database) or via sensing (eg radar avoidance in 5GHz WiFi). There may be further restrictions to reduce interference between unlicensed users such as “politeness protocols”. If there are no unlicensed users then the band is a “classically” licensed band.

The introduction of these different users can be managed by the regulator or the primary licence holder. There may be some situations in which secondary licence holders can also introduce other secondary or unlicensed users.

Hence, in designing rights for spectrum bands we need to consider the potential for a mix of licensed and unlicensed users.

The problem of defining these rights can be divided into two parts:

1. The rights for the band as a whole as defined by the sum of all the emissions within the band which need to be defined so that neighbouring users have appropriate certainty.
2. The division of these rights to the different users within the band.

The first of these has been studied in some detail in relation to licensed users. For example, in the UK a form of spectrum usage rights (SURs) has been defined and implemented which meets these criteria. These are introduced briefly below before we turn to the second part of the problem of rights definition.

Defining overall rights for the band: An introduction to SURs

An SUR is a licence which sets out the interference that a licence holder is allowed to cause in terms of signal strength as experienced by a receiver. Signal strength can be measured in terms of power flux density (PFD) which expresses the power at a certain point in terms of watts/m^2/MHz.

[7] This view was set out very presciently by Eli Noam in ’Spectrum auctions - yesterday’s heresy, today’s orthodoxy, tomorrow’s anachronism; taking the next step to open spectrum access’. Journal of Law and Economics, 1998, pp 765-780.

There are broadly three ways that one licence holder (let us term them "A") can interfere with another ("B"). Each of these mechanisms needs to be controlled by licensing.

1. *Geographical interference*. In this case both A and B are using the same frequencies but in different locations. If A moves too close to B then signals from A's transmitters can interfere with reception on the edge of B's coverage area.
2. *Out-of-band interference*. In this case, A and B are located in the same geographical area, using separate but nearby frequencies. If A's transmissions "spill out" into neighbouring bands then they can be received by B's receivers as interference.
3. *In-band interference*. Again, A and B are located in the same area with nearby frequencies. In this case, B's receivers are not perfect and also pick up some of the signal A transmits in its own bands causing interference.

The PFD term needs to be applied to each of these three different forms of interference.

- For geographical interference, the signal level generated, specified in terms of power flux density (PFD), at or beyond the geographical boundary should not exceed a set power level.
- For out of band interference the PFD measured at an agreed height above ground level should not exceed a set power level at more than a certain percentage of locations in a set area. The reason for the percentage of locations is that if just one, or a few, measurements were made they might be close to a spectrum neighbour's base station where the signal level received would be very high. Enough measurements are needed to average over a representative area.
- In-band interference can be specified and measured in an identical manner to out-of-band interference. The only difference is that the allowed PFD level would be higher, reflecting the fact that spectrum users are generally allowed to transmit much higher power levels within their bands than outside them.

This, in outline, is what comprises an SUR – a set of three PFD limits corresponding to the three types of interference. The complexity arises in firstly deciding what limits to set and secondly in verifying that they have not been exceeded. A guideline document setting out how all this can be achieved is available[8].

PART TWO: DEFINING RIGHTS WITHIN A BAND

General scope of the problem

With multiple users within a band, the levels of emissions for each need to be defined in some way such that each user has an optimal share of the total, and that the total does not exceed the SUR for the band. This might include defining:

- Interference levels between licensed users.
- Interference levels between unlicensed users.
- Interference levels from between licensed and unlicensed users.

[8] See http://stakeholders.ofcom.org.uk/binaries/spectrum/spectrum-policy-area/spectrum-management/spectrum-usage-rights/sursguide.pdf

The interference levels between licensed users can be handled within the existing SUR structure (indeed, SURs envisage licensed users negotiating with each other to optimise the interference they are allowed to generate). However, the other two types of interference require difference constructs. Each of these might need to be defined and policed in a different manner due to the differing levels of control and monitoring that are possible with licensed and unlicensed usage. In particular, the level of interference that should be allowed between unlicensed and licensed users might best be defined not in terms of one that generates minimal interference but in terms of one that optimises the economic value of the band.

