A Vehicle-based Measurement Framework for Enhancing Whitespace Spectrum Databases

Tan Zhang and Suman Banerjee
Department of Computer Sciences
University of Wisconsin-Madison, USA

1. PROBLEM AND MOTIVATION

Many users have the annoying experience of their browsers getting stuck, video streaming becoming slow and choppy, and navigation services no longer accessible, especially in big cities like Chicago and New York. Most of these common frustrations come from a spotty wireless connection. Despite more and more cellular towers being deployed today, often heard from wireless providers is the concern that the surging demand for ubiquitous connectivity, intrigued by smartphones, tablets and other mobile devices, is outpacing their network capacity [9]. A key communication resource being exhausted is the *radio frequency spectrum* over which the airwave of a wireless signal propagates. To address the ongoing spectrum crunch, various regulatory agencies across the world intend to open up additional spectrum bands for wireless connectivity.

The transition from analog to digital television broadcast rolling out worldwide frees up many TV channels previous in use. The vacated channels in the television band (512 - 698MHz), popularly referred to as TV whitespaces, are recently released in the US and several other countries for unlicensed and opportunistic usage. Since a digital broadcast can pack multiple programs in a 6MHz TV channel, a substantial amount of spectrum (up to 180MHz [1,7]) is available in TV whitespaces. As illustrated in Figure 1, the total bandwidth of unused TV channels is usually greater than the sum of the 2.4GHz WiFi band and the 700MHz LTE band. In addition, a TV-band signal can achieve a superior propagation range (>2km [2,18]) over WiFi signals. Hence, this emerging spectrum has been explored in a host of research prototypes [1,2,13,18] and commercial deployments [15] to support diverse applications.

In a whitespace network, the unlicensed, secondary devices are required to only operate in vacant channels in absence of primary devices. These primary devices belong to the owners of some frequency bands, who have purchased a license for their exclusive usage. There are two types of primary devices in the TV band, i.e., TV broadcasts and wireless microphones. To determine vacant TV channels, a FCC-preferred mechanism [5] is to have whitespace devices to query a *spectrum occupancy database*. These databases are operated by some third-party companies following the FCC's guideline. They leverage a well-known propagation model to predict the coverage boundary of each primary device as illustrated in Figure 2. For a whitespace device being outside the estimated boundaries of all the primary devices, the databases will conclude the corresponding channel to be whitespace and vice versa.

Whitespace database limitations: Existing databases have three major limitations that hamper the efficient use of TV whitespaces. First, the databases rely on a fixed, conservative propagation model to predict for all kinds of environment, causing under-utilization of whitespace spectrum over large area (<71%). Second, the databases have no attempt to distinguish the quality of whitespace channels,

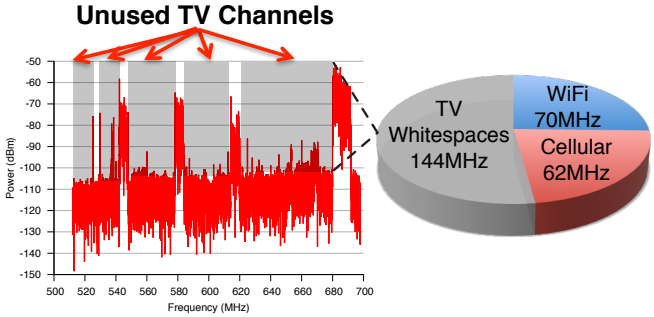

Figure 1: TV whitespace spectrum (in gray color) observed atop a building at Madison, WI. The amount of spectrum is more than the sum in the 2.4GHz WiFi band and the 4G cellular band.

which can differ significantly at a given location. This is caused by the uncoordinated interference among secondary devices sharing a same channel, along with the leakage from TV transmitters into their neighboring channels. The low-quality channel can deteriorate the performance of a whitespace network operating therein. Finally, the databases are not responsible for validating the locations reported by primary and secondary devices. Such location information is used for determining spectrum availability, and can be inaccurate due to various reasons such as incorrect entries to databases, transmitter reallocation, etc. The imprecise location information could lead to wasted communication opportunities, or even worse, harmful interference to primary devices.

