Exploring Indoor White Spaces in Metropolises

Xuhang Ying
Dept. of Information Engineering
The Chinese University of Hong Kong

Jincheng Zhang
Dept. of Information Engineering
The Chinese University of Hong Kong

Lichao Yan
Dept. of Information Engineering
The Chinese University of Hong Kong

Guanglin Zhang
Institute of Network Coding
The Chinese University of Hong Kong

Minghua Chen
Dept. of Information Engineering
The Chinese University of Hong Kong

Ranveer Chandra
Microsoft Research
One Microsoft Way
Redmond, WA 98052

ABSTRACT

It is a promising vision to utilize white spaces, *i.e.*, vacant VHF and UHF TV channels, to satisfy skyrocketing wireless data demand in both outdoor and indoor scenarios. While most prior works have focused on exploring outdoor white spaces, the indoor story is largely open for investigation. Motivated by this observation and that 70% of the spectrum demand comes from indoor environments, we carry out a comprehensive study of exploring *indoor* white spaces. We first present a large-scale measurement of outdoor and indoor TV spectrum occupancy in 30+ diverse locations in a typical metropolis Hong Kong. Our measurement results confirm abundant white spaces available for exploration in a wide range of areas in metropolises. In particular, more than 50% and 70% of the TV spectrum are white spaces in outdoor and indoor scenarios, respectively. While there are substantially more white spaces in indoor scenarios than in outdoor scenarios, there is no effective solution for identifying indoor white spaces. To fill in this gap, we propose the first system WISER (for White-space Indoor Spectrum EnhanceR), to identify and track indoor white spaces in a building, without requiring user devices to sense the spectrum. We discuss the design space of such system and justify our design choices using intensive real-world measurements. We design the architecture and algorithms to address the inherent challenges. We build a WISER prototype and carry out real-world experiments to evaluate its performance. Our results show that WISER can identify 30%-50% more indoor white spaces with negligible false alarms, as compared to alternative baseline approaches.

Categories and Subject Descriptors

C.2.1 [**Computer-communication networks**]: Network Architecture and Design-Network topology; C.4 [**Performance of systems**]: Design studies, Measurement techniques

MobiCom'13, September 30-October 4, Miami, FL, USA.

Keywords

TV white spaces, clustering algorithms, sensor placement

1. INTRODUCTION

The skyrocketing growth of mobile devices and applications has triggered a need for additional radio frequency (RF) spectrum to satisfy this demand. Since most of the frequency spectrum has been licensed for different purposes (satellite, TV, radio, RADARs, cellular, *etc.*), a recent concept of dynamic spectrum access (DSA) is being explored to provide additional spectrum with little disruption to existing licensed users and their devices.

A recent manifestation of DSA is in the TV spectrum. In 2008, the FCC passed a historic ruling that allowed unlicensed devices (similar to Wi-Fi) to operate in the locally unoccupied TV spectrum (also called the TV white spaces or simply white spaces). Devices, similar to the Wi-Fi devices of today, are required to detect the available spectrum before using it for communication. Per a 2010 FCC Second Order and Report [6], the white space devices can detect available spectrum using either spectrum sensing or by querying a geo-location web service over the Internet.

Most white space devices and standards are being designed around querying the web service for spectrum availability. This is primarily because spectrum sensing is expensive – in cost, energy consumption and complexity of the circuitry. Furthermore, it is more difficult to accurately detect TV signals using spectrum sensing at low thresholds with commercial, off-the-shelf hardware. In contrast, the geo-location approach does not require any hardware and is easier to implement. Devices report their locations to a web service, which in turn returns a list of TV channels that can be used at that location[1]. See [8] for a survey on this line of approach. However, this approach suffers from inherent inefficiency. The geo-location service uses propagation modeling to determine the available spectrum, and hence, is very conservative in the channels it returns for a given location. For example, the propagation models used by the FCC do not account for buildings, and man-made obstructions that exist in urban areas. In a measurement study in [36],

[1]To avoid interference to white space users from wireless microphone usage, wireless microphone users are suggested to register the location and frequency usage to the geo-location database. In this way, the database can exclude the frequency occupied by wireless microphone at the location and within its estimated interfering neighbors from the returned list.

an in-operation geo-location database service reports only half of vacant channels across a major city.

In this paper we carry out measurement-driven analysis and design to white space networking in indoor environments. We first carry out a large-scale measurement across 30+ diverse locations in a typical metropolis, and reveal that more than 50% and 70% of the TV spectrum are white spaces in outdoor and indoor scenarios, respectively. While there are significantly more white spaces in indoor scenarios than in outdoor scenarios, there is no effective solution for identifying indoor white spaces. Given that most people are indoors 80% of the time [24] and 70% of spectrum demand comes from indoor environments [12], it will be extremely useful if it were possible to use the vacant TV channels in indoor scenarios.

We therefore take the next step, and propose a system, called WISER (for White-space Indoor Spectrum EnhanceR), that increases the number of TV channels available for indoor white space networking, while (i) not requiring client devices to sense the spectrum, (ii) building more accurate white space database locally by integrating outdoor and indoor sensing, and (iii) not interfering with TV transmissions.

WISER enables buildings to become “white space enabled” using an innovative approach of profiling, sensor placement, and integration with geo-location databases. It provides a technique for building owners or managers to get their building profiled for white space use, based on which WISER strategically determines a few locations to place RF sensors. These sensors capture the additionally available white space spectrum in the building. Ideally we would require a very dense placement of RF sensors to obtain every additionally available channel. However, such approach can be very expensive due to the large number of RF sensors[2]. Hence, we propose an innovative clustering scheme to reduce the number of sensors needed to (i) capture the white space variation in indoor environments, yet (ii) provide most of the additional channels for indoor use.

Throughout this paper we make the following contributions:

- In Section 2, using comprehensive measurements in Hong Kong, we show that more than 50% and 70% of the TV spectrum are white spaces in outdoor and indoor scenarios, respectively. Further, in Section 3, we reveal that indoor white spaces expresses interesting location correlation and channel correlation. These characteristics provide insights in identifying and utilizing indoor white spaces.

- In Section 4, we propose an indoor white space identification system, called WISER, that allows additional white space spectrum to be used indoors. To the best of our knowledge, this is the first system that can use the additional white spaces without requiring client devices to sense the spectrum. Our approach is rather general, and can be additionally used in other spectrum monitoring applications. In Section 5, we present channel and location clustering algorithms to reduce the number of sensors used by WISER to monitor white space spectrum.

- In Section 6, we build a proof-of-concept WISER prototype, deploy it in a building floor, and evaluate its performance by real-world experiments across a four-month duration. In particular, we demonstrate that WISER can identify 30%-50% more indoor white spaces with negligible false alarms, as compared to alternative baseline approaches.

[2]For the purpose of identifying white spaces, an RF sensor consisting of USRP1, a TVRX receive-only daughter board and an antenna still costs around 1K US dollars. Hence, reducing the total sensor cost is an important system design consideration.

Given that indoor environments are crowded, and have the most spectrum demand, our system provides a new principled approach to make additional spectrum available in these environments. Last but not the least, we remark that WISER is not limited to the TV white space spectrum. It can be similarly applied to any other portion of the spectrum where dynamic spectrum access techniques will be adopted.

2. INDOOR/OUTDOOR WHITE SPACE AVAILABILITY MEASUREMENT

2.1 Objective

We carry out a large-scale indoor/outdoor white space measurement in a typical metropolitan city Hong Kong. The purpose of our measurement is two-fold. First, we aim to understand the difference between the indoor and outdoor white space availability patterns. Such understanding motivates our investigation on exploring indoor white spaces. Second, while various measurement studies have been reported in the literature, we find a large-scale indoor/outdoor measurement is missing. Such large-scale measurement is critical for properly evaluating the potential of white space networking in metropolises, where different districts observe diverse spectrum occupancy patterns.

