Allowing sensing as a supplement: an approach to the weakly-localized whitespace device problem

Kate Harrison and Anant Sahai

Wireless Foundations, EECS, UC Berkeley
{harriska, sahai}@eecs.berkeley.edu

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

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

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

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

I. Introduction

A. Whitespaces

Heterogeneous system deployments are common due to the natural spatial variation (e.g. in population). For example, cellular providers do not provide uniform service throughout the country due to variations in demand and cost of coverage. Allowing only one system to transmit in a particular band often leads to spectrum underutilization for this reason.

As the recent PCAST report [1] recognizes, our future ability to unlock spectrum for new uses is going to involve being able to exploit these gaps in coverage, called spectrum whitespaces. The FCC's recognition that databases can play a very important role in whitespace access in the television bands is something that we (and the authors of the PCAST report) expect to generalize to other bands as well. Consequently, the ideas explored in this paper are quite general, but we will explore them in the specific context of the TV whitespaces in the United States to make the discussion concrete.

B. Realistic whitespace devices: master and slave

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

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

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

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

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

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

[3]A slave device may also request that the master provide specific operating parameters. In this case, the problem is the same as in the US.

C. Contributions of this paper

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

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

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

Implementation details and data sources are given in the appendix. Source code for all simulations is at [21].

D. Related work

While there are many papers on sensing [4], cooperative sensing, location fingerprinting and other related topics, there does not appear to be any work which seriously looks at the master-slave issues we have identified nor any which addresses the inherent location uncertainty of the slave. Thus the work done by the FCC and Ofcom constitutes the majority of the prior work on this topic.

For example, [5] use a more restrictive form of location fingerprinting which includes a beacon in the primary's signal, requiring coordination and expenditure on the part of the primary. This lack of incentive alignment caused the FCC to drop the idea of beacons entirely [6, ¶70]. Moreover, the authors of [5] were primarily concerned with localizing the secondary device rather than accepting and addressing the location uncertainty.

A similar methodology which used TV signals to carry information about nearby used channels is considered in [7], but the authors missed the crucial step of using communicating with the whitespace database to reduce uncertainty.

II. The problem with the FCC's approach

Under the current FCC rules, a slave device employs a master device to help it access the whitespaces. In particular, the slave communicates its device ID to the master who relays this to the whitespace database (WSDB) along with the master's location. After verifying that the device is allowed to operate (based on its device ID), the WSDB signals that the master may share its list of available channels with the slave.

The assumption that is implicit in this mechanism is that *the slave is close enough to the master that the list of available channels has not changed*. However, a simple calculation[4] shows that even a pair of 40-mW transmitters can hypothetically communicate at more than 1 kbps at a range of 100km. While this is not an acceptable data rate for most purposes, it *is* enough to obtain the information which allows the device to access the whitespaces since the amount of master-slave communication needed for this is quite minimal[5].

The FCC makes this assumption more explicit in its 2012 comments on which fixed devices may act as master devices:

> **We are prohibiting fixed devices with an HAAT greater than the current maximum of 106 meters from providing channel lists to Mode I personal/portable [slave] devices.** This action is necessary because a Mode I device, which does not incorporate a geo-location capability, obtains a list of available channels from a fixed or Mode II device that is determined by the geographic coordinates of those devices. Under the current 106 meter limitation, **the communication distance between a Mode I device and the fixed or Mode II device that provides a channel list is relatively short**, and thus there is a low probability that a Mode I device will operate at a location where its channel list is not valid, i.e., does not meet the minimum separation distances from co-channel and adjacent channels TV stations or other protected services. However, **if the fixed device that obtains the channel list for a Mode I device operates with greater HAAT than the current rules permit, the Mode I device could operate at a greater distance from the coordinates of the fixed device where the available channel list was calculated. This will increase the chance that the Mode I device could operate at a location where the channel list is not valid.** We will therefore require that the TV bands database not provide channel lists for Mode I devices through fixed devices with an antenna HAAT of greater than 106 meters. [2, ¶19] *(emphasis added)*

At this point the problem has been tacitly recognized but it is not well understood. Television service areas (whose borders

---

[4]Here we use the theoretical Shannon capacity of a channel, pathloss from the ITU propagation model, the bandwidth of a TV whitespace channel in the United States (6 MHz), a 10 meter HAAT, and a noiseless channel (thermal noise only).

$$\begin{aligned}
\text{rate} &= W \cdot \log_2\left(1 + \frac{\text{signal power}}{\text{noise power}}\right) \\
&= W \cdot \log_2\left(1 + \frac{(\text{transmit power}) \cdot (\text{pathloss})}{\text{noise power}}\right) \\
&= 6 \times 10^6 \cdot \log_2\left(1 + \frac{(.04) \cdot (8.8 \times 10^{-17})}{2.4 \times 10^{-14}}\right) \\
&= 1274 \text{ bps}
\end{aligned}$$

[5]We believe that another implicit assumption is that the slave device will be regularly communicating with the master device (i.e. they are part of the same system) and thus such a low rate would be unacceptable. However, this need not be the case: a slave may "associate" with any master with whom he can communicate.

often mark a change in channel availability) are often—and usually correctly—thought of as large areas. For example, Figure 1 shows the approximate service areas of the television stations on channel 10. From this view, the FCC's lightweight approach to the problem appears reasonable.

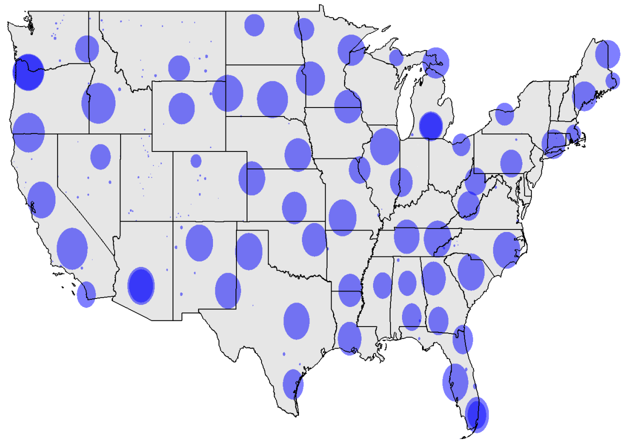


Fig. 1. Service areas for TV towers on channel 10.

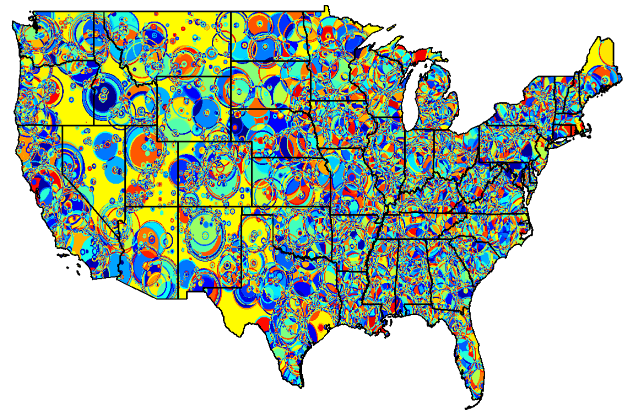
Even when we consider the aggregate effect of the television stations on all channels, the spatial variation appears smooth and predictable. For example, see Figure 4(a) in which the number of available whitespace channels is plotted.