V-Scope Approach: We explore the use of spectrum measurements to augment databases for overcoming above limitations. As shown in Figure 2, the key idea of our solution is to leverage the measured powers of a primary device to tune a propagation model (e.g., the slope α). The refined model can be used to better predict the coverage of primary devices, thereby reducing the wasted area for whitespace utilization. In addition, we refine propagation models for different secondary devices, while developing adjacent-channel models to predict the leakage of TV broadcasts. These models can be used together to estimate the quality of whitespace channels. Finally, we can leverage the observed propagation trend to "backtrack" the location of any TV-band transmitters.

To make the measurement-refined models widely applicable, we need to collect spectrum measurements over a large area – a fairly challenging and laborious task. To reduce the measurement effort, we present the design of a vehicle-based measurement framework called V-Scope. As suggested by the very name, V-Scope leverages public transit buses to carry spectrum sensors and collect measure-

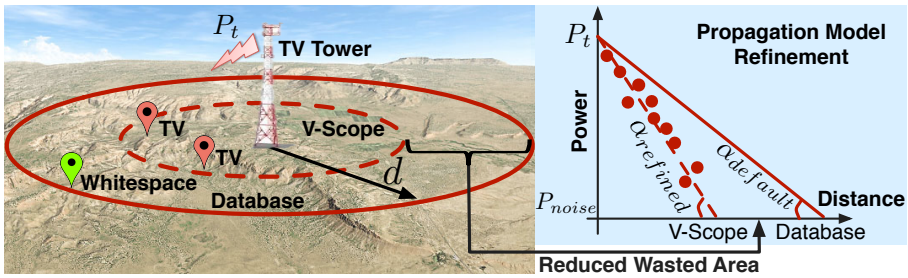

Figure 2: Databases use a default propagation model to predict the coverage of primary devices, blocking the use of whitespace spectrum in certain area. V-Scope reduces the wasted area by refining their propagation model with measurements.

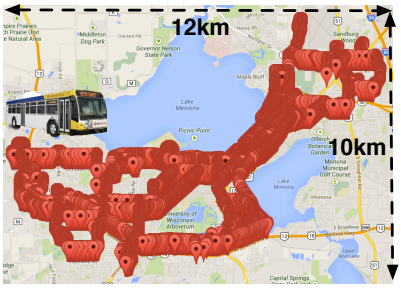

Figure 3: A snapshot of spectrum measurements collected from a metro bus over 120 square kilometer area. Each marker is a measured location.

ments opportunistically as they travel [1]. Through a 6-week deployment on a single metro bus at Madison, WI, we have been able to collect measurements at more than one million distinct locations over 120 sq. km. area as shown in Figure 3. We utilize these measurements with a suite of measurement and model construction techniques to i) better predict whitespace spectrum, ii) estimate the quality of whitespace channels, and iii) localize primary and secondary devices.

2. BACKGROUND AND RELATED WORK

Whitespace determination: Whitespace devices are expected to use one of the following two approaches to detecting primary transmission [5]. The initial approach is to have each whitespace device "sense" the primary activities based on local spectrum measurements. Such an approach can largely increase hardware cost and protocol overhead of a whitespace network, especially due to the extremely low detection threshold (i.e., -114dBm). Prior systems [1, 6–8] attempt to detect primary signals based on their spectral features, but fail to achieve the mandated threshold due to the noise in a spectrum analyzer. V-Scope uses a zoom-in pilot tracking technique (§ 3.1) to reduce the noise floor, thereby successfully fulfilling the above detection requirement. It further shifts the burden of primary detection to a few vehicle-mounted sensors in a dedicated measurement infrastructure.

Considering the difficulty and inefficiency of local spectrum sensing, the FCC advocates the alternative approach based on spectrum occupancy databases. Senseless [7] demonstrates the first database-driven whitespace network. The proposed database uses a Longley-Rice model augmented with terrain data, and is reported to incur low loss of whitespace spectrum. In the same year, the FCC mandates commercial databases to use an alternative, widely-used model (R6602 [5]). However, we find these databases to cause under-utilization of certain whitespace channels over a wide area, possibly due to the inability of capturing the fading and shadowing in an urban environment. To reclaim this spectrum wastage, V-Scope refines propagation models for the detected primary devices (§ 3.2), which can be utilized by the databases to better predict whitespace spectrum. Calibrating propagation models with measurements has been explored before [10, 11]. However, prior approaches require measurement locations to be carefully chosen, which can hardly be guaranteed in vehicular sensing. V-Scope leverages weighted regression to compensate the non-uniform measurement density, thus improving the accuracy of fitted models.