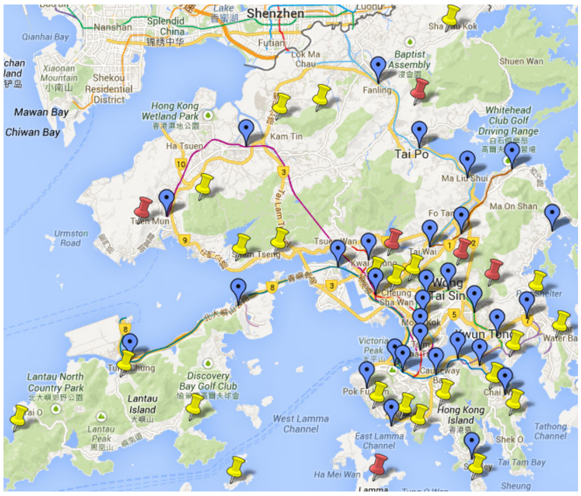


Figure 1: A map showing the six principal TV transmitting stations, 23 fill-in TV transmitting stations [5], and 31 diverse measurement locations that cover all 18 districts in Hong Kong. Principle and fill-in TV stations are labeled using red and yellow pin-shape markers, respectively. Measurement locations correspond to blue droplet-shape markers. Each measurement location is covered by 1-5 TV stations (2.2 on average) [4], and the distance to its nearest TV station ranges from 0.97 km to 9.35 km (3.25 km on average).

2.2 Methodology

2.2.1 *Equipment and Setup*

The measurement equipment consists of a USRP [9], a Log Periodic PCB Antenna, a laptop computer, and a battery bank. We use the USRP board coupled with a TVRX receiver-only daughter board and a GNU Radio platform [3] to construct a spectrum analyzer, for measuring and detecting TV signals in the 470-806 MHz TV spectrum band. We calibrate the measurement device using a RF signal generator to obtain accurate power reading in dBm. We

	Outdoor				Indoor				Indoor Bonus
	Urban	Sub-urban	Rural	Overall	Urban	Sub-urban	Rural	Overall	Overall
White Space Ratio (%)	44.1	55.9	60.9	53.6	67.9	74.7	73.3	72	18.4
Number of Vacant Channels	18.5	23.5	25.6	22.5	28.5	31.4	30.8	30.2	7.7
Total Vacant Spectrum (MHz)	148	188	204	180	228	251	246	242	62

Table 1: Summary of the indoor and outdoor white space measurement results.

use energy detection to detect analog signals by comparing the receiving power spanning 100 KHz centered at their visual carriers against -104.2 dBm. For digital TV signals, we use a feature-based detection scheme similar to that in [22] for accurate detection; the scheme is able to detect digital TV signal strength as low as 96 dBm/8 MHz. More details can be found in our measurement report [36].

2.2.2 Measurement Locations and Design

We measure the indoor/outdoor white space availability at 31 diverse locations in Hong Kong, including Hong Kong Island, Kowloon Peninsula, and New Territories. These locations cover all the 18 districts of Hong Kong, which have very different terrain and population characteristics [1]. For example, the urban area such as MongKok has the world-highest population density and the skyscrapers in the area have deep influence on the signal propagation. While in remote areas such as Yuen Lang, the population and tall-building densities are much lower. Fig. 1 shows the six principal TV stations, 23 fill-in TV stations[3] and 31 measurement locations. More details (*e.g.,* effective radiated power, polarization, *etc.*) about the TV broadcasting network of Hong Kong are available in [7]. All measurement locations are in well populated commercial/residential areas; thus our measurement results capture the representative spectrum occupancy patterns in Hong Kong.

At each location we measure the indoor/outdoor spectrum occupancy pattern of the TV bands at three time instants, one in the morning, one at noon, and one in the evening. The indoor measurements are taken inside various commercial buildings at the selected locations. The time interval between two adjacent measurements is 4 hours. In each measurement, we scan all 42 analog and digital TV channels multiple times. These channels are in the frequency range of 470-806 MHz with 8 MHz channel spacing [32]. For ease of discussions, they are labeled as CH 1 to CH 42. Using feature-based TV signal detection schemes, we analyze the measurement results and identify locally unoccupied channels. These unoccupied channels are labeled as white spaces available at the corresponding locations.

2.3 Observations

We summarize the indoor and outdoor white space measurement results in Table 1. In particular, we first group the 31 measurement locations into three areas, namely urban, sub-urban, and rural, according to the population density [1]. Then we summarize the indoor and outdoor white space results for each area.

Several observations can be made from the results. First, similar to the US and Europe, there are a large number of vacant TV channels in Hong Kong – more than 50% of the TV channels are white spaces. Together with existing white space measurement studies [13, 20, 26], our measurements confirm abundant white-space networking potential in metropolises, different areas in Hong Kong observe very different white space availability patterns. For example, in outdoor scenarios, there are on average 5 more vacant channels in the sub-urban area than in the urban area, corresponding to 54 MHz additional spectrum. Third, there are more indoor white spaces than the outdoor white space, which is mainly due to the signal attenuation because of the blocking effects of walls. In particular, in indoor scenarios, there are on average 72% of the 42 TV channels are white spaces, 18.4% higher than that in outdoors. This corresponds to 7.7 additional vacant channels and a total amount of 62 MHz spectrum. The amount of overall indoor white spaces is 242 MHz, enough to support one additional Wi-Fi service in operation.

Furthermore, the indoor white spaces are less fragmented than the outdoor ones. The average length of contiguous vacant channels is 3.84-channel in indoor scenarios as compared to 1.87-channel in outdoor scenarios. This indicates the indoor white spaces is of better "quality" than the outdoor ones, since it is easier to design wireless devices to communicate over contiguous channels than over fragmented ones.

[3]There are several tens of additional fill-in TV stations, but their precise location information is not available online, and we are unable to estimate their broadcasting coverage accordingly.

3. INDOOR WHITE SPACE CHARACTERISTICS MEASUREMENT

Complementary to our large-scale indoor/outdoor availability measurements, we also conduct intensive indoor measurements to gain necessary understanding of indoor white space characteristics.

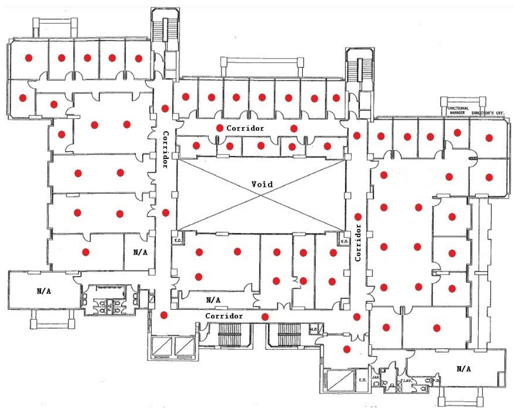


Figure 2: A map showing the 65 measurement locations on the 7th floor of a building. Each measurement location corresponds to a dot in red color on the map.

3.1 Methodology

3.1.1 Equipment and Setup

The same equipment described in Section 2.2.1 is used for indoor white space characteristics measurement, except an omnidirectional antenna with a gain of 0 dBi.