However, results such as these lull us into a false sense of security. Consider Figure 2 which highlights the variation in the list of available channels across the United States. Each unique list list of available channels[6] is mapped to a unique color, which is then plotted on a map. As can be seen in Figure 2(b) (a zoomed-in version for clarity), the list actually changes quite rapidly. This variation, especially the concentric circles, results from two things:

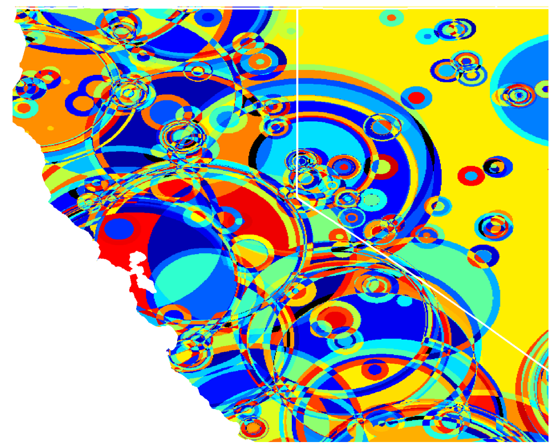
1) The co-channel and adjacent-channel excluded areas for the same protected region will differ slightly (the co-channel exclusions extend about 10 km further than the adjacent-channel exclusions). This leads to some of the concentric circles[7]
2) In many cases, the same tower will transmit on different channels (e.g. Sutro Tower in San Francisco broadcasts on over 10 TV channels [10]). Each channel has different propagation characteristics as well as a potentially-different transmit power. This means that although they may aim to serve the same market, each station will end up with a slightly different service area.

[6]It may be helpful to think of the list of available channels as a binary vector, with each entry answering the question "is this channel available for secondary use?"

[7]Of course, the real world will differ from the figures shown here as we are using the simple ITU propagation model [8] which does not take terrain into account. Incorporating terrain would mean that service areas would change from being perfect circles to being slightly (or significantly, depending on the variation in terrain) misshapen. (This can be seen by looking at the actual protected contours for TV stations, available on the FCC's website [9].)



(a) Full map



(b) California-Nevada region (zoomed version of Figure 2(a))

Fig. 2. Color-coded maps showing the variation in the list of available channels (a) across the United States and (b) in the California-Nevada region. Each color represents a unique list of available channels.

Note that we have not said anything about the size of the difference between the various lists, only that they are different. In the same way that relying on the "number of whitespace channels" map in Figure 4(a) is extremely optimistic, relying only on this very colorful map is extremely pessimistic. A more reasonable metric is the probability of wrongful transmission, where "wrongful transmission" is defined as a transmission which is against the spirit of the regulations (e.g. inside the service area of a TV station).

In Figure 3, we show the CCDF of the probability that a slave will transmit on a channel which is actually not available at his location, despite being assured by the WSDB and the master that the channel is available[8]. We have plotted several scenarios, each representing the slave's maximum distance from the master. For example, we can see that even if the slave is within 10km of the master, about 40% of the population has at least a 4% chance of wrongful transmission.

There are a few points which are important to keep in mind

[8]Details on this calculation can be found in the appendix.

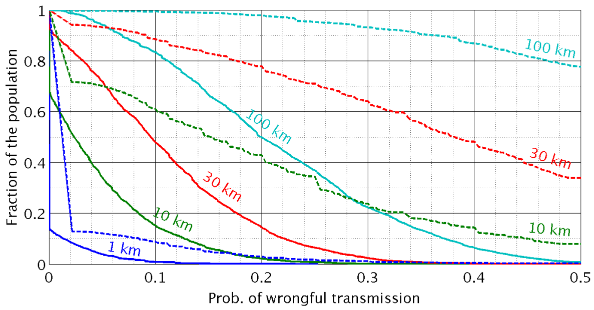


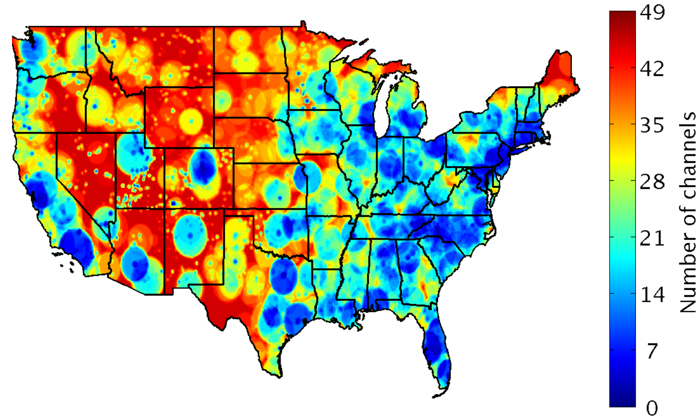
Fig. 3. CCDF of probability of wrongful transmission with an ignorant (solid lines) and opportunistic (dashed lines) secondary transmitter. Details on this calculation can be found in the appendix. (Note that the jump in the opportunistic lines from 0 to 0.02 occurs because the probabilities are discrete with minimum step size 0.02.)

when interpreting these results:

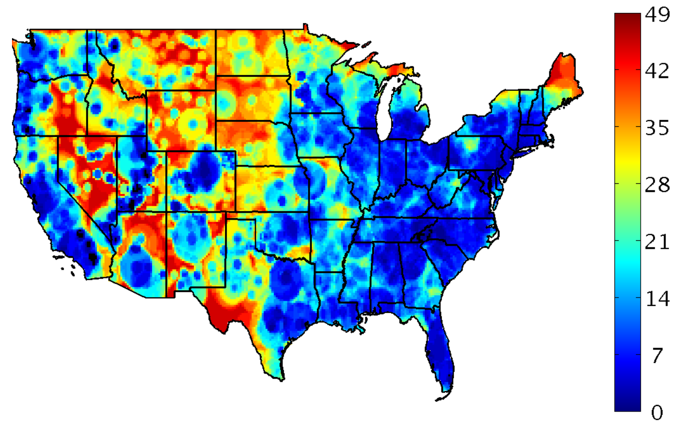
- While 10km (and especially 100km) might seem extreme, note that (1) the master is potentially quite long-range (it may, for example, be a fixed device with height up to 106 meters above average terrain [2, §15.711(b)(3)(iv)(C)]) and (2) the slave has little to no incentive to choose a nearby master if its first goal is to obtain more spectrum. We assumed uniform distributions for an "ignorant slave" (solid lines) but a slave could easily choose to "game the system" by choosing the master which gives the most favorable results ("opportunistic slave," dashed lines).
- The number of channels available to such an opportunistic slave device[9] is shown in Figure 4(c) (we assume it can only use masters within 25km). The corresponding probability of wrongful transmission (assuming the slave chooses the most favorable master but then chooses uniformly among the channels "available" to it) is shown by the dashed lines in Figure 3. The probability of wrongful transmission is alarmingly high even with an optimistic 10km limit on the master-slave range.
- The slave need not contact the master via whitespace channel. In fact, he is not even limited to communicating with the master wirelessly. He thus is not limited to whitespace emissions limits which potentially increases his range (e.g. via ham radio frequencies[10]). See §15.711(b)(3)(iv)(D) of the 2012 FCC rules [2] for more details on the master-slave communication requirements.
- Once the initial exchange is complete, bidirectional communication is no longer required (see [2, §15.711(b)(3)(iv)(D)]). In this manner, a slave device could initiate contact with a master while he is nearby, then move to another location and simply receive periodic updates from the longer-range master. We believe this is an unintentional loophole in the regulations.