Estimating whitespace channel quality: Existing databases are not designed to predict the quality of whitespace channels. Instead, whitespace networks [1, 13] are expected to perform this task by leveraging spectrum sensors equipped with their end devices, which would again increase the hardware cost and protocol overhead. V-Scope enables the databases to estimate channel quality on behalf of whitespace networks, thereby obviating the sensing overhead. Such a function is realized by refining propagation models of secondary devices (§ 3.2) and constructing novel leakage models (§ 3.3) based on spectrum measurements collected from a few vehicular sensors.

Localizing TV-band devices: Existing databases do not attempt to validate the locations reported by primary and secondary devices. V-Scope uses the measured power of a device to pinpoint its location, thereby providing database operators additional means to validate such information and detect spectrum violators. Prior RSSI based localization techniques [3, 12] use all the measurements *indiscriminately* to construct a propagation model, which can be biased by environmental variation leading to large localization error. V-Scope enhances this approach by carefully selecting measurements from certain radiation sectors of a transmitting device, while constructing a sector-specific model to improve accuracy (§ 3.4).

3. APPROACH AND UNIQUENESS

Our proposed system consists of a server and multiple clients as shown in Figure 4. Each client is mounted on a (public) vehicle and uses a spectrum analyzer to collect spectrum samples in all the TV channels. Based on the captured spectrum, the client detects different types of primary and secondary devices, and estimate their power in real-time. It uploads the detection results and the measurement locations to our server over some wireless networks. Upon receiving these measurements, the server refines propagation models for each primary and secondary device, and constructs a leakage model for each TV broadcast. The databases can use the refined propagation models of primary devices to better determine whitespace spectrum at any location, while using the models constructed for secondary devices and the leakage model to predict the interference power in whitespace channels.

We next highlight some design features in V-Scope for addressing the two major challenges in augmenting databases. The first is to accurately detect weak primary signals and measure their power as will be discussed in § 3.1. The second is to efficiently utilize these (vehicular) measurements for constructing various models, which will be covered in later subsections.

[1] Of course, public transit buses are just one of many possibilities for vehicles that can carry spectrum sensors; other potential examples are mail delivery trucks, taxicabs, and many other third party services that scour different city roads.

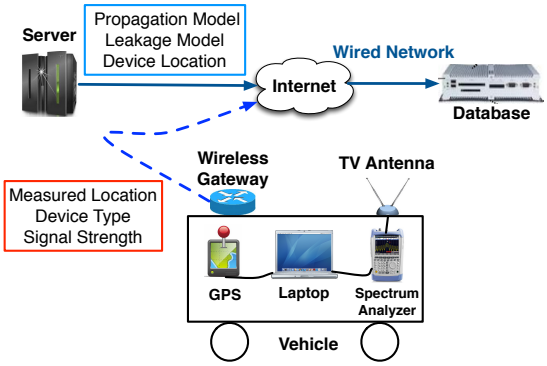

Figure 4: System architecture of V-Scope.

3.1 Zoom-in pilot tracking algorithm for primary detection

A client performs the dual functions based on the spectrum samples of a TV channel, i.e., detecting primary and secondary signals, and measuring their power. This power information is used to establish the ground-truth spectrum availability for evaluating the accuracy of databases, and to refine their propagation model (§ 3.2). We find this task to be particularly challenging in detecting primary signals due to the stringent detection threshold (-114dBm) imposed by FCC [5]. Figure 5 illustrates this challenge and our unique solution for measuring TV signals. The same concept can be applied to detecting licensed microphones as well.

As shown in Figure 5 (left), a spectrum analyzer uses a finite number of sampling bins (FFTs) to represent the spectrum of a TV channel. Each bin has an amplitude equal to the total power integrated over its spectral interval. This power is contributed not only by a primary signal, but also by the thermal noise in a spectrum analyzer. The noise power is generally higher than that of a weak primary signal, leading to a flat spectrum that represents the envelop of overwhelming noise (noise floor) rather than this signal.

To detect this primary signal beneath noise, we develop a *zoom-in pilot tracking* algorithm that hinges on two key observations. First, the spectrum of a primary signal has certain parts that can be detected more easily than others. For example, a TV signal presents a peak at the beginning part of its spectrum called pilot, which is more prominent from noise as illustrated in Figure 5 (left). Second, collecting spectrum at a narrower bandwidth reduces the noise in each sampling bin because *lesser* noise is aggregated over a *finer* spectral interval. Exploiting both facts, we configure a spectrum analyzer to capture only a fraction of a TV channel around the pilot frequency. This effectively reduces the noise floor in the zoom-in spectrum (Figure 5 (right)), producing a peak feature that can be used to reliably detect TV signals at the mandated threshold.

