Measurement-Augmented Spectrum Databases for White Space Spectrum

Ayon Chakraborty and Samir R. Das
Computer Science Department
Stony Brook University
Stony Brook, NY 11747, USA
{aychakrabort, samir}@cs.stonybrook.edu

ABSTRACT

Spectrum databases used to estimate TV white space availability often provide inaccurate and largely conservative estimates as they are primarily based on empirical propagation models. This leads to 'loss' of white space spectrum that is critical in urban areas with large spectrum demand. While alternatives are possible in terms of incorporating direct spectrum measurements, the measurement locations must be judiciously chosen so that measurement effort is not prohibitive. Fundamentally, this boils down to addressing the estimation accuracy vs measurement effort question. We present a rigorous data driven analysis to address this using measurement data collected in parts of New York City metro area. We show that it is possible to develop models that estimate whether the current database estimates are reliable in a given location. Following this, we provide a recipe for developing a 'measurement-augmented' spectrum database that takes the help of measurements where needed and falls back on the current propagation model-based database technique in the rest of the areas. The final takeaway is that it is possible to improve database accuracy significantly with only modest amount of measurements.

Categories and Subject Descriptors

C.4 [**Performance of Systems**]: Design Studies, Measurement Techniques, Modeling techniques, Reliability, availability, and serviceability.

Keywords

TV Whitespace, Spectrum Databases, Propagation Model.

1. INTRODUCTION

According to the rules for white space (WS) spectrum access, 'secondary' (unlicensed or lightly licensed) use is allowed only when it does not adversely affect 'primary' (i.e., licensed incumbent) communications. To enable this, the recommended solution by the FCC in connection to TV band spectrum (so called TVWS)

is to use one of the approved *spectrum databases* [9].[1] All secondary users are required to consult such a database prior to accessing any WS channels for communication. The database uses (i) detailed knowledge of TV transmitter characteristics and (ii) specific propagation models to estimate the TV signal power at any given location. This in turn provides the 'protection contours' for each TV channel – secondary operation is allowed only outside these contours.

Clearly, the accuracy of such databases is crucial for successful operation in the WS. 'False negative' estimates by the database mean that WS cannot be exploited even when it does not hurt primary communications. This leads to lost spectrum opportunities. 'False positive' estimates, on the other hand, mean that primaries are not protected leading to violation of regulations. Recent measurement studies have reported that the spectrum databases are prone to varying degrees of inaccurate estimates [22, 23, 24, 19]. The inaccuracy is somewhat biased towards generating more false negative errors as opposed to false positives. This is not entirely unexpected as the databases 1) rely entirely on empirically derived propagation models and 2) these models are tuned to provide more conservative estimates of protection regions due to legal issues pertaining to false positive estimates (more on this in Section 3). The actual database question aside, multiple measurement studies in the TV bands also independently point out limitations of propagation modeling approaches in estimating availability of TVWS channels for secondary use [11, 16, 22, 23, 24, 19].

Lost White Space: The database inaccuracies lead to a serious loss of spectrum opportunity in *urban* regions, specifically in large metro areas that are also population hotspots: 1) a large number of TV channels in these regions mean that fundamentally WS availability is poor [12] and 2) the spectrum demand is large due to dense populations. Losing spectrum opportunity precisely where more spectrum is needed does not bode well for the business case for TVWS. Our measurement study in parts of the New York City metro area [6] shows that roughly 40%–75% of available WS spectrum is lost (Section 2).

1.1 Measurement Augmentation

The problem can be easily remedied if the spectrum databases can use actual measurements [23, 24] instead of standard propagation models.[2] This is viable as the primary occupancy in the TV bands do not change often and thus the measurements do not need to repeat frequently in the same location. However, the challenge here is to limit the spatial extent of such measurements as

---

[1]There are several such database providers, e.g., Spectrum Bridge [7], Google [1], i-Connective [2], Keybridge [4], etc.
[2]We do not discuss regulatory issues that can arise here; we limit ourselves to technical issues only.
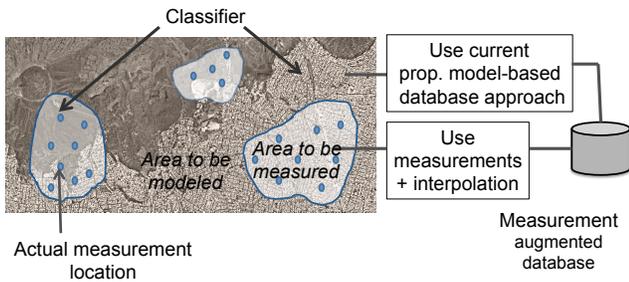
Figure 1: Overview of the proposed measurement-augmented spectrum database approach.