[9] Although we tend to think of a model in which the master and slave devices are tightly coupled (e.g. via proprietary protocols which vary by manufacturer), we envision a much richer and liberal ecosystem in which, for example, master and slave devices interoperate naturally.

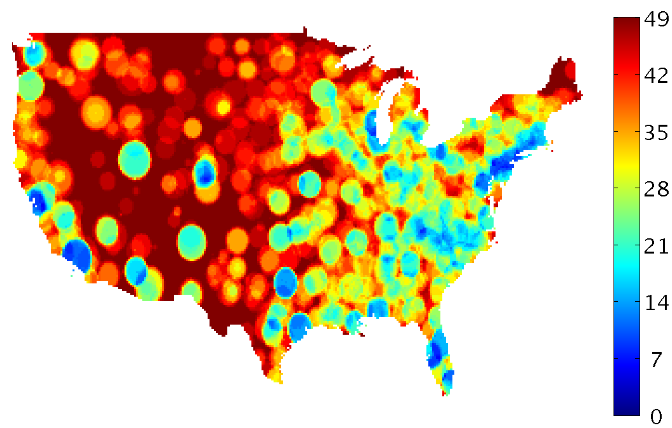
[10] He could technically communicate with the master over the Internet, but we believe this is an unintentional loophole.



(a) Number of channels available to a device with perfect geolocation



(b) Number of channels available to a slave which knows only that it is at most 25km from the master



(c) Number of channels available to an opportunistic slave which contacts the "most desirable" master (i.e. the one which reports the most available whitespace channels) within 25km

Fig. 4. Estimated number of whitespace channels available with (a) perfect geolocation and (b) location uncertainty. We are not imposing the artificial restriction that mobile WSDs (e.g. slaves) only use channels 21-51, but we expect that the same basic picture will remain.

III. Ofcom's approach to the master-slave problem

The next step to solving this problem has been taken by Ofcom. We will describe this solution and its shortcomings in this section.

The UK's Ofcom has partially accounted for this problem in their rules, using the service area of the master to estimate the maximum distance between the slave and the master [3, ¶5.11].

> For this, the WSDB will use the TVWS availability data obtained from Ofcom and the channel usage parameters received from the master WSD (see 5.35.5) to calculate the coverage area in which slave WSDs are likely to operate. It will then calculate the generic operational parameters that apply within the coverage area based on a number of default (conservative) device parameters. [3, ¶5.38.2]

They then appear to use this estimated distance to create a set of locations which is highly likely to contain the true location of the slave. Using these potential locations, they can calculate which channels[11] will be safe for the slave to use as long as he is at one of those potential locations.

Although there are some subtleties to estimating the master-slave distance (e.g. using the antenna gain of the slave device [3, ¶5.85]), this solution can certainly be made conservative enough to adequately protect the primary's systems. One ridiculous example of this is to assume that the slave is always within 1000 km of the master: while it is an extremely conservative bound, it has a very high probability of being correct.

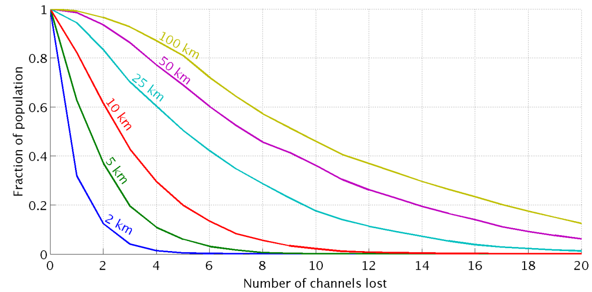
IV. The problem with Ofcom's approach

The problem with Ofcom's approach is that a tradeoff has been established: protecting primaries vs. providing reasonable opportunities to slave devices (as compared to devices with geolocation). For example, we see in Figure 4(b) that assuming a maximum master-slave distance of 25km drastically reduces the number of whitespace channels available to slave devices in the United States. We have already argued in Section II that even 25km is a conservative estimate, so it is providing poor service to both primaries and secondaries.

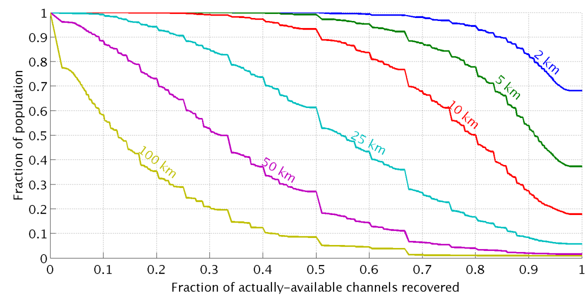
For simplicity, we will assume a fixed master-slave maximum distance (i.e. that the slave is within R km of the master) as opposed to one which varies based on the characteristics of the master and slave devices. Figure 5 quantifies this problem for various values of R in the context of the United States.

In Figure 5(a), an empirical CCDF (complementary cumulative distribution function) weighted by population shows the raw number of channels lost to the slave due *solely* to his lack of geolocation capability. As we expect, the number of channels lost increases with R. For example, at $R = 2$ km, the median person is losing about 1 channel as opposed to 2 channels at $R = 10$ km.

[11]Note that part of the motivation for Ofcom's regulations are its vision of variable power limits which depend on the device's location. To understand the subtleties of this, please refer to [11].



(a) Number of channels lost using Ofcom-like regulations



(b) Fraction of actually-available channels which are available to slave devices using an Ofcom-like regulations

Fig. 5. CCDFs (by population) showing the impact of Ofcom-like regulations

However, channels have different value in different places depending on their scarcity: in urban areas where whitespace channels are sparse, each channel is worth more. In contrast, whitespace channels are less valuable (individually) in rural areas because they are relatively abundant. Thus in Figure 5(b) we look at what fraction of the actually-available (i.e. available with perfect geolocation) whitespace channels a slave device would be able to use under Ofcom-like rules. Again, with $R = 2$ km we see that most people will be able to use most of the actually-available channels whereas with $R = 10$ km the median person will only recover about 80% of their potential channels. This metric is quite useful as it provides a number-of-channels-agnostic approach to quantifying these rules.