The detected pilot seems insufficient for estimating the power of a TV signal. This is because a TV signal has a much wider band than a zoom-in spectrum. To include the power of the uncovered spectrum, we observe a fixed power offset (≈15dB) between a TV signal and its pilot, which has been specified by TV standards [4]. This offset can be added to the pilot power (i.e., sum of those peak bins) to derive the total power of a TV signal. We then compare this total power with the -114dBm threshold to determine whitespace channels. We find this algorithm can achieve above 95% accuracy in detecting different types of primary signals at a wide range of power (up to -120dBm), and take less than a second to process all the 30 TV channels at a location.

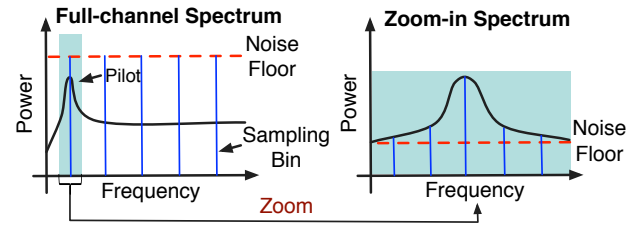

Figure 5: Illustration of TV signal detection. A narrower band of spectrum is captured around a TV pilot (left), producing a clear peak from reduced noise floor (right).

3.2 Region model for predicting in-band power

Using the power measurements of a device, V-Scope refines the parameters of a given propagation model. This model can be used to accurately predict its signal strength at any given location. Most of propagation models are based on a well-known, *linear* trend of the decaying signal strength over an increasing propagation distance. What is unknown is the slope and intercept (α, ϵ) of this linear function, which depend on a specific propagation environment. Thus, V-Scope use measurements of a device to determine these parameters such that the tuned model matches best the measurements in a target environment. This can be achieved by least-square regression, with the objective of minimizing the sum of the differences between the measured power and the predicted power at each measured location, called fitting errors. V-Scope leverages the following techniques to improve this regression based approach as illustrated in Figure 6.

Region-specific model: V-Scope first groups measurements into road segments and calculates an individual set of model parameters for these segments (Figure 6 (a, b)). The motivation is that these regions are likely to have different propagation characteristics due to environmental variation. By tuning a separate model for each region, this region-specific model can capture local environment better than a global model comprising a single set of parameters.

Weighted regression fitting: In refining a model for each region, we need to compensate the non-uniform density of measurements collected from (public) vehicles driving at a varying speed. The sparsely measured area has fewer measurements contributing to the sum of fitting errors, causing a biased model making large prediction errors in this area. To remove this fitting bias, we develop a weighted fitting algorithm that assigns a different weight (importance) to each measurement. A higher weight can inflate the fitting error of a measurement. These higher weights are therefore assigned to sparser measurements such that a same *weighted sum* of fitting errors is achieved across all the locations (as indicated by the circles in Figure 6 (c)). This can lead to an unbiased model predicting all the locations equally well.

3.3 Adjacent channel model for predicting TV broadcast leakage

Built on the region models of TV broadcasts, V-Scope can further predict their leakage in the adjacent channels caused by imperfect filtering [2]. The core intuition of our solution comes from an inherent power relationship between a TV signal and its leakage at any reception location. As illustrated in Figure 7, a TV signal and its leakage can be viewed to be emitted from two *collocated* signal sources. The power of these sources differs by a constant δ, which depends on the filtering characteristics of a TV transmit-

[2]The adjacent channels refer to the two immediate neighboring channels that are above and below a TV broadcast channel.