(a) New Jersey (NJ)

(b) Long Island (LI)

Figure 2: Database inaccuracies on a per-channel basis in the NJ and LI datasets. Note that false negatives are abundant, though false positives are rare.

measurements are expensive to the database operator.[3] Our goal in this work is thus to pose a fundamental question: *Can the spectrum database predictions be significantly improved with only a modest amount of measurements?*

We address this question with a data-driven analysis using spectrum measurements collected on 30 DTV channels at 5K+ locations in the NYC metro area - one of the major population hotspots in US (Section 2). Our specific contribution is a recipe to 'augment' current spectrum databases with actual measurement data. Unlike prior work that uses very extensive measurements to completely replace the current propagation model-based databases [23, 24], we propose a middle ground: measurements are used *strategically only in specific regions*; the rest of areas can still use the current database approach. This strikes a good balance between accuracy and cost in terms of measurement effort. Figure 1 provides a broad overview.
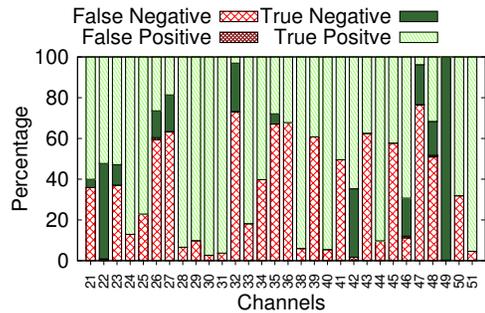
The centerpiece of the technique is a classifier model (Section 4) that predicts the accuracy of the current propagation model-based database estimates using factors that are poorly accounted for in these models (Section 3). This enables us to separate out the regions where the database accuracy could be poor. We use sparse measurements followed by standard interpolations in these regions, but stick to current database estimates in the rest of the areas (Section 5). Finally, we show that this approach as a whole *is able to recover a significant amount of lost white space with only modest measurement effort*.
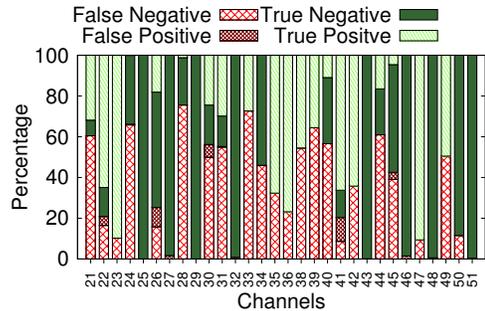
2. MEASUREMENTS

We conducted measurements in two regions within New York City metro area [6] – 1) in the cities of Hoboken, Jersey City, Newark and Summit in New Jersey (NJ) and 2) parts of Suffolk and Nassau counties in Long Island (LI) in New York covering approximately 250 and 150 sq. kms, respectively. All these regions are urban with significant population and economic activities. The measurements are taken from the road with a vehicle mounted spectrum sensor (ThinkRF WSA4000 [8]). We scanned the UHF spectrum that includes the DTV channels 21 through 51 (512 MHz - 698 MHz except channel 37 which is used for astronomical purposes). The measurement yielded about 150 K+ data points in aggregate for 5K+ locations sampled. Each data point consists of a particular channel's FFT information tagged with corresponding GPS coordinates.

To set a stage for further analysis, we first determine the WS spectrum availability based on these measurements. Specifically,

we use FCC's guidelines to determine whether WS channels are deemed available to *portable devices* hypothetically operating at the measurement locations [9]. For such devices, the FCC-mandated sensing threshold to detect a TV signal is quite low (-114 dBm). The noise in the spectrum sensor itself can make it challenging to detect weak TV signals. Even for the high-end spectrum sensor that we use, capturing a TV signal over a 6 MHz channel gives rise to overwhelmingly high noise floor compared to the required -114 dBm threshold. To address the high sensitivity needed, we leverage a salient feature of TV signals, viz., presence of a pilot signal at 309KHz offset above the lower edge of the 6MHz TV channel and the fact that the power of the pilot has a specific relationship to the channel's aggregate power.[4] We zoom into a very narrow band (first 488KHz of a TV channel) to detect the pilot's presence and power with a very low noise and use this to calculate the channel's aggregate power. A similar technique on the same sensor hardware was also used in [23, 24].