Figure 6 shows the percentage of places which, purely as a result of location uncertainty, lose *all* access to the whitespaces. For now, it is sufficient to note that Ofcom's approach scales very poorly with the uncertainty radius: at $R = 100$km, 15% of the country loses access to whitespaces that would have been available with geolocation capabilities.

We wish remind the reader that in order to be completely safe, the regulator must choose a distance beyond which they think it will be *impossible* for a slave to communicate with a master. Consider for a moment that a 100-mW transmitter (100 mW is the maximum EIRP for personal/portable devices) can theoretically communicate at 3.2 kbps to another device 100 km away[12]. This means unnecessarily barring slaves from a great number of channels which otherwise could have been

[12]Assumptions: channel 2, 10.1m HAAT, clean channel. Even a 40-mW device could communicate at almost 1.3 kbps.

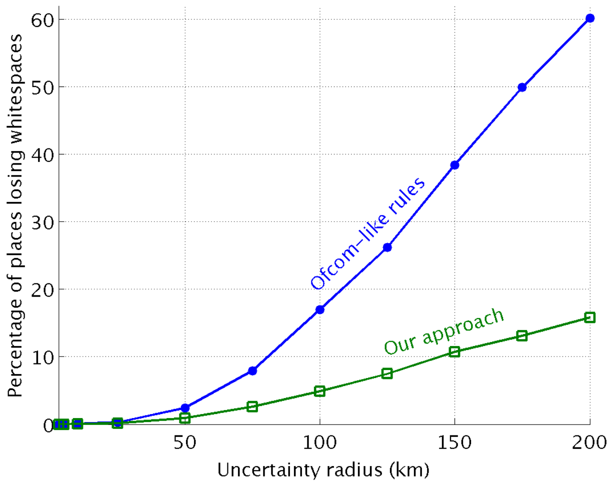


Fig. 6. Percentage of places that lose access to the whitespaces under various approaches.

safely used[13]. Keep in mind that even a device which actually has geolocation capabilities will be forced to operate as a slave if it cannot successfully locate itself (e.g. GPS fails indoors).

In the next section, we will suggest a simple enhancement to Ofcom's approach which will greatly increase the opportunities for slave devices while adequately protecting primary systems.

V. Proposed enhancement to Ofcom's approach

Let's summarize what we have learned so far:

1) A slave device could conceivably be located tens or hundreds of kilometers from its master.
2) The list of available channels can vary rapidly with location.
3) Using the master's location for the slave is inherently unsafe due to the first two points.
4) Ofcom's solution of using only those channels available in the slave's set of potential locations drastically reduces the slave's opportunity, forcing a choice between primary and secondary quality of service.

We believe that this tradeoff is actually a symptom of the way that people are thinking about solving this problem, rather than an unavoidable fact of nature. Instead of constraining the set of possible slave locations to being a circle, why not allow it to take any shape at all? For example, it is clear how the database should respond to a slave device which can reliably say "I'm in either San Francisco or New York City, but I don't know which." The size or shape of the set of potential locations (which we term the "uncertainty region") is actually *completely irrelevant*.

The question now becomes: how can we meaningfully reduce the size of the uncertainty region? We have previously established in Figure 2 that the list of available channels has high spatial variation[14]. There is an entire field which capitalizes on spatial variation in radio signals: location fingerprinting.

Location fingerprinting in conjunction with radio environment maps helps a user to pinpoint his location by correlating signal measurements at his location with those in a pre-populated database[15]. For example, WiFi fingerprinting uses the names and signal strengths of nearby wireless access points to estimate a user's location. This is done by Apple in its iOS when GPS is not available [17].

Note that our approach is to actually allow the consideration of a *simpler* problem than the traditional localization problem.

[13] An alternative is to mandate a minimum spectral efficiency for master-slave communications, e.g. 2 bps/Hz.

[14] Moreover, its variation is actually correlated to the phenomenon we'd like to detect (i.e. available channels).

[15] **Location fingerprinting**: Location estimation is a general problem which has been addressed extensively. This work has been done with an eye toward sensor networks. In this application, many sensor nodes are deployed to gather information about the environment. The nodes may be statically configured or may be moved from time to time. Determining the exact position of each sensor requires significant effort and potentially hundreds of nodes may be deployed, hence automatic position-finding significantly reduce the costs of sensor networks.

For example, [12] and [13] consider the problem of determining the relative position of nodes in a network. [14] considers a similar problem where a small portion of the nodes know their absolute locations; RSS and TOA are used to estimate distances.

Another technique for location estimation is location fingerprinting. Location fingerprinting (LF) is a localization technique which uses pre-existing (computed or measured) signal strength maps—typically of WLAN signals—in conjunction with measurements from the device itself. In particular, the measurements are checked against the maps to determine locations in which the measurements are self-consistent.

The accuracy of LF depends heavily on the map data as well as the usage environment. For example, WLAN LF works quite well in urban areas where WiFi hotspots are in abundance (and the maps are likely to be of good quality and recent) whereas it performs poorly in rural areas.

Note that the TV bands (as well as AM and FM radio bands) have ideal qualities for use in LF: (1) stations are relatively static over time; (2) stations are at fixed locations; (3) signals in these bands propagate well.

LF requires map data which is extensive and accurate. Thus, the biggest challenge in LF is creating and maintaining the database of map data [15]. There are two main techniques for database creation: (1) radio environment mapping and (2) use of propagation models to predict the signal strength at various locations. While (1) is more accurate (assuming the landscape doesn't change significantly), it requires substantially more data-gathering effort.

Radio environment maps: The now-conventional approach to whitespace databases assumes that ground-truth (regarding which channels are free to use and where) is obtained via registration and reliable propagation models. However, while registration is easy to assume for TV transmitters in first-world democracies, many other systems and countries cannot rely on such information being easily, publicly accessible.

There is extensive research on how to create and maintain radio environment maps (REMs). Radio environment maps synthesize information from many sensors to create an estimate of the signal strength at a wide variety of locations. REMs are typically used for location fingerprinting, network planning, and signals intelligence.

Creating an accurate REM is a difficult problem for several reasons:

- Gathering accurate data is a costly proposition as it requires many measurements, particularly in fast-changing environments (e.g. cities) [15].
- Synthesizing such large amounts of information may be quite difficult and time-consuming, depending on the collection method [16].
- If sensing information comes from multiple types of devices, the different device characteristics (e.g. sensitivity) must be accounted for.
- Depending on the application, the signal strength may vary with time (e.g. in public safety bands which are used intermittently), further compounding the problem.

We relax these problems by *not* requiring full localization and instead attempt to simply reduce the uncertainty in a device's location. This uncertainty—however large or small—is acknowledged by the whitespace database in that it will only permit the device to transmit on channels which are available for secondary use *everywhere* the device *could* be located. In this way reduced uncertainty will increase the whitespace opportunity for that device but any amount of uncertainty can be accommodated. Thus we can simply "solve" the location fingerprinting problem to the desired fidelity and not worry about fully solving these problems.