3.1.2 White Space Threshold

In our measurements, we label a TV channel as locally unoccupied if the corresponding channel receiving power is less than

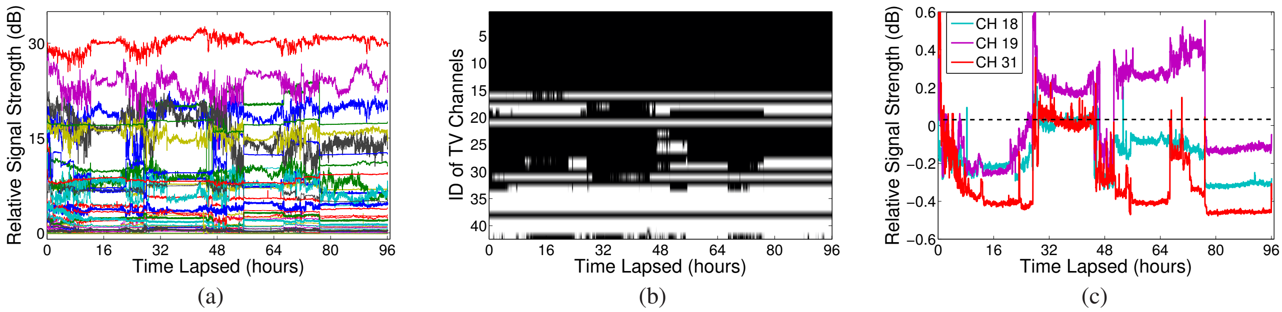


Figure 3: (a) Relative signal strength (in dB) for all 42 TV channels in a 96-hour window. (b) Channel occupancy statuses for all 42 channels in the same 96-hour window. White spaces are denoted in white, and occupied channels in black. (c) Three TV channels (CH 18, 19 and 30) that move up and down around the white space thresholds (*e.g.,* the dotted line). They become white spaces from time to time.

or equal to a pre-set white space threshold, and as occupied otherwise. This energy-detection based identification method is faster but less sensitive as compared to those used in our indoor/outdoor large-scale measurements. We adopt the method to do fast profiling of the white space availability at a large number of dense locations in an operational building. We set the white space thresholds to be -84.5 dBm/8 MHz for the digital TV signals and -104.2 dBm/100 KHz for analog TV signals (centered at their visual carriers). These values are the maximum readings out of detected power spanning the corresponding bands in all empty channels based on long-time sensing results. While these white space thresholds used for determining channel vacancy may seem aggressive due to hardware limitations, we believe that the observations drawn from the measurement results are general and our system is not tied with these thresholds. We note that if improved hardware is in use to allow a highly-conservative sensing threshold of -114 dBm as suggested by FCC, then the identified unoccupied channels are safe to use, in the sense that secondary users using these channels will not cause interfere with primary TV users within a considerably large neighborhood. See [6] for detailed discussions on setting interference-safe sensing threshold for TV white space networking.

3.1.3 Measurement Locations and Design

As shown in Fig. 2, a total of 65 indoor locations on one floor of a building are selected for measurement. At each indoor location, every channel is scanned five times. We observe <1% difference in recorded signal strengths beyond five-time measurements. During each scanning, each channel is measured for 0.2 seconds. The recorded signal is then processed by a GNURadio FFT program with a bin size of 2048 and a resolution of 3.9 KHz. We then record the maximum values observed at each bin for the channel during each scanning and compute the average value for that bin. For a digital channel, we compare the total channel power over 8 MHz centered at the middle of a TV channel against the white space threshold of -84.5 dBm. For a analog channel, we compare the power spanning 100 KHz centered at the frequency of its visual carrier against -104.2 dBm.

We conduct two types of indoor measurements, namely *indoor long-time sensing* and *indoor one-time profiling*. Indoor long-time sensing is for understanding temporal features of indoor white spaces. We measure the received signal strengths for all 42 TV channels at a typical indoor location consecutively for 96 hours. Indoor one-time profiling is to probe spatial features of indoor white spaces for a typical indoor environment. We mount the measurement equipment onto a movable cart and run a Python script to scan all TV channels automatically. We then profile the 65 indoor locations one by one. The duration for profiling one location is about 1.5 minutes, and the whole process lasts roughly three hours. We can obtain a 65x42 matrix containing absolute signal strengths (in dBm) for 42 channels at the 65 indoor locations. To facilitate comparison, we convert the absolute signal strengths to the relative signal strengths by comparing them against the corresponding white space thresholds. We collect a total of 13 one-time profiling data in two periods in four months.

3.2 Long-Time Indoor Sensing Results

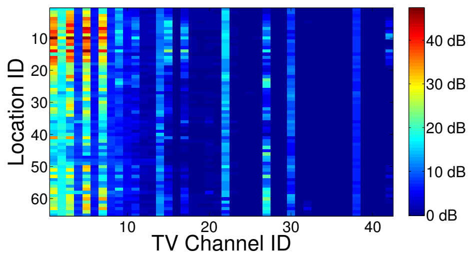
There are several observations from the long-time indoor sensing. First, from Fig. 3a, there are *strong* channels whose relative signal strength exceed a large threshold most of the time. For instance, we observe 14.29% channels with >10 dB relative signal strength 95% of the time across the 96-hour interval, which means they are *long-term occupied* at the typical indoor location. Second, for *weak-to-normal* channels, their signal strengths express a *short-term stable yet long-term unstable* pattern. Third, we observe intermittent white spaces availability from Fig. 3b. This results from the temporal signal strength variation in weak-to-normal channels. These observations suggest that to extract the maximum indoor white space potential, it suffices to identify strong channels via long-time sensing and then focus resources to track the slow-varying white space availability of weak-to-normal channels.

3.3 One-Time Spectrum Profiling Results

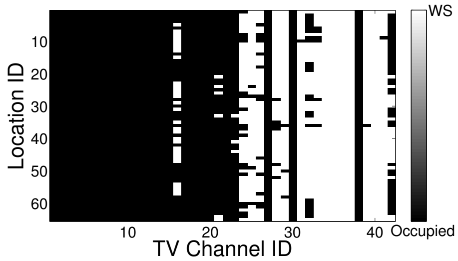
We have the following observations from one-time spectrum profiling results. First, as shown in Fig. 4, a channel may express *spatial variation* in its relative signal strength and white space availability across different indoor locations. Such *spatial variation* is caused by complex outdoor and indoor signal propagation and attenuation patterns due to various factors, for example the blocking effects of indoor walls. This implies potential white space loss if indoor white space availability is directly inferred from outdoor availability results.

Second, seen from Fig. 5a, for a given channel, there is also strong correlation in signal strengths and white space availability across different locations. This suggests that we can infer the channel vacancies of multiple correlated locations from those of one or a few representative locations. We refer this observation as indoor white space *location correlation*.

In addition, seen from Fig. 5b, we observe that for multiple channels, there are strong correlation in their signal strengths and white space availability patterns across all locations. The similarities among these channels can be due to that they share the same trans-



(a)



(b)

Figure 4: (a) Typical spatial map of channel relative signal strengths extracted from one-time spectrum profiling data. We observe signal attenuation patterns across different locations. (b) Typical spatial map of indoor white spaces (denoted in white, occupied channels in black) extracted from one-time spectrum profiling data. We observe different channel vacancy patterns across different locations.

mission tower, similar loss and attenuation patterns in their propagation. This suggests that we can group these "similar" channels together into a group, and infer the vacancies of this group of channels from those of a representative channel in this group. We refer this observation as indoor white space *channel correlation*.

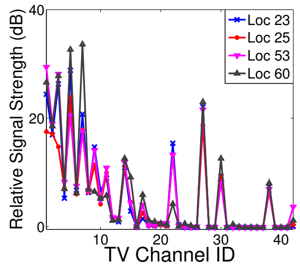
The above observations on location correlation and channel correlation are drawn based on one-time profiling results. They have the potential to form useful guidelines in designing indoor white space identification systems. One important question stands on the way, however, is *whether the correlation among channels and locations are stable across different time epochs*.

3.4 Stability of Indoor Channel Correlation and Location Correlation

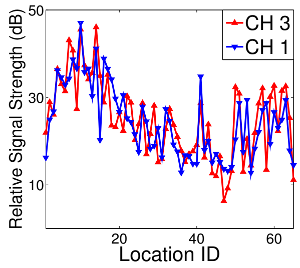
To study the stability of the channel correlation and location correlation, we cluster channels and locations according to their similarity using one-time profiling data collected across seven days in the first two-week period. We then compare one of the clustering results (*e.g.,* the first day) with other clustering results. If the channel correlation and location correlation are stable over time, then we should see consistent clustering results in these seven days. We should otherwise observe discrepancy in clustering results.