In Figure 2 we provide per-channel statistics of spectrum database inaccuracies assuming measurements as ground truth. Specifically, Spectrum Bridge [7] is used as the database provider.[5] In this figure, false positive (negative) indicates that the database predicts the channel to be available (not available) for secondary use, but the measurement detects presence (absence) of signal. True positive (negative) indicates that the database is correct in predicting the channel to be available (not available). The database is checked for each measurement location and the database esti-

---

[3]Various mechanisms are possible including deploying actual spectrum sensors [14], deploying sensors in public vehicles [23, 24] or drive tests like cellular operators, etc. All are expensive in terms of actual labor or cost.

---

[4]The pilot signal's power is less than the channel's aggregate power by approx. 11.3dB.

[5]Per FCC, all database providers should use the same modeling approach and thus should provide identical results.

mate for each channel is classified into one of the four categories – false/true positive/negative. The figure shows aggregated per-channel statistics. We make a couple of observations:

1. False negative predictions are frequent, averaging 35% and 32% for NJ and LI respectively. Also more than 70% of all measured locations show prediction error in atleast one channel. On the other hand, false positives are relatively rare (almost absent in NJ but present in a few channels in LI). This points to the general conservative nature of the database estimates leading to loss of WS. According to the measured data, the loss is staggering: *Approximately 75% and 40% of available WS is lost in NJ and LI respectively due to incorrect prediction by the database*, where the loss is measured as a fraction of channels inaccurately predicted to be unavailable, aggregated over all locations.

2. The observations in specific categories (false positive, etc.) are clustered and not randomly distributed.[6] This highlights the possibility that such regions can be identified in advance via an out-of-band modeling approach. This in turn provides the possibility of *advance partitioning of the entire area of interest into two types of regions – $Area_{Measure}$* (current database approach is unreliable and measurements are preferred) and *$Area_{Model}$* (database is reliable). See Figure 1. This forms the essential basis of the approach we present below.

3. ANALYZING PREDICTION ERRORS

We now follow up on the above observations by taking a deeper look at possible causes of the database inaccuracies. The first step is understanding what modeling approach is internally used in the spectrum databases and the second step is identifying factors responsible for erroneous estimates.

3.1 Modeling Approach in Databases

Per recommendation from FCC, the approved spectrum databases use a specific propagation modeling approach (called *F-curves* [20]) to identify the protection contours for TV channels. F-curves are statistical models derived from empirical measurements and are specified by operating band, effective radiated power (ERP), antenna height above average terrain (HAAT), antenna beam parameters, location and time percentage reliability requirements. For the latter, FCC specifies (50,90).[7] F-curves are used to estimate a contour boundary in each radial direction[8] such that the estimated field strength at the boundary point is 41dBu (this value is related to (50,90) above). These distances are again weighted by antenna patterns if they are not uniform.

One reason why F-curves overestimate the contour boundary is due to the use of 'average terrain' instead of fine-grained terrain information. However, in our experience more advanced approaches such as the Longley-Rice [13] that do use more granular terrain information do not appear to fare much better overall. We did an independent set of analysis to establish this. Figure 3 show scatter-plots for received power according to the Longley-Rice model and the measured power for two cases (channel available and not available). Note very poor correlation for the channel available case (TV signal absent), while some degree of correlation is indeed present when the channel is not available (TV signal detected).[9] Thus, Longley-Rice may not be very good in predicting white spaces.

---

[6] The analysis of spatial correlation in Section 5 demonstrates this.
[7] This means that the signal from the TV station is present at 50% of the locations for 90% of the times within the contour boundary.
[8] Radial HAAT is generally calculated every 1°azimuth.
[9] Not shown here, F-curves show poor correlation in both cases.
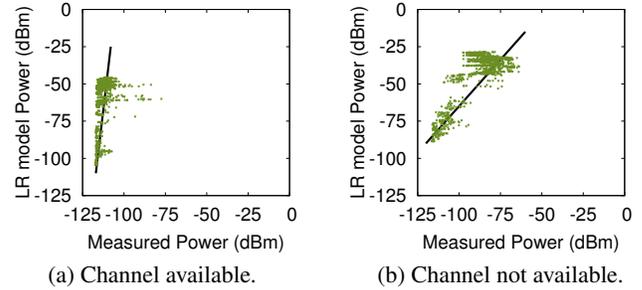

(a) Channel available. (b) Channel not available.

Figure 3: Modeled power according to Longley-Rice vs. Measured power. Note very poor correlation for the channel available case (TV signal absent).