We strengthen our argument with the following facts:

- Most devices are transceivers rather than transmit-only devices. Thus they have the prerequisites for rudimentary receiving (i.e. sensing) capabilities.
- Although sensing for the *absence* of TV signals is quite hard [18], it is in fact relatively easy for a device to determine if it could *receive* TV. TV signals are at or above 15dB SNR inside the service area, something that is quite easy to sense[16].
- The list of channels available for secondaries has a lot of variation, as seen in Figure 2(a). Similarly, the list of channels on which users can successfully watch TV also varies a lot (omitted for brevity).

When we combine the three facts above with location fingerprinting, a potential solution is quite clear: a slave can reduce the size of its uncertainty region using an easy-to-generate report on which TV channels it receives[17] (i.e. someone could watch TV on that channel there).

So how does it work? Let us first explain our process with a toy model which will also be used later. We will then formalize our proposal.

A. Illustrative example

A simple example is illustrated in Figure 7. This example shows how the uncertainty region of a secondary device changes as it learns more about its environment.

Figure 7(a) simply sets up the example. It shows the service areas of towers on two channels, a red channel and a blue channel. True to real life, these service areas may or may not overlap.

Figure 7(b) shows in gray the initial uncertainty region of a particular secondary device. The size of this initial region is determined by the device's estimated maximum distance to the master and, since no other information is known yet, takes the shape of a circle. The secondary's true location is marked by a star. Notice that TV signals for both channels will be strong at this location.

In Figure 7(c), we see the initial reduction in the size of the uncertainty region. Having sensed a strong TV signal on the red channel, the secondary can use this information to exclude areas where TV signals on the red channel are expected to be weak[18]. Notice that the resulting uncertainty region is now disconnected (which is *not* a problem) and, more importantly, reduced in size.

The final figure, Figure 7(d), shows the secondary's uncertainty region after sensing a strong TV signal on the blue channel. Again, regions where blue-channel TV signals are expected to be weak are excluded from the uncertainty region. We see now that the uncertainty region is just a fraction of its initial size, demonstrating the power of our location-fingerprinting-like technique.

A side note: our assumption is that the slave simply notices when it receives TV on a particular channel (thus it is in the service area for *some* tower on that channel). However, since it will be within the service area, one might suggest that it actually decode the TV signal and determine the station ID, further reducing its uncertainty. While this is possible in theory, it might not be practical: including an ATSC decoder module might be too expensive to justify for devices which are not already intended for watching TV—for the same cost, a GPS unit would solve the same problem (when outdoors).

B. Details of our proposed solution

We propose that the slave be allowed to use sensing techniques[19] to determine the channels on which it can receive TV in order to reduce the size of its uncertainty region. Although the slave would also benefit from identifying those channels on which it cannot receive TV, a device which can do this reliably may as well be a sensing-only device. Our solution hinges on the fact that the presence of a signal is much easier to detect reliably than the absence of one.

Once the device has identified some channels on which he receives TV, it communicates these to the WSDB (through the master). The whitespace database uses this information in conjunction with its radio environment maps and the slave's maximum distance to master in order to narrow down the set of possible slave locations. The WSDB does so through Algorithm 1 which follows the approach described in Section V-A.

VI. Evaluation of proposed solution

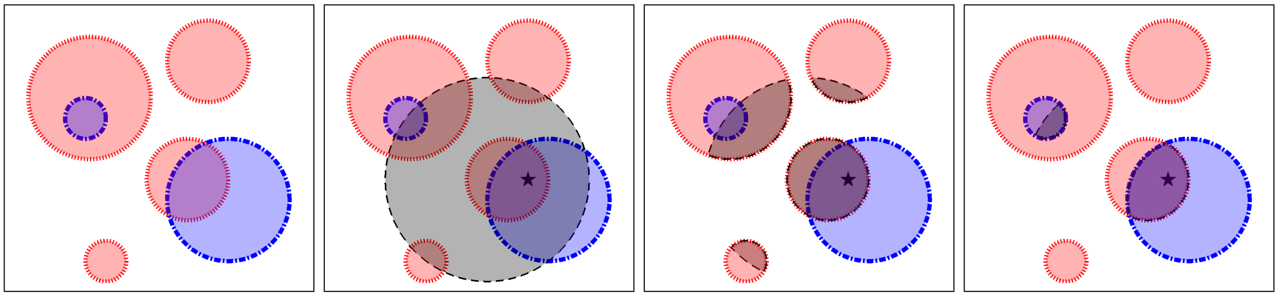
First, we compare our proposed approach with the Ofcom-like method from Figure 5 and the results are shown in Figure 8. For high values of R, learning the channel states is incredibly valuable (e.g. the median person goes from recovering just over 10% of his actually-available channels to

[16]At first glance, it may appear as though this solution applies particularly well only to the TV whitespaces. However, it is not unreasonable for a device to use TV signals to reduce its location uncertainty in order to gain access to a *different* set of whitespaces. After all, the PCAST report [1] suggests that all whitespaces should use a common database.

[17]Notice how we took a somewhat useless or disappointing result in the sensing-only world—"oh no, that channel's in use... I guess I'll try another one"—and turned it into useful information!

[18]In this paper, we consider the sensing output to be either "can receive TV" or "cannot receive TV." In reality, at minimum a third option—"cannot tell"—should be allowed. Conceivably, more precise outputs could be used.

[19]We recognize that this is actually not as simple as it sounds. For example, the slave will need to determine that he hears a TV signal as opposed to noise from other whitespace devices. However, in most cases he should receive either TV signals *or* WSD signals, not both: recall that WSDs cannot transmit within the service area of a TV station.



(a) The service areas of towers on two channels, red (dotted lines) and blue (dashed lines), are shown.

(b) Uncertainty region of secondary is shown in gray. Its true location is marked by a star.

(c) Uncertainty region after the secondary senses a strong TV signal on the red channel.

(d) Uncertainty region after the secondary also senses a strong TV signal on the blue channel.

Fig. 7. To illustrate our approach, we present a toy example as a series of figures. In (a), the service areas of TV towers on two channels, red and blue, have been marked. (b) shows in gray a sample "base" uncertainty region for a secondary device which knows only its distance from the master. A star marks the secondary's true location. After sensing a strong TV signal on the red channel, the secondary's uncertainty region can be reduced, as shown in (c). Notice that the region is noncontiguous but reduced in area. The secondary subsequently senses a strong TV signal on the blue channel which further reduces its uncertainty region, as shown in (d).