To filter out the noise in profiling data while maintain necessary information about signal strengths, we first quantize the relative signal strengths as described in Section 4.5.3. We then apply the clustering algorithms described in Section 5, and compute the Rand Index [33] of the clustering results to exam their consistency. The Rand Index is commonly accepted as an objective criterion for comparing two clustering results. A Rand Index of 1 indicates two identical clustering results, and 0 indicates total disagreement.

To facilitate the illustration for Fig. 6, we define *a channel group* as a group of channels that exhibit similar signal strengths across



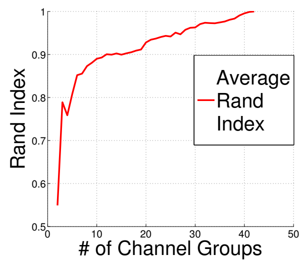
(a)



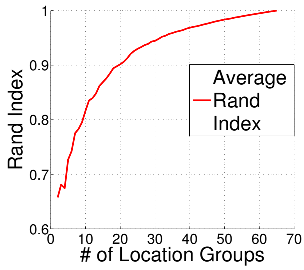
(b)

Figure 5: (a) Location 23, 25, 53 and 60 receive close signal strengths from many TV channels. Each line represents the relative signal strengths received at an indoor location for 42 channels. We observe strong location correlation. (b) CH 1 and 3 exhibit very close signal strengths at different indoor locations. Each line represents the relative signal strengths for a channel at 65 indoor locations. We observe strong channel correlation.

different indoor locations, and *a location group* as a group of locations that receive similar signal strengths for a particular set of channels. As shown in Fig. 6a, when we have three or more channel groups, the average Rand Index reaches 0.79 and keeps rising. It means that we would obtain very similar channel clustering results, and the channel correlation is very consistent across different days. Fig. 6b shows that the average Rand Index has exceeded 0.80 when there are 10 or more location groups for a group of channels (*i.e.,* CH 1, 3, 5 and 7), which indicates that location correlation also exhibits consistency in time. These observations imply that *the correlations among channels and locations are very likely to be stable in time*, which is critical in gauging our designs of the indoor white space identification system, to be discussed in the next section.



(a)



(b)

Figure 6: (a) Channel correlation is consistent across different days. The input is a matrix of quantized relative signal strengths for 42 channels at 65 locations. We would obtain highly similar channel clustering results with more then three channel groups. (b) Location correlation consistency across different days. The input is a matrix of quantized relative signal strengths at 65 locations for 4 channels (*i.e.,* Ch 1, 3, 5 and 7). We would obtain highly similar location clustering results with more than five location clusters. In both figures, clustering result of data in the first day serves as a main reference. Both curves start with two groups. The Rand Index is 1 when there is only one group.

3.5 Summary

In summary, our results from long-term sensing, one-time profiling, and correlation stability verification reveal the following important observations:

- Strong channels are long-term stable; weak-to-normal channels are short-term stable but long-term unstable.
- *Location correlation*: given a channel, there are strong correlations in its signal strength and white space availability across multiple indoor locations; these correlations are stable in time.
- *Channel correlation*: there are strong correlations in multiple channels in their signal strengths and white space availability patterns across all locations; these correlations are stable in time.

The first observation suggests that to extract the maximum indoor white space potential, it suffices to identify strong channels via long-time spectrum sensing and then track the slow-varying white space availability of weak-to-normal channels. The rest two observations suggest that we can focus on monitoring representative channels at representative indoor locations, and inferring the availability of other channels at other locations by exploiting the location and channel correlations. These observations will be utilized in designing indoor white space identification systems in the next section.

4. WISER - WHITE SPACE INDOOR SPECTRUM ENHANCER

In this section, we first explore the design space of an indoor white space identification system. We then present WISER's architecture, which consists of real-time sensing module, white space database, and indoor positioning module.

4.1 Design Objective and Design Space

A well-performed indoor white space identification system needs to (i) minimize false alarms (*safety*), and (ii) identify as many white spaces correctly as possible (*efficiency*). In addition, the sensor cost is also a major consideration, as RF sensors can be expensive. We define the following three metrics to evaluate a system:

- *False Alarm Rate (FA Rate):* the ratio between the number of channels that a system mis-identifies as vacant and the total number of vacant channels that the system identifies. A system with lower *FA Rate* is safer.
- *White Space Loss Rate (WS Loss Rate):* the ratio between the number of channels that a system mis-identifies as occupied and the total number of actually-vacant channels. A system with lower *WS Loss Rate* is more efficient.
- *Sensor Cost:* the total cost of all RF sensors in use.

Generally, there are several approaches to identify indoor white spaces. The *outdoor-sensing-only (OS-Only)* approach performs spectrum sensing locally on the rooftop using one outdoor sensor, and only reports channels that are available outdoors for indoor use. Despite of the low sensor cost, this approach is too conservative to be efficient, as it fails to take into account of the significant attenuation on TV signals due to the blocking effects of indoor obstacles (*e.g.*, walls).

Under the second approach called *one-time-profiling-only (OTP-Only)*, we only profile indoor locations once, then label every channel at each location as either 1 (vacant) or 0 (occupied), and store such information in the white space database for later retrieval. However, this approach assumes the indoor white space availability to be time-invariant, which does not hold according to our indoor white space measurements in Section 3. As a result, this approach fails in both safety and efficiency; in particular, violating safety would cause severe harm to primary users.

The third approach is *sensor-all-over-the-place*. This approach deploys dense indoor sensors to monitor indoor white spaces. Although this scheme can achieve the ideal safety and efficiency, it comes with a forbidden sensor cost.

In summary, none of the above approaches could achieve a satisfactory performance at low cost. To tackle this problem, we propose an indoor white space identification system called WISER with excellent safety and efficiency at lower sensor cost as compared to the above baseline solutions.

4.2 WISER Overview

As shown in Fig. 7, WISER consists of three components, namely real-time sensing module, white space database, and indoor positioning module. WISER takes users' locations as the inputs and outputs the indoor white space availability at the given locations.

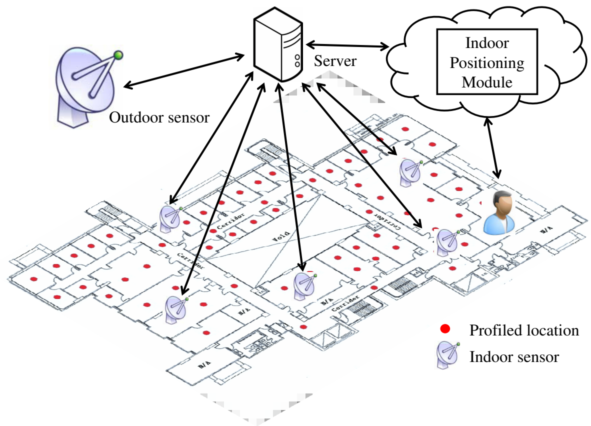


Figure 7: Architecture schema of WISER.

4.3 Indoor Positioning Module

In WISER, a user determines its indoor location by using the indoor positioning module. Indoor positioning has been widely explored in recent years. Many wireless technologies can be integrated into indoor positioning, including IR, ultra-sound, RFID, WLAN, and Bluetooth. See [17] for a recent survey. Since indoor positioning is developing into a mature technology, WISER simply uses one existing system as the indoor positioning module. It is conceivable that the indoor positioning accuracy will affect WISER's performance. We will study the relationship between indoor positioning accuracy and the performance of WISER in Section 6.

4.4 White Space Database

The white space database receives real-time channel signal strength reported from the sensors in the real-time sensing module, as well as the corresponding up-to-date indoor white space availability. After obtaining its indoor position, user queries the white space database to retrieve a list of vacant channels for communication with a separate white space networking infrastructure. To handle dynamics of wireless microphones, wireless microphone users are suggested to register the location and frequency usage to the white space database, so that WISER can exclude the frequency occupied by wireless microphone at the location and within its estimated interfering neighbors from the returned list.