3.2 Factor Analysis

With the above understanding, we hypothesize that certain propagation related parameters overwhelmingly influence the database accuracy. Our goal now is to explore this influence systematically. We identify three key parameters that have significant influence on the accuracy.[10]

Distance from Transmitter: For a given channel, this is the distance of the nearest TV transmitter operating in that channel. This parameter is important due to the fact that it is likely that the database will be able to predict accurately if the nearest transmitter is too close (channel not available) or too far (channel available) from the location.

Modeled Power: This represents the estimated received power for a channel as predicted by the FCC F-Curve Model. Just like above, too small or too large modeled power indicate that the database is likely accurate.

LOS Obstruction Length: This is the portion of the 'line-of-sight' from the closest TV transmitter at the specific channel that is 'obstructed' due to terrain irregularity. To understand this construct an imaginary straight line from the transmitter antenna to the location. Parts of this line could be 'obstructed' due to terrain elevations. We are interested in the total such obstruction length. To determine this, we use elevation profile from NASA's SRTM 1-arc-second terrain database (30m resolution) [5].

3.3 Types of Prediction Errors

Figure 4 shows the fraction of locations measured with respect to each of these parameters for specific accuracy categories (namely, false positive (FP), false negative (FN), true positive (TP) and true negative (TN)). NJ and LI data are presented separately for the purpose of illustration. To understand these plots, consider the database prediction – channel available or not available. These predictions can be correct or incorrect. Accordingly, channel available (not available) decisions contribute to FP and TN (FN and TP) categories. Thus, if the plot for FP is reasonably separate from that for TN, it is possible to 'learn' where the database is likely to make incorrect decisions for channel available. Similar is the case for channel not available. We do see that there is a separation between at least one of these pairs of plots for all three parameters and sometimes the separation is quite distinct. This indicates that these parameters could be effective in 'teasing apart' locations where the

---

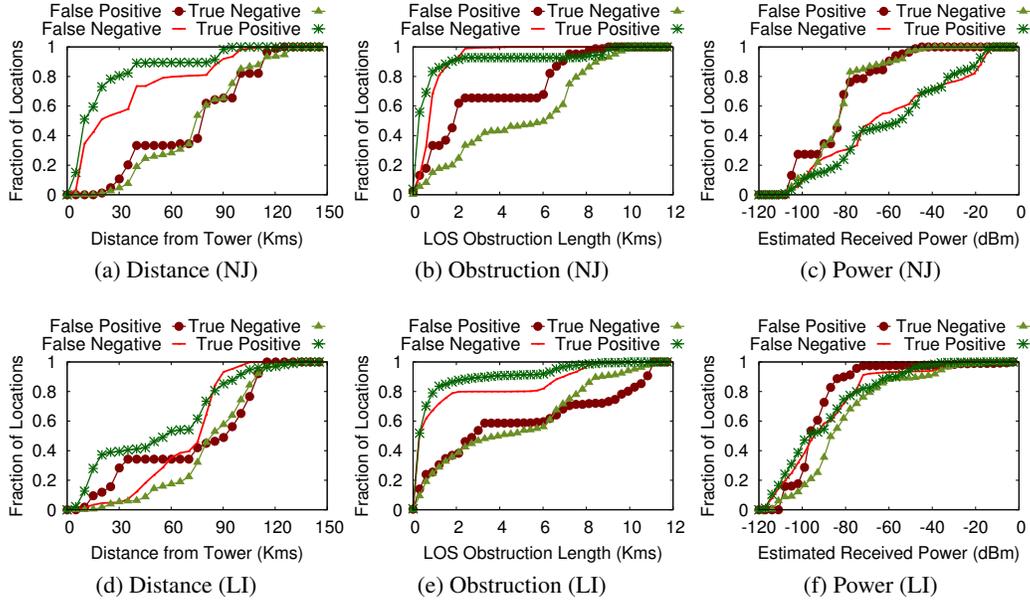[10] We explored several other factors like Antenna Height, ERP etc., but their influence seems lesser.

Figure 4: Influence of three factors, viz., Transmitter Distance, Line-of-Sight Obstruction lengths and Modeled Power on the prediction accuracies.
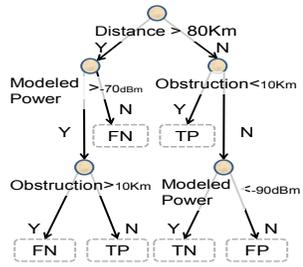

Figure 5: Example of a decision tree classifier.

database estimate is likely incorrect. Our general approach will be to use measurements in these locations as opposed to 'correcting' database estimates in some other way.