The mechanisms to achieve these constraints depend heavily on whether there is knowledge of the unlicensed users and an ability to control them. In most cases unlicensed users are uncontrolled – there is no way of knowing how many there are in any given location at any given time, and no way of controlling their actions as a result. Hence, defining the interference levels that they can cause can only be based on scenarios of deployment which may prove to be inaccurate. Alternatively, if there is some mechanism to monitor and control unlicensed users such as through a database access approach, then there it becomes possible to control interference levels to some degree. The next section discusses two different technical approaches to unlicensed access of bands with licensed users that might be adopted in future.

Emerging technologies that allow sharing between licensed and unlicensed users

Recent technical developments have led to the possibility of sharing between licensed and unlicensed users. Principally these include ultra wideband (UWB) and cognitive or white space access (we will use the term “cognitive” here). Each of these is briefly introduced below prior to a discussion of the implications for licensing.

UWB makes use of the trade-off that exists between bandwidth and power. Essentially, a given range can be achieved with less power if more bandwidth is employed. UWB takes this to an extreme with ultra-low power levels spread across very broad bands of spectrum to enable data rates of many Mbits/s but with a range that will typically only be within a room. The premise of UWB was that although the signal would span multiple licensed bands that the interference generated as a result to licensed users would be so low as to be unnoticeable. However, with no clear definition of “unnoticeable” or “harmful” discussions around appropriate conditions of access were complex and protracted – an issue we will return to later. UWB is sometimes generically referred to as an “underlay” – a signal that sits under the licensed use of the spectrum.

Cognitive devices make use of the gaps in transmissions from licensed users. These gaps may be geographical – areas where the licensed user is not operating, or temporal – times when the licensed user is not operating. The devices then transmit at power levels that may be in line with those of the licensed user as long as it is clear that doing so will not cause interference. As with UWB, defining acceptable interference levels has proven contentious. An important point for our development of rights is that proponents of cognitive access appear to be assuming a database model where the cognitive device determines its location and consults a database to determine which spectrum is free in that locality. Cognitive access is sometimes termed “overlay” although perhaps “inter-lay” would be a more accurate terminology.

In both cases of UWB and cognitive the assumption has been that access should not materially impact the licensed user. In a situation where the licensed user is already operating this may be the only pragmatic way to proceed. However, if a green-field licence were to be awarded it may be that the usage rights could be structured so as to optimise a mixed model of usage including licensed and unlicensed use. As well as considering the interference that the unlicensed users can generate to the licensed user it might be important to consider the reverse as well.

We will be taking these factors into account as we define our rights model – in the next section we assume that interference can be controlled and determine what the level should be while in subsequent sections we explore the implications of uncertainty in the level.

Setting the optimal level of interference between licensed and unlicensed use

Varying the level of interference that the unlicensed users can cause has four key effects – for example if it were increased it would:

1. Increase the overall interference generated in the band with the implication that the licensed user can generate less interference if the total is to stay within the SUR limits.
2. Increase the utility to the unlicensed user since they can now transmit further or at higher data rates. This might translate into a need for less infrastructure.
3. Decrease the utility to the licensed user since their system will now suffer more interference, reducing the range or data rate. This might translate into a need for more infrastructure.
4. As a result of (1) and (3) the licensed users will also generate less interference towards the unlicensed users, increasing the unlicensed user's utility from the band in something of a circularity.

The effect of changing the interference level may be very direct, for example in the case of UWB where a higher interference level would directly translate to a greater range. Or it may be less clear cut, for example in the case of cognitive, where the devices try to avoid interference and hence increased interference levels may only modify usage around the periphery of the white spaces.