4.5 Real-time Sensing Module

Given an indoor environment, the real-time sensing module performs real-time outdoor and indoor spectrum sensing, and reports the results to the white space database. To realize this functionality, we first conduct one-time spectrum profiling at a sufficient number of indoor locations. To exploit the observation in long-time measurement (Section 3.2), we then group the channels into strong and weak-to-normal channels. We also identify some permanently available channels based on the long-time outdoor sensing, and focus on the remaining weak-to-normal channels. Then, we perform channel-location clustering to leverage channel and location correlations (Section 3). Last, we place indoor sensors based on the results of channel-location clustering to optimize the use of a limited number of sensors and then send real-time data to server periodically.

4.5.1 One-Time Spectrum Profiling

One-time spectrum profiling aims to learn the indoor white space characteristics, including channel and location correlations. These locations should be as dense as possible. For example, in our study, 65 measurement locations are profiled including almost every room and corridor on an indoor floor, as shown in Section 3.4. For a non-profiled location, its indoor white space availability is assumed to be the same with the nearest profiled one. The profiling data will serve as the main reference to determine indoor sensor locations. Ideally, we should profile all measurement locations simultaneously. However, since we observed that indoor white spaces tend to be stable with moderate variation in the short term (*e.g.,* hours), we base WISER deployment on the asynchronous one-time profiling within three hours.

4.5.2 Channel Grouping

We first group TV channels into two basic classes: strong channels and weak-to-normal channels (including permanently vacant channels) based on one-time and long-time measurements. Weak-to-normal channels have a higher chance to be white spaces due to signal variation and attenuation.

We motivate the simple grouping scheme with the following reasons. First of all, such classification is general. In fact, for a particular area, TV services are broadcast via one or more major TV towers in several channels. These channels usually carry strong signals and are distinguishable from others. In addition, due to stable TV broadcasting arrangement in the long run (*e.g.,* years), the obtained grouping result will also be stable accordingly unless dramatic changes occur (*e.g.,* shutdown of a TV tower). Moreover, it is efficient to group channels into such two basic classes. Because by exploiting "always on" nature of some TV channels, indoor sensors can focus on other time-varying channels rather than these stably strong ones.

4.5.3 Channel-Location Clustering

The channel-location clustering aims to obtain *channel-location (CL) groups*, each of which contains locations with similar signal strengths for channels with similar propagation patterns. In our current solution, we obtain CL groups by first clustering the channels according to their received signal strengths across the locations. Then, for each channel cluster, we group the locations according to the received signal strength distribution among those channels. More details will be described in Section 5. Nonetheless, conducting joint channel-location clustering is an alternative method, and is left for future work.

To reduce the undesirable impact of noise, we introduce a *uniform quantizer* for received signal strengths before conducting channel-location clustering, which is defined as below:

$$S_q = \begin{cases} -1, & S_r \leq 0 \\ \left\lfloor \frac{S_r}{Q} \right\rfloor, & S_r > 0 \end{cases}$$

where S_r is the received relative signal strength compared to the pre-set white space thresholds, Q is the quantization step (*e.g.,* 5 dB), and S_q is the quantized relative signal strength.

4.5.4 Indoor Sensor Placement

Given a set of CL groups, we aim to deploy indoor sensors to fulfill (i) *coverage* requirement: every CL group is covered by at least one sensor, and (ii) *performance* requirement: safety is guaranteed and efficiency is maximized, given a number of indoor sensors. The brute-force way is to enumerate all possible combinations to find the best one. Instead of adopting this inefficient approach, we propose a greedy algorithm that could achieve local optimum with guaranteed system performance, as described in Section 5.

To avoid interference with primary users, indoor white space availability shared within each CL group is derived from real-time sensing results in a conservative way. First of all, if a particular CL group is covered by multiple indoor sensors, indoor white spaces are determined by comparing the maximum received signal strength for a channel with the white space thresholds as follows:

$$WS(C_x) = \begin{cases} 1, & S_{max}(C_x, L_{sensors}) < WSThr_x \\ 0, & otherwise \end{cases}$$

where $WS(C_x)$ means the white space availability for channel x at all locations in the CL group, $S_{max}(C_x, L_{sensors})$ means the maximum received signal strength among multiple sensor locations for channel x in the CL group and $WSThr_x$ is the white space threshold for channel x.

Second, if a CL group is covered by only one indoor sensor, a *protection range* (in dBm) needs to be added to minimize false alarm rate as follows:

$$WS(C_x) = \begin{cases} 1, & S(C_x, L_{sensors}) < WSThr^*_x \\ 0, & otherwise \end{cases}$$

where $WSThr^*_x = WSThr_x - PR_x$, and the PR_x is the protection range for channel x for this CL group. To exploit the statistics of received channel signal strength deviation in a CL group in training data sets, we obtain PR_x by multiplying a constant (3 by default) with the standard deviation of all received signal strengths in channel x at locations in the CL group. This constant measures how conservative we are and can be used to control the FA rate.

5. ALGORITHM

In this section, we first discuss the intuition of algorithms we use in deploying real-time sensing module. We then present the channel-location clustering and indoor sensor placement algorithms.

5.1 Algorithm Intuition

As revealed in Section 3.4, there exists consistent correlation among channel signal strength at different locations. An intuitive way to extract such correlation is clustering. Given one-time profiling data, we treat signal strengths across different indoor locations for a channel as a feature vector, and cluster similar channels together to obtain channel clusters. Then for each channel group, we treat signal strengths in different channels of the channel group at a location as a feature vector, and perform location clustering accordingly. In this way, we obtain CL groups (recall that CL group stands for channel-location group). Within each CL group, the channels share similar signal strength at the locations; hence, intuitively we can use one sensor per CL group to monitor signal strengths of the channels (in the group) at the locations (in the group). Following the above intuition, we design Algorithm 1 and Algorithm 2. As the next step of the above intuition, we need to deploy one sensor (or more) per CL group (*i.e.,* coverage requirement), and enable WISER to correctly identify as many indoor white spaces as possible (*i.e.,* performance requirement). These two requirements are very important in designing our indoor sensor placement algorithm (Algorithm 3) and its ranking mechanism (Algorithm 4). More details will be discussed later in this section.

5.2 Proposed Algorithm

The proposed algorithm takes quantized relative signal strength and the number of indoor sensors as inputs, and outputs a list of sensor locations. In the case of multiple feasible sensor placement options, the one with least FA rate (with first priority) and least WS Loss rate is chosen, namely "least-false-alarm-first" criterion.

5.2.1 Channel-Location Clustering

As stated in Section 4.5.3, the channel-location clustering aims to cluster locations that receive similar channel signal strength for certain channels into the same group. We do so by first grouping the channels according to their receiving signal strengths at all indoor locations, and then within each channel group clustering the locations according to their receiving signal strengths of this group of channels. In our algorithm designs, we adopt Ward's minimum variance method [23] as the linkage criterion, and Euclidean distance as the similarity metric owing to their widespread popularity.