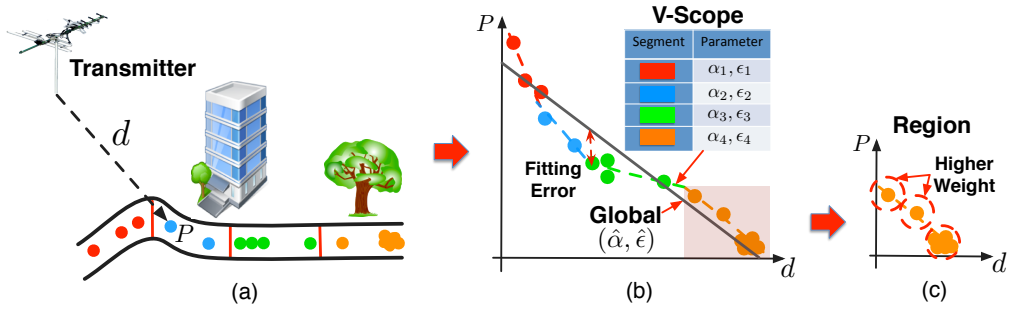

Figure 6: Illustration of V-Scope's model refinement procedure. The measurements of each device are first binned into road segments (a). An individual set of model parameters is calculated for these segments (b). In solving these parameters, different weights are assigned to measurements for compensating their non-uniform density (c).

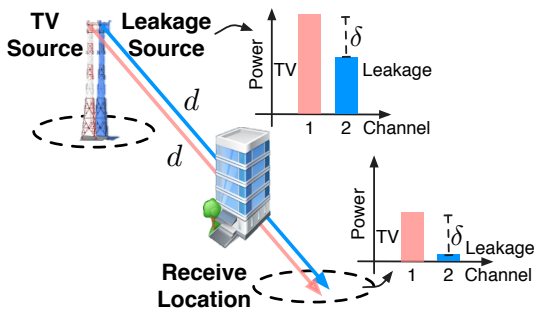

Figure 7: Illustration of predicting the adjacent-channel leakage of a TV broadcast. A TV signal and its leakage have a constant power offset δ at any reception location, which can be determined by measurements.

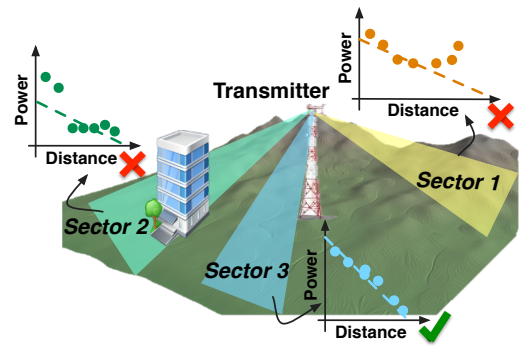

Figure 8: Illustration of the sector-based localization procedure. The measurements are partitioned into radiation sectors. Only measurements in sector 3 are selected for localization because they present a better linear propagation trend.

ter. To any reception location, the emitted signals traverse along a *same* path, hence experiencing similar attenuation imposed by the environment. As a result, the received power of these signals also differs by δ. V-Scope estimates δ for each TV broadcast using the measured in-band and adjacent-channel power. The estimated parameter will be added to the (predicted) power of a TV signal at a given location to derive its leakage power. The leakage power is combined with the inference power from secondary devices to calculate the total noise power in a whitespace channel as its quality.

3.4 Sector based localization

In constructing above models, the location information about the transmitting primary and secondary devices is needed for calculating the propagation range of measurements. When such information is not available for a device or awaits validation, V-Scope leverages an extended version of propagation model fitting to solve its location and those propagation parameters (α, ϵ). To do so, it first selects measurements in certain *radiation sectors* as illustrated in Figure 8. This is motivated by the common observation that the surrounding environment can significantly perturb the large-scale propagation trend observed in outdoor measurements. For example, the measurements in Sector 1 and Sector 2 match poorly with the expected linear propagation trend due to terrain elevation and obstacles respectively. These noisy measurements could bias the estimated distance to a target transmitter, introducing large localization error (up to 100m in § 4.3). Fortunately, we find the environmental effect to be unlikely to affect all the transmission directions of a signal. Thus, V-Scope only selects a few sectors (e.g., Sector 3) with measurements that present a highly linear trend for localization. It further assumes a different set of propagation parameters for each chosen sector to take account their distinct propagation characteristics. The transmitter location and all these model parameters are solved using an optimization method. We will show in § 4.3 that the combination of the sector-based measurement selection and model fitting can significantly improve the localization accuracy.

4. RESULTS AND CONTRIBUTIONS

In this section, we highlight some representative results about V-Scope based on our vehicular measurements collected over a 120 square-km area. We quantify the gain of V-Scope over a FCC-approved database [14] that predicts whitespace channels based solely on a propagation model. Overall, we find V-Scope can reclaim spectrum wastage of the database at up to 59% locations, select all the suitable whitespace channels at 72 – 83% locations, and localize a whitespace device with an error about 27m.