How could an optimum be found? The Coasian point is that the two sides could bargain (assuming that unlicensed users could form a coalition; otherwise it would be an n-person negotiation). For example, a generous allocation of rights could be given to the commons, and the individual licensee could then buy then out. There is a further market solution in which the band is given in its entirety to one party, which then chooses an allocation which maximises its revenue. Think, for example of a band which can house a primary user and an overlay. Provided the overlay can efficiently be monetised, the holder of the rights will have an incentive to find the optimal allocation.

With UWB the regulator imposed a level of interference that UWB could cause based on calculations of transmitter powers that would not cause any noticeable interference to licence holders. Hence, rather than finding an optimal solution, the "new entrant" unlicensed use had to fit in around the existing licensed usage. The situation with cognitive is still unfolding but appears similar in that cognitive devices are not allowed to generate any significant interference to the licence holders. Given the difficultly of changing licenses that have already been issued perhaps this cautious approach is not surprising but it does appear to have contributed, perhaps significantly, to the failure to date of UWB because of the very low power levels allowed.

Could a market solution have worked here? The key issue is whether the licensed use already existed. If the bands were vacant then it would have been possible to set any level for UWB transmissions and reflect this in any licenses issued. If, for example, these licenses were auctioned then the auction price should reflect the cost of the interference to the licence holders. A "generous" allocation could then be made to UWB and in principle the licence holders could have negotiated with the UWB users, or perhaps more practically, the manufacturers and associated standard bodies, to modify this if appropriate. However, the complexities of negotiating with multiple parties remain and might act as an obstacle. If the licences already exist then the regulator will be inclined towards a level that is generous to the licence holders. In this case, with a nascent technology and industry the chance of multiple potential UWB users negotiating a raised limit with potentially multiple licence holders seems slim.

With cognitive the situation is less clear cut. The impact of changing the interference allowed changes the amount of spectrum that becomes "white space". If there is plentiful white space then the impact of allowing more interference may be small whereas if there is very little the impact could be great. Interference control can be fine grained and rapid through changing parameters in the database. Nevertheless, as with UWB, allocations that are more or less generous to the licensed user are possible with the same practicalities applying to existing licensed users. It is possible that the medium of the database might be a mechanism to make negotiations between multiple parties simpler since it effectively provides a communications path between them.

Setting the optimal levels between unlicensed users

As discussed above, the setting of optimal interference levels between different unlicensed users is unlikely to be possible through market mechanisms. The application of economic regulatory approaches is also likely to be highly uncertain due to the difficultly in predicting and controlling unlicensed use. Simpler, more pragmatic approaches might be better.

Interference between unlicensed users can occur between users of the same technology (eg WiFi users) and users of different technologies (eg between WiFi and BlueTooth). In general, same technology interference is managed and controlled through the technical standard. For example, WiFi has protocols that cause one transmitter to "back off" if it hears another transmitting on the same channel so as to avoid interference. It tends to be in the interests of technical standards bodies to ensure that any unlicensed technologies that they develop share the spectrum equitably with other users of the same technology otherwise their products are unlikely to be successful. Given that it is rarely possible to ascribe different economic values to different users of the same unlicensed technology, and even more difficult to implement unequal sharing, it is reasonable to assume that same-technology interference can be managed without regulatory intervention.

Interference between different technologies is more problematic. In some cases, different technologies do not share well – for example one might occupy a channel for long time periods denying access to another. With no restriction on the number of technologies that can be introduced in a band and innovative new concepts emerging over time, detailed regulatory activity is problematic. In previous work[9] Ofcom

[9] See http://stakeholders.ofcom.org.uk/consultations/scc/statement/ . This sets out that a fair wireless user is one that shares the resources equitably with other users, and behaves appropriately according to its needs. To behave fairly a device should implement a method to

considered this problem and suggested that any new technology being introduced in a band should demonstrate equitable sharing with existing users unlicensed users of the band. This assumes that all unlicensed uses are of equal value. In principle, it might be appropriate to allow some technologies or uses greater interference generation than others. In practice, this seems extremely difficult to engineer and monitor and hence equitable sharing may be the best outcome that can be achieved.