4. PREDICTING DATABASE INACCURACIES

We use the idea outlined in Section 3 to learn a classifier that takes the above three parameters as features[11] and produces a detection accuracy type (i.e., FP etc) as the output. While various approaches are possible, we use a Decision Tree Classifier (C4.5) as the decision boundaries are easily interpretable and thus manual corrections can be done more easily if necessary. A Decision Tree classifier produces *rule-like constructs*: by way of example, if the distance to the transmitter is more than, say, 100 km and the estimated received power is less than, say, -80 dBm, then it is highly likely the database is going to be inaccurate. See Figure 5 for an example.

Challenges: There are several challenges in building a useful classifier. First, we need to make a decision whether we should learn a separate classifier for each channel. While a separate classifier

for each channel may handle frequency specific effects better than an aggregated classifier over all channels, the former is likely to be somewhat region specific. If in a different region, where TV channel occupancy is very different, channel specific classifier may not work well. We have investigated this aspect at length (not reported here for brevity) and finally settled on the aggregate classifier. It is only incrementally worse for the two regions we studied (NJ and LI), but is much superior (to channel-specific) when the model is learnt in one region and applied to the other. Second, even within a region, the amount of training data must be minimized to reduce measurement effort.

To analyze the above issues, we use a series of cross-validations. We use small contiguous portions (as opposed to random)[12] of the measurement data for training and run k-fold cross-validations where the folds are different contiguous partitions of the data set. We also run cross-validations across the two regions – NJ and LI. Figure 6 summarizes the results in terms of the expected classifier accuracy.[13]

Classifier Accuracy: There are two broad takeaways from Figure 6. 1) When the model is tested in the same region, the accuracy is very good with even with small amount of training. Just about 10% training data achieves about 85% accuracy. The cross validation results across regions (model learnt at NJ and applied to LI and vice versa) are also encouraging – roughly about 70% accuracy with about 20% data. Note that the straight line distance between these regions is roughly 150 km. Though more study from diverse geographic region is needed to prescribe a guideline for exactly how much to train, these results show a significant opportunity. In contrast, existing approaches require training per 100 m [23] thus requiring several orders of magnitude more effort.

---

[11]Note that these parameters come from independent data sets and do not require measurements.

[12] Contiguous data collection is much easier than random.

[13]Note that this is distinct from the database accuracy. This simply looks at whether the classifier is correctly classifying a test data in the four categories: FP, FN, TP, TN.
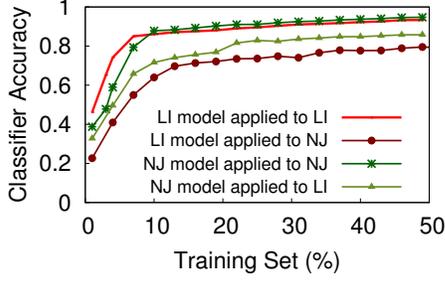
Figure 6: Decision tree classifier accuracy for predicting database errors.

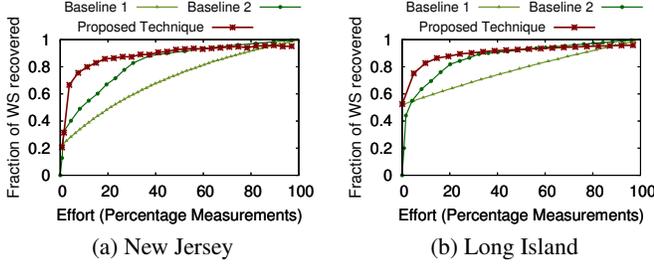

(a) New Jersey (b) Long Island

Figure 7: Accuracy vs. measurement effort.

5. MEASUREMENT AUGMENTATION

We run the Decision Tree classifier trained with 10% of the measured data on the NJ and LI data sets to partition all locations into two subsets - $Area_{Model}$ and $Area_{Measure}$ (details in Algorithm 1). In subset $Area_{Model}$, the database accuracy is high (TP and TN). In subset $Area_{Measure}$ it is low (FN and FP) and thus predictions from these locations preferably should come from actual measurements. Further all locations in $Area_{Measure}$ do not need to be strictly measured – spatial subsampling followed by a suitable interpolation technique can be used to reduce the measurement effort (details in Algorithm 2). There is already a growing literature on this topic specifically targeting radio environment mapping [10, 18, 17].

Accuracy vs. Effort: Figure 7 shows overall accuracy vs effort for a couple of baseline mechanisms along with the proposed approach. Instead of showing the channel availability prediction accuracy (as in Figure 2) we show fraction of available channels correctly identified. The 'effort' in the plots here is measured in terms of number of locations that are actually measured (normalized as a fraction of all locations). The following cases are evaluated:

(i) *Baseline 1:* A random set of locations use measurements and rest use the current database for prediction.