Algorithm 1 Potential WSDB algorithm to calculate list of truly available channels for slave device

Require: T: list of channels on which TV is received at the slave ($|T|$ may be 0)
Require: L: precise location of the master device
Require: R: estimated maximum distance between slave and master ($R \geq 0$)
Ensure: Slave device transmits only on channels which are truly available
{Determine the uncertainty polygon}
Uncertainty polygon, P, initialized to a circle of radius R centered at L
for all t in T **do**
  Let A_t represent the coverage area of t
  $P \leftarrow P \cap A_t$
end for
{Determine the channels available everywhere in the uncertainty polygon}
List of channels available for slave use, C, initialized to the empty set
for all whitespace channel, W **do**
  if W available for whitespace use everywhere in P **then**
    $C \leftarrow C \cup \{W\}$
  end if
end for
return C

recovering over 65% of the same channels in the $R = 100$ km case). Naturally, the differences are smaller for small values of R since there was less room for improvement with the original Ofcom-like method.

Now we've shown that there are real gains to be had from using this not-very-complicated system, let's see what's going on in more detail. We performed a similar test (via Monte Carlo simulations) on real-world data[20] and the results are shown in Figure 9. Each subfigure shows a different original uncertainty distance so as to be comparable to the results shown earlier. Each subfigure shows the results for slightly different scenarios:

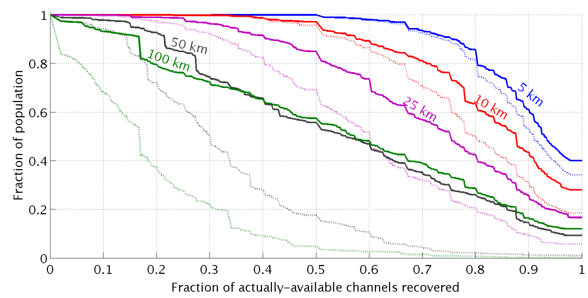


Fig. 8. CCDFs (by population) comparing our approach (solid lines) to an Ofcom-like approach (dotted lines) for each uncertainty radius R.

- **LrxO** ("learn rx [received channels] only," blue solid line): learn about channels on which TV is received[21]. If TV is not definitely received on a particular channel, we conclude nothing after "learning" its channel state. (Note: it would be more accurate but less useful if this line simply stopped after we had learned all of the received channels. However, we extend it horizontally for the purposes of comparison with the other lines.)
- **LrxF** ("learn rx [received channels] first," red solid line): learn first those channels on which TV is received (it's faster to sense these channels so in reality even a sensitive slave would do this) and then confirm that the other channels do not receive TV.
- **LA** ("learn all," purple solid line): learn all channel states

[20] Information on our methods can be found in the appendix.

[21] Within this set of channels, channel states are learned in a random order which varies for each point in the Monte Carlo simulation.

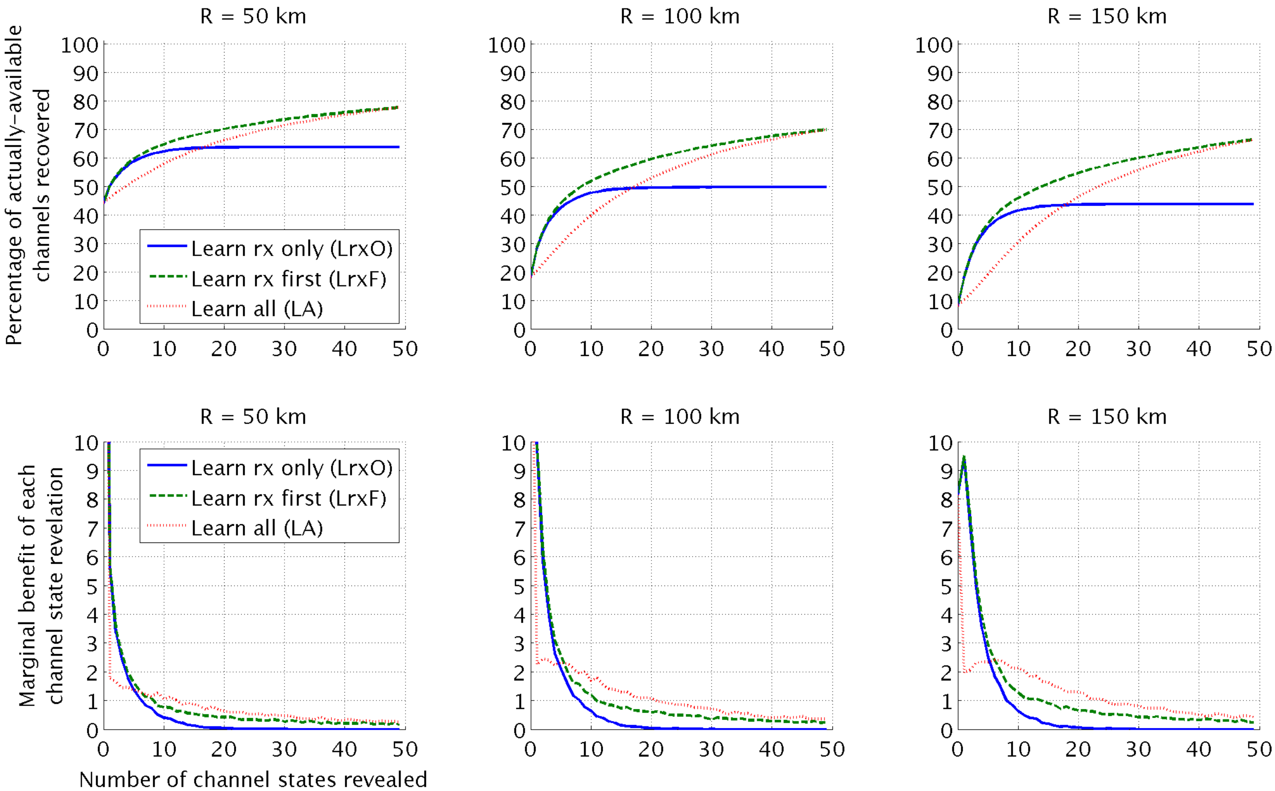


Fig. 9. Percentage of recovered channels as a function of the number of channel states revealed under various schemes. Solid lines represent absolute recovery percentage while dashed lines indicate the marginal benefit of learning a new channel state.

in a random order.

The x axes of the plots represent the number of channel states which have been learned by the slave. In other words, $x = 0$ is equivalent to Figure 7(a), $x = 1$ is equivalent to Figure 7(b), etc. The y axes list the percentage of channels that can be recovered (as compared to the number of channels recoverable with perfect geolocation).