Algorithm 1 Channel Clustering

1: **Input:** S: the $M \times N$ training data set
{M is the total number of locations, N is total number of channels to be clustered}
k: the number of channel clusters
2: **Output:** k channel clusters with sizes of $M \times n_1$, $M \times n_2$, ..., $M \times n_k$, where $\sum_{i=1}^{k} n_i = N$
3: **if** $k == N$ **then**
4:     **return** N channel clusters with a size of $M \times 1$
5: **end if**
6: let each channel object ($M \times 1$) be a cluster
7: compute the proximity matrix
8: **repeat**
9:     merge two "closest" clusters based on the linkage criterion
10:     update the proximity matrix
11: **until** k clusters remain
12: **return** k channel clusters with sizes of $M \times n_1$, $M \times n_2$, ..., $M \times n_k$, where $\sum_{i=1}^{k} n_i = N$

As mentioned in Section 5.1, we need deploy a given number of sensors for CL groups. One straightforward way to determine the number of CL groups is to let it be equal to the number of sensors. With a larger number of CL groups, we may not guarantee the coverage requirement; with fewer CL groups, it tends to be a waste of sensor resources, which is undesirable. Provided a number of CL groups, the next question is how to decide the desirable number of channel clusters, and also the desirable number of location clusters for each channel group. There are various ways to probe the true number of clusters, including a model-based approach [16]. In this study, we are more interested in the relationship between the number of indoor sensors and system performance instead of seeking for the true number of channel or location clusters. In our current solution, we permutate all possible numbers of channel clusters, and assign numbers of location clusters to channel clusters that are proportional to sizes of channel clusters. For instance, given 20 CL groups, we may try out 3 (or other possible numbers from 1 to 20) channel clusters, and obtain clusters with sizes of 8, 16 and 16 via channel clustering. Then, we will assign 4, 8 and 8 location clusters to 3 channel clusters accordingly before location clustering. Note that when multiple combinations are possible, WISER will choose the one according to the "least-false-alarm-first" criterion.

Algorithm 2 Location Clustering

1: **Input:** S_i: the $M \times N_i$ channel cluster
{M is the total number of locations to be clustered, N_i is total number of channels}
k_i: the number of location clusters
2: **Output:** k_i location clusters with sizes of $m_1 \times N_i$, $m_2 \times N_i$, ..., $m_{k_i} \times N_i$, where $\sum_{j=1}^{k_i} m_j = M$
3: **if** $k_i == M$ **then**
4:     **return** M location clusters with the size of $1 \times N_i$
5: **end if**
6: let each location object ($1 \times N_i$) be a cluster
7: compute the proximity matrix
8: **repeat**
9:     merge two "closest" clusters based on the linkage criterion
10:     update the proximity matrix
11: **until** k_i clusters remain
12: **return** k_i location clusters with sizes of $m_1 \times N_i$, $m_2 \times N_i$, ..., $m_{k_i} \times N_i$, where $\sum_{j=1}^{k_i} m_j = M$

The channel clustering algorithm is illustrated in Algorithm 1. As shown in Line 6-7, we let each channel object be its own cluster at the initial stage and compute pair-wise proximity matrix (*i.e.,* Euclidean distance). Then we apply Ward's minimum variance method to merge two "closest" clusters and update the proximity matrix accordingly (Line 9-10). We repeat the process until a required number of clusters is obtained. In case of multiple training data sets as the input, we can extend the dimension of channel objects by treating signal strengths across different locations across different days for a channel as a 2-dimension vector, and perform channel clustering in the same way. One potential advantage of training with multiple data sets is that we can improve clustering accuracy and be robust to noise in individual data sets. The location clustering algorithm (Algorithm 2) is conducted in a similar way, except that we let the vector containing signal strengths in different channels at a location be a location object during the whole process. We find our cluster algorithms achieving decent performance in experiments to be discussed in Section 6. However, they do not necessarily achieve the best possible performance for the problems considered in this study. It would be an interesting direction to consider alternative linkage criteria and similarity metrics in the clustering algorithm design to further improve the performance.

5.2.2 Indoor Sensor Placement

As mentioned in Section 4.5.4, we need to meet coverage and performance requirements in determining sensor location based on the result of channel-location clustering. To avoid brute-force enumeration of indoor sensor locations (*e.g.,* C_{65}^{N} given 65 indoor locations and N indoor sensors), we introduce a greedy sensor placement algorithm as shown in Algorithm 3. Initially, all CL groups are not covered, and all indoor locations are candidates for sensor placement. We rank candidates according to their ranks (Line 5). More details about the ranking algorithm (Algorithm 4) will be illustrated later. In Line 6-12, we check for each candidate one by one, starting from the highest ranking one, whether any not-covered CL groups can be covered if we place one sensor at that location. If yes, we remove the covered CL groups from the not-covered list, add this location to the sensor location pool and repeat the process for the remaining candidate locations; otherwise, we exam the next candidate until a suitable one is found. Our algorithm terminates the while loop (Line 4-13) when all initially not-covered CL groups are covered by at least one sensor. Since one indoor location may belong to different CL groups with respect to different channel groups, it is possible that we are able to cover multiple CL groups by deploying one sensor, and cover all CL groups with fewer sensors. Hence, upon the termination of the while loop, our algorithm guarantees that all CL groups will be covered. If there are still additional quota for sensor locations, we rank remaining candidates again and choose higher-ranking locations consecutively until the expected number of sensor locations is reached (Line 14-17).

Algorithm 3 Indoor Sensor Placement

1: **Input:** N: given sensor quota
2: **Output:** L: a set of sensor locations
3: **Initialize:** $L \leftarrow NULL$
4: **while** there are not-covered Channel-Location groups **do**
5:     $C \leftarrow$ remaining ranked candidate locations
6:     **for** $i = 1$ to length(C) **do**
7:         **if** $C[i]$ can cover any not-covered Channel-Location group **then**
8:             $L \leftarrow L + C[i]$
9:             $N \leftarrow N - 1$
10:             **break**
11:         **end if**
12:     **end for**
13: **end while**
14: **if** $length(L) < N$ **then**
15:     $C \leftarrow$ ranked remaining candidate locations
16:     $L \leftarrow$ first $(N - length(L))$ candidates in C
17: **end if**
18: **return** L

The ranking algorithm is crucial in fulfilling the performance requirement, as illustrated in Algorithm 4. At each stage, we rank candidate locations by comparing their "contribution" to the overall performance, which is quantized in items of evaluated FA rate and WS Loss rate after their individual participation in the current sensor pool. In Line 5-8, we first add a candidate location to existing sensor location pool, and then compute corresponding FA rate and WS Loss rate f' and w'. We repeat the process for all candidate locations, and then compare their FA rates and WS Loss rates. By applying the "least-false-alarm-first" criterion, we will finally obtain the desirable ranking for given candidate locations (Line 10).

6. PROOF-OF-CONCEPT WISER PROTOTYPE AND EXPERIMENTS

In this section, we deploy a proof-of-concept WISER prototype and conduct actual experiments across four months to evaluate its performance. The WISER prototype takes user indoor positions as inputs, and outputs a list of vacant channels at the positions. Our objectives are: (i) to evaluate the performance of the deployed WISER prototype, and compare it with alternative solutions in a typical indoor scenario; (ii) to demonstrate its ability of providing accurate

Algorithm 4 Candidate Location Ranking

1: **Input:** S:already selected sensor locations, M:candidate locations
2: **Output:** M': ranked candidate locations
3: **Initialize:** $F \leftarrow NULL$ {FA rate}, $W \leftarrow NULL$ {WS Loss rate}
4: **for** $i = 1$ to $length(M)$ **do**
5:     $S' \leftarrow S + M(i)$
6:     $[f', w'] \leftarrow$ evaluated performance for S'
7:     $F[i] \leftarrow f'$
8:     $W[i] \leftarrow w'$
9: **end for**
10: $M' \leftarrow$ ranked M according to (F, W) {Candidate locations with less FA rate (first priority) and less WS Loss rate get higher ranking}
11: **return** M'

indoor white space information across a long period of time (in this case four months) and that there is no need to re-configure WISER frequently; (iii) to demonstrate how to choose a suitable number of indoor sensors for the target indoor scenario; (iv) to illustrate the impact of indoor positioning errors on WISER's performance.

6.1 Scenario and Settings

6.1.1 Scenario

We build a proof-of-concept WISER prototype for one floor of a typical office building to evaluate its performance using actual experiments. As shown in Fig. 2, this scenario contains individual office rooms with various sizes, common rooms of larger sizes, and corridors, all of which are separated by walls or glass. A total of 13 one-time spectrum measurements were taken during the months of October 2012 to March 2013 with the same equipment and procedures as described in Section 2.2.1.