(ii) *Baseline 2:* A random set of locations use measurements and the rest uses interpolation for prediction.[14]

(iii) *Proposed technique:* Briefly, for each channel the classifier is used first to partition the universe of locations into $Area_{Model}$ and $Area_{Measure}$. Then K locations are picked that belong to $Area_{Measure}$ for the most number of channels. These are actual measurement locations and K is indicative of measurement effort. Rest of the locations in $Area_{Measure}$ for each channel use

---

[14]We use a simple linear interpolation technique called Inverse Distance Weighting or IDW [3]. [10] has shown linear interpolation techniques can achieve almost as good performance as more sophisticated geostatistical techniques such as Ordinary Kriging [18].
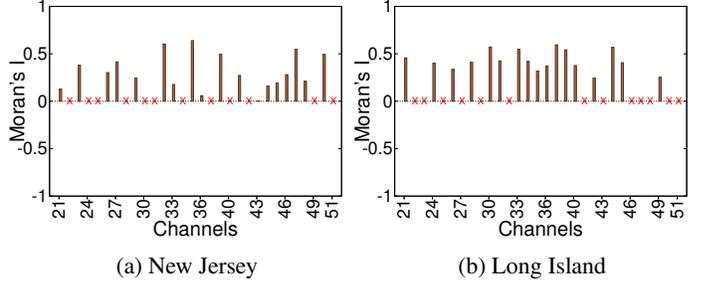

(a) New Jersey (b) Long Island

Figure 8: Moran's-I statistic representing the spatial autocorrelation among the measurement locations shown in a per-channel basis. (The statistic is not calculated for channels lacking enough data points.)

interpolation using the available measurements as in (ii) above. All other locations (i.e., $Area_{Model}$) use standard propagation modeling as in current databases. Algorithms 1 and 2 have a more formal presentation.

Algorithm 1 Measurement area calculation for all channels

1: $L \leftarrow$ All discreet locations in *rasterized* area /*30m x 30m cells*/
2: $C \leftarrow$ All Channels under consideration /*21 through 51, except 37*/
3: **for all** $l \in L$ **do**
4:     **for all** $c \in C$ **do**
5:         $Feature_{l,c} \leftarrow < Distance_{l,c}, Obstruction_{l,c}, ModelPower_{l,c} >$
        $decision \leftarrow CLASSIFIER(Feature_{l,c})$ /*Decision Tree*/
6:         **if** $decision = TP$ or TN **then**
7:             Add location l to $Area_{Model,c}$
8:         **else if** $decision = FP$ or FN **then**
9:             Add location l to $Area_{Measure,c}$
10:         **end if**
11:     **end for**
12:     Add $Area_{Measure,c}$ to $Area_{Measure}$
13: **end for**

Algorithm 2 Channel occupancy decisions

1: $K \leftarrow EFFORT$ /*Number of measurement points*/
2: $M_K \leftarrow$ Most occurring K locations from $Area_{Measure}$
3: $Powers(M_K) \leftarrow$ Set of measured powers at all points in M_K
4: **for all** $l \in L$ **do**
5:     **for all** $c \in C$ **do**
6:         **if** $l \in M_K$ **then**
7:             $Power_l \leftarrow Measure(l)$ /*use actual measurements*/
8:         **else if** $l \in Area_{Measure,c}$ **then**
9:             $Power_l \leftarrow Interpolate_{IDW}(Powers(M_K))$ /*interpolate such a point from measured points*/
10:         **else**
11:             $Power_l \leftarrow PropagationModel(l, c)$ /*use existing spectrum database propagation model to estimate received power*/
12:         **end if**
13:         $decision_{l,c} \leftarrow Rules_{FCC}(Power_l)$ /*apply necessary rules to determine spectrum occupancy*/
14:     **end for**
15: **end for**

Figure 7 presents the accuracy vs effort of the three techniques above based on the measurement data. Note that with 0% effort (no measurement), the overall performance is the same as the current databases (roughly 25% (NJ) and 60% (LI) of available channels correctly predicted). The two baseline techniques show incremental improvement with more measurements. However, the proposed approach improves in accuracy quickly with more effort, reaching over 80% with only about 10% effort.
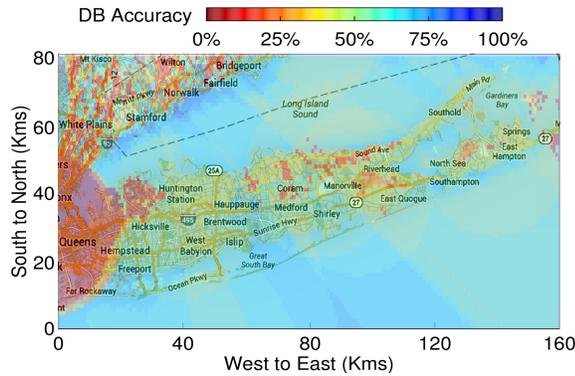
Figure 9: Heatmap showing estimated database accuracy in an extended region. The region shows eastern part of NYC, Long Island, parts of upstate NY and CT.