The solid lines represent the absolute recovery percentage, which is naturally quite interesting. However, it is also interesting to see what the marginal benefit of learning each channel state is and this is shown by the dashed lines (colors match). Let's notice a few things about these plots:

1) As expected, the absolute recovery percentage (top plots) decreases as the uncertainty radius R increases.
2) LrxO (the solid blue line) provides the most benefit (as compared to LA, the dotted red line) for awhile until there are no more channel states to learn. At that point, continuing to learn anything (as opposed to receiving no new information) obviously wins out.
3) If we were allowed to continue learning channels after we ran out of channels with the LrxO method (shown by LrxF, the dashed green line), we would continue to do better than LA. It appears that more information is revealed if you know you're within a TV's service area than if you're not (comparing LrxF and LA). This may be because being inside of a service area is a low-probability event (look at the sparsity of towers on each channel). It may also be because the state of the adjacent channels can be more easily predicted (there is a high probability that they are also unavailable for secondary use due to adjacent-channel exclusions).
4) This last conclusion is also supported by the graphs of the marginal benefit of learning each channel state (bottom plots). The marginal benefit of learning the state of a random channel (LA, the dotted red line) is initially lower than that of learning a channel on which TV could be received (the other two lines). As more channel states are revealed, this ordering switches. This is because (1) the LA approach is still learning a mixture of received and not-received channel states while (2) the LrxF approach is learning only not-received channel states. The marginal benefit of the LrxO approach obviously goes to zero because no further channel states are revealed once all TV-received channels are known.

A. Connection to location fingerprinting

Location fingerprinting (LF) is a technique used to determine a device's location based on its knowledge about the environment. This is achieved by comparing its observations with global data while attempting to find a unique location at which its observations are consistent.

Our proposed approach is very similar to location fingerprinting in that it also uses local observations matched against global data. There are two main differences, though: (1) we already have the global data in the TVWS databases and (2) the *problem* is different. The goal of LF is to uniquely pinpoint a device's location whereas our approach is to identify a set of channels which are safe to use at any of the potential locations. It is in fact possible to find a common list of safe channels despite having very distant potential locations.

The key point is this: location uncertainty is *not* the same as uncertainty in the channel list. Heretofore we have discussed the location uncertainty as if it is the only important thing. However, what one *actually* cares about[22] is the uncertainty in the channel list: if we know which channels are safe to use, it doesn't matter where the slave is located. This is important because it is possible to have a high location uncertainty but have a good idea of which channels are safe to use. So note that we will not necessarily know where the slave is even if we know which channels it can safely use. In fact, even in the basic case where the slave knows only its maximum distance R to the master, it can still recover some channels without knowing its location very precisely, as shown in Figure 5(b).

VII. Consequences: added value for databases

In this section, we discuss some of the opportunities for database providers to add value under our proposed system.

A. Channel state discovery

Naturally the order in which channel states are revealed will have an effect on how quickly and efficiently the uncertainty can be reduced. Databases seeking to add value could interact with the client (rather than just receiving a list of channel states) to tell him which channel state to sense next (e.g. "can you definitely receive TV on channel C?"). This would reduce the discovery cost (in time and energy) for the slave by telling him where to concentrate his efforts.

B. Types of fingerprint information

Databases could also compete on how much and what kinds of information they can safely fuse. For example, databases might contain additional radio environment maps (e.g. for air traffic controller frequencies or cellular bands) which would let slave devices use fingerprint information from other bands. If we accept the vision of the PCAST report (shared whitespace infrastructure), it is likely that databases will already contain such information.

C. Types of location information

In future work, we plan to discuss the idea of using multiple services for localization. For example, a slave device might find himself in the following situation:

- Within R_1 km of master 1.
- Within R_2 km of master 2.
- Able to detect strong TV transmissions on channels $C_1, ..., C_n$.

A database which can safely fuse a wide variety of location information would clearly be of use to slave devices.

VIII. Conclusion

This paper has described in some detail the FCC's and Ofcom's approaches to safely enabling slave devices (i.e. devices which do not have geolocation capability). We have described the most glaring problem with the FCC's approach, namely its indifference to the distance between the master and slave. We have shown quantitatively that this is potentially a significant problem, especially if the whitespace device chooses its master opportunistically.

We have also described Ofcom's approach to this problem, in which they estimate the distance between the master and the slave. We have shown a major downside to this approach: it necessarily establishes a tradeoff between primary protection and secondary quality of service.

We went on to propose an enhancement to Ofcom's approach which leverages existing capabilities in secondary devices to safely allow them to recover much of what was unnecessarily lost with Ofcom's approach. We showed qualitatively that our proposed approach could provide gains over existing approaches. This lessens the tension between primary and secondary because now the distance can be very conservatively estimated without significantly negatively impacting the secondary device.

Although we have proposed an enhancement that appears to provide gains for all parties, the approach itself is not actually the most important part of this work. Instead, we wish to convey the message that localization services can take many forms, as seen, for example, in the iPhone which aggregates data from GPS, WiFi fingerprinting, and cellular positioning. In fact, our proposal is very similar to assisted GPS as seen in cell phones. Localization is an integral part of whitespace devices, and constraining it to a narrow definition (e.g. 50-meter accuracy) is unnecessarily restrictive to innovation.

In the regulatory sense, localization should be loosely defined as "the ability to provide information from which a set of *potential* locations can be computed." Databases can then choose if and how to fuse this information, potentially with guidance from a regulator or standards body. Since the databases are software-upgradeable and overseen by the regulator, any bugs or loopholes can be quickly patched. In the worst-case scenario, a particular type of information can be deemed unsafe and hence ignored.

[22] It is certainly possible that there are reasons beyond safety for wanting to locate the slaves. However, we take the viewpoint that safety and deployability are the main concerns of the regulators.

IX. Future work

We'd like to end with a few notes on some of the work that still needs to be done.

A. Consequences of incorrect channel state estimates

We have not analyzed the consequences of incorrect channel state estimates, though we have tried to minimize the likelihood of such an event by suggesting that the slave identify only those channels on which it can definitely receive TV (where the chances for a false positive are quite low due to the high SNR conditions). However, the potential for harmonics in nearby channels complicates this as well.

Incorrectly identifying a channel state will, with our simple proposed algorithm, certainly result in an incorrect location estimate because of the nature of the algorithm. However, if the slave identifies enough channel states, an inconsistency may eventually be discovered (e.g. "there are no places where you can watch channel 2 and channel 5 so one of your states must be wrong") so that at least the error can be recognized. Future work in this area may draw on results in the location fingerprinting literature to address this problem in a robust way.

B. Errors caused by propagation model inaccuracies

It has frequently been noted [19] that propagation models are inaccurate when compared to the real world. Hence the database may incorrectly estimate the service areas of the TV towers (though it could be seeded with the predicted contours from the FCC) and thus incorrectly estimate the slave's list of available channels.

However, slave devices need not give the databases a binary answer to "do you receive TV on channel C?". Instead, they could feed back the signal strength sensed on that channel and the database can apply the threshold. Since the slaves will also be transmitting their device IDs which link them to a specific make and model of device (to check that they are allowed to transmit), the database could also use information on the sensitivity of their sensing to build up a map of signal strengths throughout the nation. In fact, there is already work in this field [16], [20]. Thus over time we can refine the information in the databases to minimize this problem. An important thing to note is that the quality of our estimates will improve with the number of devices, thus improving the accuracy at the same rate as accuracy becomes important.

X. Acknowledgements

We thank the U.S. NSF (CNS-403427, CNS-0932410, CNS-1321155, ECCS-1343398 and a Graduate Research Fellowship) for their support.