6.1.2 Settings

We set the white space threshold to be -84.5 dBm/8 MHz for digital TV channels and -104.2 dBm/100 KHz for analog channels. To capture real-world outdoor sensing, we perform outdoor long-time sensing for 30 hours by placing the equipment on the rooftop. We observe 5 channels who signal strengths are always far below the white space threshold, and hence treat them as outdoor-sensor-reported white spaces. We also remove 11 long-term strong channels (whose received signal strengths are 10 dB higher than the white space threshold) that have little chance to be white spaces indoors even in the presence of signal attenuation, and focus on tracking the availability of the remaining channels to maximally extract the indoor white space potential.

To ensure the quality of channel-location clustering , we use two sets of one-time profiling data as inputs into the clustering algorithm. We are aware of the potential over-fitting problem in clustering, and with only one one-time profiling data set, the deployed WISER tends to cause severe false alarms in the future prediction. It turns out that WISER can effectively alleviate the potential over-fitting problem by using two data sets. We use the remaining 11 data sets for actual experiments.

6.2 Performance Evaluation

For any indoor WS identification system, we are interested in the FA Rate and WS Loss Rate as defined in Section 4.1. The objectives of this experiment are (i) to quantify the performance of our WISER prototype using the above two metrics, (ii) to compare it with the performance of WISER with those of two alter-

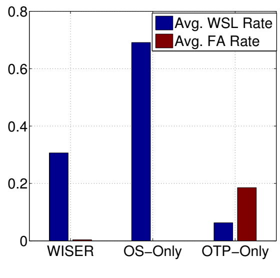
Exp. ID	1	2	3	4	5	6	7	8	9	10	11	Avg.
Extra White Space (%)	30.14	44.99	37.79	34.20	31.76	50.37	39.62	35.51	35.20	40.00	43.40	38.47
WISER False Alarm Rate (%)	0.16	0.23	0.14	0.15	0	1.08	1.09	0	0.15	0.41	0.50	0.36

Table 2: Extra white spaces and FA rate achieved by our WISER prototype as compared to the OS-Only approach in eleven experiments across four months.

natives OS-Only and OTP-Only discussed in Section 4.1, and (iii) to demonstrate WISER's ability to providing accurate white space information in a long period of time.

In this experiment, the WISER prototype consists of a total of 20 indoor sensors and one outdoor sensor. OTP-Only labels channels at each location as 0 (occupied) or 1 (available) based on one of the training data sets for the WISER prototype, and use this information to report white spaces in the future. We assume perfect indoor positioning accuracy at this stage and will study the impact of indoor positioning errors in Section 6.4.

We have several observations. First, as shown in Table 2, our WISER prototype is able to identify 30-50% more white spaces with negligible average FA rate (*i.e.,* 0.36%) as compared to OS-Only. Fig. 8a shows that OTP-Only is not suitable for identifying indoor white spaces in long term due to its unacceptable false alarm rate, and WISER outperforms OS-Only as shown in Fig. 8b. Second, WISER is able to provide accurate indoor white space information with negligible false alarms across a long period of time – four months in our experiments. Note that as WISER is built on channel-location correlations that are determined by surrounding environments, such as building structures, nearby buildings, *etc.*, it needs to be re-calibrated when environment changes substantially (*e.g.,* a skyscraper is built nearby). Although WISER may give a FA rate of 1% occasionally, one can systematically lower WISER's FA rate by using more training data sets as inputs and using more conservative signal strength threshold in determining white spaces. We believe that our principled design behind WISER is general and does not depend on particular choices of system design parameters and inputs.



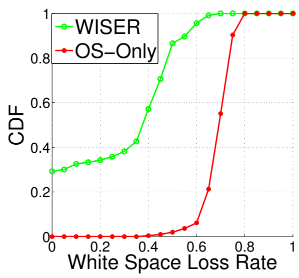
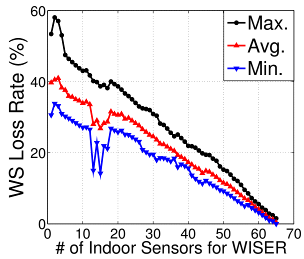


Figure 8: (a) Averaged FA rate and WS Loss rate for WISER, OS-Only and OTP-Only. The FA rate is 18.53% for OTP-Only, which may cause severe interference to primary users. As compared to OS-Only, WISER is able to report 38% more white spaces on average with negligible false alarm rate. (b) Averaged CDF curves of WS Loss rate for WISER and OS-Only. With OS-Only, 90% locations would suffer a WS Loss rate of 60%. This number is <5% if using WISER. Note that we excluded OTP-Only from this comparison, as OTP-Only will result in high FA rate despite of its seeming high WS Loss rate, which makes itself undesired in the first place. Also, we did not plot the CDF curves for the FA rate for WISER and OS-Only, as both of them result in negligible FA rate and their CDF curves for the FA rate are not very informative.

6.3 Number of Indoor Sensors to Use in WISER

For a specific indoor scenario, it is important to balance between WISER performance and the sensor cost. To understand their relationship, we vary the number of indoor sensors during WISER deployment and evaluate WISER performance in actual experiments under the same setting.

As shown in Fig. 9, the WS Loss rate and FA rate decrease accordingly with more sensors. Note that the average FA rate drops to 0.4% when there are over 18 indoor sensors and remain rather low afterwards. To meet certain safety requirement (*e.g.,* less than 0.4% average FA rate in this example), the optimal number of indoor sensors would be 18; if the worst-case FA rate needs to be lower than 1.0%, WISER requires a minimum of 20 indoor sensors. In practice, we would jointly consider safety, efficiency and sensor cost to determine the optimal number of indoor sensors required. It also implies that for a typical indoor scenario, we could conduct multiple sets of one-time profiling, using some of them for WISER deployment and others for probing the suitable sensor number.



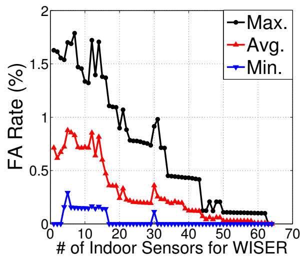


Figure 9: (a) The average, minimum (best-case), and maximum (worse-case) WS Loss rate v.s. the number of indoor sensors. The more indoor sensors are used, the lower WS Loss rate WISER is able to achieve. (b) The average, minimum (best-case), and maximum (worst-case) FA rate v.s. the number of indoor sensors. The FA rate tends to be high with fewer indoor sensors used, and gets lower when the number of indoor sensors increases. The average rate becomes stable when there are more than 18 indoor sensors.

6.4 Impact of Indoor Positioning Errors

Previous experiments show that WISER has good performance if the indoor positioning is accurate. In practice, indoor positioning may incur errors. Therefore, it is important to evaluate the performance of our WISER prototype in the presence of indoor positioning errors.

Ideally, WISER returns the accurate indoor WS information at the user's position that is determined by the indoor positioning system without any false alarms. However, due to positioning errors, the user could appear in anywhere within the circle that is centered at the reported position with a radius of a certain positioning error (in meter). Positioning errors are problematic as they can manifest themselves in terms of both white space loss and (even worse) false alarms. To avoid false alarms, WISER returns the commonly

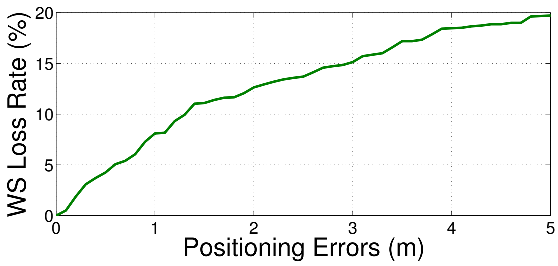


Figure 10: Additional WS Loss rate (averaged) introduced by indoor positioning errors. The WISER prototype consists of 20 indoor sensors and one outdoor sensor.

available channels within the circle, which would result in an additional loss of white spaces. In this way, indoor positioning errors will not result in additional false alarms. To quantify this impact of positioning errors, we compute WS Loss rates of WISERs with and without positioning errors respectively, and measure this additional loss against the positioning error. We show the average case for the 65 locations in Fig. 10. Note that with a positioning error of 5 meters, we would incur an average of 20% additional WS loss. To ensure the efficiency of WISER, the indoor positioning system should incur positioning error less than a couple of meters.