Spatial Clustering: Figure 7 does not show how clustered the actual measurement points are. In fact, it may appear that Baseline 2 does not do too poorly as it converges quickly as well - though at a lower rate than the proposed technique. However, in Baseline 2 the locations are random, while they are clustered in the proposed technique. This saves on actual measurement effort if drive tests are used. To demonstrate this, we show the spatial autocorrelation property of the measurement points in terms of 'Moran's I' in Figure 8. Moran's I [15] is a commonly used measure of spatial autocorrelation. A concept of distance is used to indicate proximity and is used as 'weights.' Moran's I is defined as:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_i (x_i - \overline{x})^2}, \quad (1)$$

where x is the random variable studied, $\overline{x}$ being the sample mean, x_i's are the observations. w_{ij} is the weight associated with each pair (x_i, x_j) and N is the number of observations. In our case, x_i is either 1 or 0 (depending on whether it is a measurement point), and w_{ij} is the distance between points x_i and x_j. We see that Moran's I is always > 0 and often substantial – between 0.3-0.6, demonstrating significant amount of spatial auto-correlation. On the other hand, in Baseline 2, Moran's I is 0 by definition.[15]

5.1 Scale-up Study

It will be interesting for the reader to understand what really happens if we scale up the area of study. To do this, we take a fairly large, 160 km × 80 km wide region covering part of the New York City metro area [6] and use the decision tree classifier (we use the LI based classifier as majority of the scale-up area lies in LI) to estimate the database accuracy for each channel throughout this region. Figure 9 shows the accuracy in the form of a heatmap. The accuracy numbers here indicate the percentage of all channels that the classifier predicts to be correctly estimated (TP or TN) by the database. Note the poor accuracy throughout NYC and immediately surrounding areas (large reddish patch on the left). Note also the general clustering of poor accuracy areas providing a guide where measurements should preferably be taken. Also, of note the large swaths of blueish high accuracy areas in most parts of Long Island and surrounding water. This primarily arises to a relatively flat terrain there relative to upstate NY and CT where a patchwork of low accuracy areas dot the map (see the northwest corner). An operator interested in providing augmented database service in this region will need to measure only the most red areas either accord-

---

[15]Moran's I can range from -1 to +1; -1 indicates perfect disperson, 0 indicates random, +1 indicates perfect correlation.

ing to a budget constraint, or using a heuristics. A simple heuristics could incrementally increase the measured area until the interpolation performance converges.

6. RELATED WORKS

Several recent measurement papers study limitations of propagation modeling approaches in detecting channel availability in TVWS. For example, authors in [16] consider several models, viz., Free-Space, Egli and Longley-Rice and show that while the former models perform poorly Longley-Rice is significantly better when used with terrain information. However, measurements reported in [11, 19] have poor experience with Longley-Rice much like our work. In [22], authors show that available TVWS in indoor vs outdoor locations could be significantly different demonstrating potential inability of current database systems to expose indoor WS.

Spatial interpolation techniques for radio environment mapping have been studied with a goal to adequately utilizing sparse spectrum sensing (e.g., [18, 10, 21]). Specifically, in [10] authors show that very few sample points may be adequate when coupled with an appropriate spatial interpolation technique. However, these points are not necessarily clustered and so drive tests to measure these locations can still incur a significant cost. The vScope system [23, 24] proposes leveraging public transport for spectrum measurements in TVWS. However, vScope is entirely opportunistic and measures all locations where the contracted public transport (with on-board sensor) travels. While the power of interpolation techniques has been demonstrated in [10, 21], neither [10, 21] nor [23, 24] give clear guidelines about the actual choice of measurement locations for doing an interpolation. Our approach remedies such shortcomings.

Finally, though not directly related to white space, SpecNet [14] developed a prototype to collect real time measurements from spectrum analyzers around the world put into a networked system. While such a system may be useful in connection to what we proposed here, the deployment must be strategic and much denser than proposed in the original paper.