Appendix
Methods

This appendix contains details on the methods used to produce the results in this paper. While we believe that these details are important, we expect that they will not interest the average reader, hence they are situated in the appendix. All relevant code can be found after publication at [21].

We use three sources of data for our models:

- Population data from the US 2010 census [22], [23]
- ITU propagation model [8]
- TV station assignment data from the FCC [24]

We also assumed for simplicity that WSDs can transmit on any of the whitespace channels, rather than being artificially limited to, e.g., channels above channel 20. The current FCC rules for TV whitespaces allow only fixed WSDs (i.e. not slaves) to transmit on channels 2-20.

Probability of wrongful transmission

1) Ignorant secondary: The probability is calculated thusly:

- Suppose the master is at location L and that we know the slave is within distance R of the master. Draw a circle of radius R around L. This represents the set of possible slave locations, S.
- For each channel C that is *available* for secondary use at location L, calculate the fraction of the area of S in which channel C is *not* available for secondary use. Call this fraction $P_{wrong,C}$. This represents the probability that the slave will wrongfully transmit given that it is operating on channel C and assuming that his location is uniformly distributed within S.
- Average $P_{wrong,C}$ across all channels which are available at the master. This is the probability that the slave transmits wrongfully given that it chooses a channel uniformly from the list provided to it by the master.
- The above process calculates the probability of wrongful transmission at a location L. Repeat this process for many locations across the United States, then calculate the CCDF weighted by population. This is what is represented by solid lines in Figure 3.

2) Opportunistic secondary: For the probability of opportunistic-and-wrongful transmission, we create an "opportunistic channel list" (call it O) by marking a channel as available for secondaries if it is available *anywhere*[23] within S. The slave is assumed to choose uniformly among the channels in O, so the probability of wrongful transmission in this case is simply

$$\frac{|O| - |C|}{|O|}$$

Note that since $|O| \leq 49$ (there are 49 TV channels), the minimum nonzero value of this fraction is $\frac{1}{49}$ which accounts for the initial jump in the dashed lines in Figure 3.

Evaluation of proposed enhancement

Our real-world simulations are Monte-Carlo style for tractability reasons. In general, we use a discretized map of the United States, as illustrated in [25]. All data (e.g. available channel lists, TV service areas, population data) is discretized

[23]For practical reasons such as numerical accuracy, we require that the channel be available for secondary use in at least 1% of S.

before being processed. Typically we use 3200 x 4800 points to cover the entire US, which means that the typical point represents roughly one square kilometer of area. In some cases, we evaluate on a much smaller grid (e.g. 200 x 300) for tractability but the underlying data (e.g. channel lists) comes from the higher-resolution map.

REFERENCES

[1] 2012 PCAST report. [Online]. Available: http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast%5Fspectrum%5Freport%5Ffinal%5Fjuly%5F20%5F2012.pdf

[2] "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Third Memorandum Opinion and Order," Federal Communications Commision, Tech. Rep. 12-36A1, Aug. 2012. [Online]. Available: http://hraunfoss.fcc.gov/edocs%5Fpublic/attachmatch/FCC-12-36A1.pdf

[3] "TV white spaces: A consultation on white space device requirements," Nov. 2012, http://stakeholders.ofcom.org.uk/binaries/consultations/whitespaces/summary/condoc.pdf.

[4] M. B. Weiss, S. Delaere, and W. H. Lehr, "Sensing as a service: An exploration into practical implementations of dsa," in *New Frontiers in Dynamic Spectrum, 2010 IEEE Symposium on*. IEEE, 2010, pp. 1–8.

[5] "Comments from the Communications Research Centre Canada: In the Matter of Unlicensed Operation in the TV Broadcast Bands," http://apps.fcc.gov/ecfs/document/view?id=6518724125.

[6] "In the Matter of Unlicensed Operation in the TV Broadcast Bands: Second Report and Order and Memorandum Opinion and Order," Federal Communications Commision, Tech. Rep. 08-260, Nov. 2008. [Online]. Available: http://hraunfoss.fcc.gov/edocs%5Fpublic/attachmatch/FCC-08-260A1.pdf

[7] S. M. Mishra, "Maximizing available spectrum for cognitive radios," Ph.D. dissertation, University of California, Berkeley, 2009.

[8] "Method for point-to-area predictions for terrestrial services in the frequency range 30 MHz to 3 000 MHz," International Telecommunications Commission (ITU), Recommendation ITU-R P.1546-3, 2007.

[9] Http://transition.fcc.gov/Bureaus/MB/Databases/tv_service_contour_data/readme.html.

[10] Http://sutrotower.com/tower-technology/.

[11] K. Harrison and A. Sahai, "Potential collapse of whitespaces and the prospect for a universal power rule," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 316–327.

[12] G. Giorgetti, S. K. Gupta, and G. Manes, "Wireless localization using self-organizing maps," in *Proceedings of the 6th international conference on Information processing in sensor networks*. ACM, 2007, pp. 293–302.

[13] L. Doherty, L. El Ghaoui *et al.*, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2001, pp. 1655–1663.

[14] N. Patwari, A. O. Hero III, M. Perkins, N. S. Correal, and R. J. O'dea, "Relative location estimation in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 8, pp. 2137–2148, 2003.

[15] B. Li, Y. Wang, H. K. Lee, A. Dempster, and C. Rizos, "Method for yielding a database of location fingerprints in WLAN," *IEE Proceedings-Communications*, vol. 152, no. 5, pp. 580–586, 2005.

[16] J. Riihijarvi, J. Nasreddine, and P. Mahonen, "Demonstrating radio environment map construction from massive data sets," in *Dynamic Spectrum Access Networks (DYSPAN), 2012 IEEE International Symposium on*. IEEE, 2012, pp. 266–267.

[17] P. A. Zandbergen, "Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning," *Transactions in GIS*, vol. 13, no. s1, pp. 5–25, 2009.

[18] R. Tandra and A. Sahai, "SNR walls for signal detection," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, no. 1, pp. 4–17, 2008.

[19] C. Phillips, D. Sicker, and D. Grunwald, "Bounding the error of path loss models," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 71–82.

[20] V. Atanasovski, J. van de Beek, A. Dejonghe, D. Denkovski, L. Gavrilovska, S. Grimoud, P. Mahonen, M. Pavloski, V. Rakovic, J. Riihijarvi *et al.*, "Constructing radio environment maps with heterogeneous spectrum sensors," in *New Frontiers in Dynamic Spectrum Access Networks (DySPAN), 2011 IEEE Symposium on*. IEEE, 2011, pp. 660–661.

[21] Https://github.com/kate-harrison/whitespace-eval.

[22] See http://factfinder2.census.gov/faces/nav/jsf/pages/index.xhtml.

[23] See http://www.census.gov/geo/www/cen_tract.html.

[24] See transition.fcc.gov/mb/video/tvq.html.

[25] K. Harrison, "Cognitive radios in the tv whitespaces: challenges and opportunities," 2011.