Methodology: We apply a 5-fold cross-validation by using 80% randomly chosen measurements to construct all the V-Scope models. The fitted models are used to predict the power of TV broadcasts, whitespace devices and adjacent-channel leakage at the remaining measured (testing) locations. We compare the predicted power of TV broadcasts with -114dBm to determine whitespace channels, while using the predicted power of secondary devices and adjacent-channel leakage to estimate their quality. We further obtain ground truth results based on measurements at these testing locations. To evaluate localization, we collect vehicular measure-

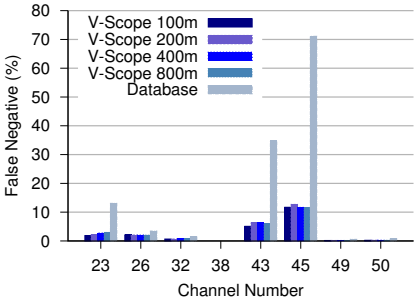

Figure 9: Fraction of locations under-utilized by the database and V-Scope in predicting TV whitespace spectrum.

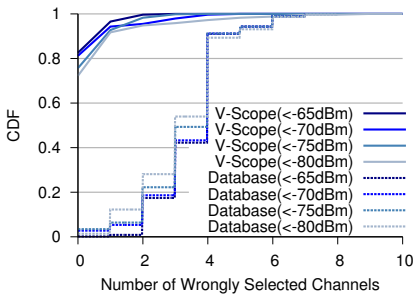

Figure 10: Number of whitespace channels wrongly selected at different locations under various quality constraints.

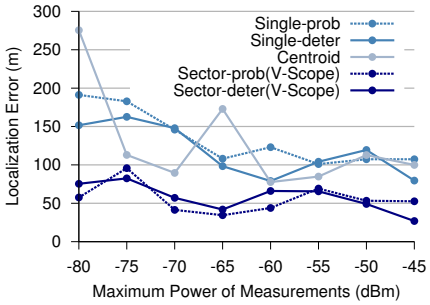

Figure 11: Accuracy in localizing our whitespace transmitter.

ments for a whitespace device mounted atop a building. We calculate the difference between its actual location (determined by GPS) and the estimated location, called localization error.

4.1 Predicting TV whitespace spectrum

We define two types of errors in determining whitespace spectrum, i.e., *false positive* and *false negative*. A false positive is a location where an occupied channel (with power $\geq$ -114dBm) is mis-predicted as whitespaces, whereas the opposite being a false negative. Figure 9 shows the false negative rates of the database and V-Scope models fitted with different region sizes in predicting all the TV broadcast channels. We note 13–71% false negative rates of the database in predicting half of the broadcast channels, indicating they are unnecessarily blocked for whitespace utilization over a large area. We further analyze the spatial distribution of these prediction errors and find that most of them result from over-provisioning the protection contour of TV broadcasts by the database. For V-Scope, we note a 100m region model can reclaim most of the spectrum wastage by the database (up to 59% locations). We also report that both the database and V-Scope achieve very low false positive rates of less than 0.3%, suggesting their efficacy in protecting primary devices.

4.2 Selecting suitable whitespace channels

Using the whitespace channels predicted above, we evaluate the accuracy of V-Scope and the database in selecting appropriate ones under different power constraints. There are about 22 – 30 whitespace channels at each location to be selected from. A whitespace channel is deemed to be suitable if its interference power is below a given power threshold. Such a quality constraint can be easily estimated by network operators based on parameters such as the distance of wireless links, transmission power, and the minimum signal-to-noise ratio (SNR) for decoding a received signal.

Figure 10 shows the CDF of the number of wrongly selected channels at different locations for both approaches under different quality constraints. Note that a 5dB increase in two consecutive power constraints can lead to a 15 – 30Mbps lower data rate for the state-of-the-art 802.11n technology [16]. Without attempting to distinguish channel quality, the database can select all the appropriate channels at less than 2% of the locations, and have 3 – 4 channels wrongly selected for 50% of the locations. In half of the mis-predicted channels, we find the noise power to deviate by at least 12dB, which can cause up to a 240Mbps drop in data rates in those commercial 802.11n radios [16]. In contrast, V-Scope correctly selects all the qualified whitespace channels at 72% – 83% locations, and mis-predicts at most 1 channel for 92% – 97% locations. The significantly higher accuracy suggests that V-Scope can help avoid most of the performance penalty on a whitespace network due to mis-selecting an channel of inappropriate quality.