7. CONCLUSIONS

If WS spectrum is to succeed commercially, inaccuracies in spectrum databases must be addressed. While previous work has recognized spectrum measurements as a way to address this issue, our work specifically develops mechanisms to reduce such measurement effort. The proposed approach relies on the current propagation model-driven databases where they are likely to be accurate and uses strategic measurements where they are not. The experimental results in parts of New York City metro area show that doing only modest amount of measurements – about 10% (25%) of the area studied that are also reasonably clustered – can recover above 80% (90%) of available WS.

Open issues do remain that we are pursuing in our ongoing work. They include 1) understanding how much data is needed to train the proposed classifier for a very large region, e.g., an entire state or country; 2) optimizing actual driving routes for measurements instead of number of locations as done in this paper.

8. ACKNOWLEDGMENTS

9. REFERENCES

[1] Google Spectrum Database. `http://www.google.com/get/spectrumdatabase/`.

[2] Iconectiv White Space Database. `http://www.iconectiv.com/spectrum-mgmt/white-spaces/index.html`.

[3] Inverse Distance Weighting. `http://en.wikipedia.org/wiki/Inverse_distance_weighting`.

[4] Keybridge White Space Database. `https://keybridgeglobal.com/whitespace/`.

[5] NASA SRTM Data,. `lta.cr.usgs.gov/SRTM2`.

[6] New York City Metro Area. `http://en.wikipedia.org/wiki/New_York_metropolitan_area`.

[7] Spectrum Bridge website. `http://spectrumbridge.com`.

[8] ThinkRF Spectrum Analyzer. `thinkrf.com/`.

[9] Second report and order and memorandum opinion and order in the matter of unlicensed operation in the TV broadcast bands. FCC ET Docket 08-260, Nov. 2008.

[10] A. Achtzehn, J. Riihijarvi, and P. Mahonen. Improving accuracy for TVWS geolocation databases: Results from measurement-driven estimation approaches. In *Proc. IEEE DySPAN*, 2014.

[11] A. Achtzehn, J. Riihijarvi, G. M. Vargas, M. Petrova, and P. Mahonen. Improving coverage prediction for primary multi-transmitter networks operating in the TV whitespaces. In *Proc. IEEE SECON Conference*, 2012.

[12] K. Harrison, S. M. Mishra, and A. Sahai. How much white-space capacity is there? In *Proc. IEEE DySPAN*, 2010.

[13] G. Hufford, A. Longley, and W. Kissick. A guide to the use of the Irregular Terrain Model in the area prediction mode. Technical report, NTIA, 1982.

[14] A. Iyer, K. K. Chintalapudi, V. Navda, R. Ramjee, V. Padmanabhan, and C. Murthy. SpecNet: Spectrum sensing sans frontieres. In *Proc. NSDI*, 2011.

[15] P. A. P. Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37:1733, 1950.

[16] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl. SenseLess: A Database-Driven White Spaces Network. *IEEE Transactions on Mobile Computing*, 11(2):189–203, Feb. 2012.

[17] C. Phillips. *Geostatistical Techniques for Practical Wireless Network Coverage Mapping*. PhD thesis, University of Colorado, 2012.

[18] C. Phillips, M. Ton, D. Sicker, and D. Grunwald. Practical radio environment mapping with geostatistics. In *Proc. IEEE DySPAN*, 2012.

[19] A. Saeed, K. A. Harras, and M. Youssef. Towards a characterization of white spaces databases errors: an empirical study. In *Proc. ACM WinTech*, 2014.

[20] White Space Database Administrator Group. Channel Calculations for White Spaces - Guidelines (1.29). `http://spectrumbridge.com/wp-content/uploads/2014/08/Database-Calculation-Consistency-Specification.pdf`, 2013. Date accessed: 2014-10-17.

[21] X. Ying, C.-W. Kim, and S. Roy. Revisiting TV coverage estimation with measurement-based statistical interpolation. `http://www.dynamicspectrumalliance.org/assets/Revisiting_TV_Coverage_Estimation_with_Measurement-based_Statistical_Interpolation_08202014.pdf`, 2014. Date accessed: 2014-10-17.

[22] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chenanant, and R. Chandra. Exploring indoor white spaces in metropolises. In *Proc. ACM MobiCom*, 2013.

[23] T. Zhang and S. Banerjee. Inaccurate spectrum databases?: Public transit to its rescue! In *Proc. ACM HotNets*, 2013.

[24] T. Zhang and S. Banerjee. A Vehicle-based Measurement Framework for Enhancing Whitespace Spectrum Databases. In *Proc. ACM MobiCom*, 2014.