Controlling levels of unlicensed interference

As mentioned above, much of the efficacy of mixed usage rights depends on the extent to which the interference generated by unlicensed users can be controlled. With accurate control, levels can be set at the optimal points discussed above. With limited control it may be necessary to leave some margin to allow for uncertainty such that the expected interference levels are below optimal levels.

Control could be enacted through the database proposed for cognitive access. Such a database could be implemented by the regulator, although in practice regulators typically do not have the IT resources available and are likely to "outsource" such database provision to accredited third parties who may be willing to provide the service for free if it fits their business needs. Unlicensed devices would apply to a database and get granted access under certain conditions (power levels, location, time). If another device applies in the same location it is given a lower power (or none at all) so the total interference levels stay under the license conditions. Or alternatively it might prove possible to contact the first device to have its power reduced to share equally. Alternative approaches might learn the expected demand in a given location and time and divide up the interference accordingly, adapting their behaviour from day to day as demand varies.

Such control would be more accurate if there were a feedback mechanism whereby the interference actually generated could be monitored and compared with that allowed. This is generally difficult to achieve, requiring a dense network of monitoring devices which can differentiate between wanted and interfering signals (a complex task generally). However, it may be that the licensed user or even other unlicensed users can extract metrics from their network that provide some indication of the interference environment.

Where unlicensed devices do not consult a database or in some other way make contact with a central "planning" function the ability to control interference is much reduced.[10] It might be possible for there to be a broadcast "interference level" report which causes the devices to modify their transmit powers or similar, but it may require an enforcement mechanism to deal with selfishness.

PART THREE: SUMMARY AND CONCLUSIONS

Conclusions

The key task in the next stage of spectrum management is to adapt regulation to the prospect of widespread sharing, on a much more sophisticated basis than that

---

become aware of other users of the same resources, not monopolize the resources so that other users cannot access them and implement a method to reduce its channel occupancy when there is congestion.

[10] Several writers have investigated what has been called 'the price of anarchy', see R Johari and N Tsitlikis, 'Efficiency loss in a network resource allocation game' Mathematics of Operations Research 2004, pp. 407- 435.

phrase is used today. It is not the job of a spectrum regulatory agency to mandate new technologies, but it should get out of the way of public or private spectrum users which want to innovate.

However, there is a role for the regulator to take steps to expand the area of choice within which public and private sector users can operate. This is best done in general by enhancing the flexibility of usage rights, which itself is best achieved by enhancing the freedom to trade them in the dimensions of time, space, level of interference and priority of access, by subdividing, re-aggregating etc. However, there are considerable ‘transactions cost’ impediments to trading where unlicensed users are involved. This creates a role for the regulator pro-actively to investigate different allocations, to make provisions for the most promising to occur and to incorporate in refarming exercises and in primary assignments based on auctions configurations of usage rights which might favour promising avenues of spectrum use. By this means the regulator does not seek to determine the future history of innovation but to shorten the birth pangs of new technologies.

What might this mean in practice? At the very least it might mean that auctioned licenses should be in a format such as SURs and that they should have an explicit allowance for interference from unlicensed usage allowing subsequent introduction of technologies such as UWB , cognitive or some other future technology into the band. But given the difficulties of negotiations with unlicensed users it also suggests that the regulator should in some appropriate way seek to represent their interests, facilitating negotiations around changes of levels of interference or entry conditions into the band. As an example, when auctioning, say, the 800MHz band, the regulator might format the licenses as SURs, setting an interference level that allowed UWB operation and cognitive access into the band. They might then formulate licence exemption regulations for the band with allowed power limits and other conditions of entry that modelling suggested would keep unlicensed usage within these limits. Finally, they might encourage a forum where the licence holder and representatives of the unlicensed users could come together to negotiate possible changes to these limits or if this were not possible, act as a proxy for the unlicensed users based on an understanding of their economic incentives.