4.3 Localizing whitespace transmitters

We compare following techniques in localizing our whitespace transmitter. *Single-deter* and *Single-prob* are two state-of-the-art techniques [3, 12] based on a deterministic propagation model and a probability model respectively. These models are constructed using all the measurements indiscriminately. *Sector-deter* and *Sector-prob* are our sector-based versions of these techniques, which are adopted in V-Scope. *Centroid* estimates the device's location using the geometric center of those strong measurements (top 5dB).

Figure 11 shows the error of different algorithms based on measurements below different power thresholds. We observe that *Single-deter* achieves a low error of 27m using all the measurements (below -45dBm). Under different power thresholds, our sector-based algorithms improve *Single-deter* and *Single-prob* by 1.2 – 3$\times$ and 1.5 – 3.5$\times$, because they carefully select measurements from a few sectors and use them to develop a sector specific model. *Sector-deter* also refines *Centroid* by 1.2 – 4.1$\times$. The error increases with weaker measurements for all the techniques because they present an indistinct propagation trend. Nevertheless, we note a more moderate performance degradation in our algorithms that can extract a relatively accurate trend from less noisy measurements.

5. CONCLUSIONS AND FUTURE WORK

We presented a vehicle-based measurement framework for TV whitespaces called V-Scope. Our system leverages spectrum sensors mounted on public vehicles to collect spectrum measurements during the drive. These measurements are aggregated and utilized to refine various propagation models, which can enable spectrum databases to better predict whitespace spectrum, estimate its channel quality, and validate the location of primary and secondary devices. A preliminary version of our system was reported in a position paper [17]. As our future work, we intend to extend our wardriving approach to managing indoor whitespaces by leveraging additional information about building layouts and construction material. We also plan to explore various scalability challenges in deploying and managing such a measurement-enhanced database.

6. REFERENCES

[1] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, and M. Welsh. White space networking with wi-fi like connectivity. In *SIGCOMM*, 2009.

[2] R. Chandra, T. Moscibroda, P. Bahl, R. Murty, G. Nychis, and X. Wang. A campus-wide testbed over the TV white spaces. *MC2R*, July 2011.

[3] K. Chintalapudi, A. Padmanabha Iyer, and V. N. Padmanabhan. Indoor localization without the pain. In *MobiCom*, 2010.

[4] A. T. S. Committee. Atsc digital television standard - part 2: Rf transmission system characteristics, December 2011.

[5] FCC. Unlicensed operation in the tv broadcast bands, second memorandum opinion and order, 2010.

[6] H. Kim and K. G. Shin. In-band spectrum sensing in cognitive radio networks: energy detection or feature detection? In *MobiCom*, 2008.

[7] R. Murty, R. Chandra, T. Moscibroda, and P. V. Bahl. Senseless: A database-driven white spaces network. *IEEE Transactions on Mobile Computing*, 2012.

[8] S. Narlanka, R. Chandra, P. Bahl, and J. Ferrell. A hardware platform for utilizing tv bands with a wi-fi radio. In *LANMAN*, 2007.

[9] New York Times. Carriers warn of crisis in mobile spectrum, 2012.

[10] C. Phillips, D. Sicker, and D. Grunwald. Bounding the error of path loss models. In *DySPAN*, 2011.

[11] C. Phillips, M. Ton, D. Sicker, and D. Grunwald. Practical radio environment mapping with geostatistics. In *DySpan*, 2012.

[12] S. Rayanchu, A. Patro, and S. Banerjee. Catching whales and minnows using wifinet: Deconstructing non-wifi interference using wifi hardware. In *NSDI*, 2012.

[13] S. Sen, T. Zhang, M. Buddhikot, S. Banerjee, D. Samardzija, and S. Walker. A dual technology femto cell architecture for robust communication using whitespaces. In *DySpan*, 2012.

[14] Spectrum Bridge Inc. Spectrum bridge database. http://www.spectrumbridge.com/Home.aspx.

[15] Spectrum Bridge Inc. TV whitespaces success stories. http://www.spectrumbridge.com.

[16] Ubiquiti Inc. Datasheet of ubiquiti sr71-a mimo cards. http://dl.ubnt.com/sr71a_datasheet.pdf.

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

[18] T. Zhang, S. Sen, and S. Banerjee. Scout: An asymmetric vehicular network design over TV whitespaces. In *ACM HotMobile*, 2013.