7. RELATED WORK

Existing works on TV White Space mostly focus on outdoor scenarios. These works include setting up a single Wi-Fi like AP system [10], implementing a network system for dynamic spectrum access [14, 37], and designing a prototype system for spectrum sensing [11]. In this paper, we develop a white space identification system for indoor scenarios where outdoor solutions do not perform well. We first carry out a large-scale measurement of outdoor and indoor white spaces, and then design WISER based on our observations. In fact, several measurement studies have investigated the opportunity for using white spaces in outdoor environments. Some of them focus on metropolitan cities including Chicago [26, 34], Singapore [20] and Guangzhou [13]. However, those measurements are limited in scope, whereas our outdoor measurement is a large-scale one that better capture the spectrum occupancy patterns across the entire city. Meanwhile, there are existing works focus on quantitative analysis on the availability of white spaces. In [29], the authors provide a theoretical analysis on the spectrum opportunity of TV white spaces in United Kingdom. [21] investigates the UHF white spaces in 11 Europe countries via an estimated methodology. Both the measurements and quantitative results show that there exists abundant white space spectrum in those cities. Different from their quantitative analysis based on propagation models or coverage maps, we base our analysis on more accurate local spectrum measurements.

As spectrum sensing is no longer compulsory due to sensing efficacy, the FCC mandated the use of a geo-location service at a given location [2]. The geo-location approach has been widely studied [15, 18, 25, 28]. These works use databases to store dynamic white space information for clients querying. However, in metropolitan cities, such as Hong Kong, wireless environment tends to be more complicated due to blocking effects of high density buildings and uneven distribution of TV towers in Hong Kong [36]. Actually, by comparing calculated results using different models and ground-truth measurements, the model-based approach fails to identify many vacant TV channels, and local spectrum measurements may obtain more accurate white space availability results. In [19], the authors investigate the white space availability in a different perspective. They quantify the white space availability in terms of pollution from existing television stations and self-interference, as well as the expected transmission rage of white space devices.

Although 70% of spectrum demand comes from indoors [12], there is little prior work on indoor white space availability. In [30], the authors assess the feasibility of white space devices in short-range indoor environment with focus on interference mechanisms. In [31], the authors perform a series of measurements in both laboratory and real environments to verify an indoor TV white space opportunity prediction model. They both focus on how to utilize indoor white spaces instead of how to identify them. An evaluation of spectrum availability in both indoor and outdoor environments is presented in [35]. The authors perform local spectrum sensing in the band from 20 MHz to 3 GHz on the rooftop and in a room of an office building, and confirm that considerably less occupancy was measured. In [27], an experimental spectrum sensor test-bed is built to investigate indoor radio environmental map, and it is recognized to be extremely challenging unless sophisticated conditions are considered. In this paper we analyze the temporal and spatial characteristics of indoor white spaces, and propose an innovative system to identify additional indoor white spaces.

8. CONCLUSIONS AND FUTURE WORK

White spaces promise to provide additional spectrum for wireless applications. However, the current rulings in the US and Canada (and being considered elsewhere), use a geo-location database to determine spectrum availability, which although works reasonably well, has severe limitations in both outdoor and indoor settings.

In this paper, we carry out a measurement-driven approach to explore white space networking potential in Hong Kong, with the focus on indoor environments. We show that the indoor white spaces have different characteristics from the outdoor ones. For example, there are more contiguous unutilized TV channels indoors, which are able to support high bandwidth communication. We then propose a system, called WISER, that is able to utilize the additional white space channels without requiring the clients to actually sense the wireless medium. Our system is incrementally deployable. Any interested building can adopt our techniques to make the building "white space enhanced" and utilize the additionally available vacant TV channels.

There are several interesting and important future directions that could be explored. First, it would be very interesting to extend our 2-dimensional system design, deployment and experiments to 3-dimensional (*i.e.,* from one floor to multiple floors in the building) by jointly exploring channel, location and space correlations. Second, it is also important to evaluate WISER's performance on different test-beds (*e.g.,* halls, residential buildings, *etc.*) other than office buildings. Third, it is also interesting to improve the indoor sensor placement by taking into account of "hotpot" locations with human activities.

Acknowledgements

We would like to thank Wei Zhang, for his help with clustering analysis, and our shepherd, Moustafa Youssef, for helpful comments in improving the presentation of the paper. In addition, we are grateful to the anonymous reviewers for their insightful comments. The work described in this paper was partially supported by National Basic Research Program of China (Project No. 2012CB315904 and 2013CB336700) and several grants from the University Grants Committee of the Hong Kong Special Administrative Region, China (Area of Excellence Grant Project No. AoE/E-02/08 and General Research Fund Project No. 411010 and 411011).

9. REFERENCES

[1] Districts of hong kong. http://en.wikipedia.org/wiki/Districts_of_Hong_Kong.

[2] Fcc frees up vacant tv airwaves for “super wi-fi" technologies and other technologies. http://www.fcc.gov/sptf/reports.html.

[3] Gnu radio. http://gnuradio.org/redmine/projects/gnuradio/wiki.

[4] Hong kong digital terrestrial tv coverage & reception database. http://app1.ofca.gov.hk/apps/ubs/map.asp.

[5] Location of the digital terrestial television (dtt) stations and the estimated coverage in hong kong. http://www.digitaltv.gov.hk/general/pdf/coverage.pdf.

[6] Second memorandum opinion and order. http://hraunfoss.fcc.gov/edocs_public/attachmatch/FCC-10-174A1.pdf.

[7] Television broadcasting network of hong kong. http://www.ofca.gov.hk/filemanager/ofca/common/industry/broadcasting/television/free_tv/tvbnet.pdf.

[8] Tv white space geolocation database. http://ieee802.org/19/pub/Workshop/5_Gibson-ComSearch.pdf.

[9] Universal software radio peripheral. http://www.ettus.com.

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

[11] R. Balamurthi, H. Joshi, C. Nguyen, A. Sadek, S. Shellhammer, and C. Shen. A tv white space spectrum sensing prototype. In *Proc. of IEEE DySPAN*, 2011.

[12] V. Chandrasekhar, J. Andrews, and A. Gatherer. Femtocell networks: a survey. *Communications Magazine, IEEE*, 46(9):59 –67, september 2008.

[13] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li. Mining spectrum usage data: a large-scale spectrum measurement study. In *Proc. of ACM MOBICOM*, 2009.

[14] S. Deb, V. Srinivasan, and R. Maheshwar. Dynamic spectrum access in dtv whitespaces: Design rules, architecture and algorithms. In *Proc. of ACM MOBICOM*, 2009.

[15] X. Feng, J. Zhang, and Q. Zhang. Database-assisted multi-ap network on tv white spaces: Architecture, spectrum allocation and ap discovery. In *Proc. of IEEE DySPAN*, 2011.

[16] C. Fraley and A. E. Raftery. How many clusters? which clustering method? answers via model-based cluster analysis. *The Computer Journal*, 41(8):578–588, 1998.

[17] Y. Gu, A. Lo, and I. Niemegeers. A survey of indoor positioning systems for wireless personal networks. *Communications Surveys Tutorials, IEEE*, 11(1):13 –32, quarter 2009.

[18] D. Gurney, G. Buchwald, L. Ecklund, S.Kuffner, and J.Grosspietsch. Geo-location database techniques for incumbent protection in the tv. In *Proc. of IEEE DySPAN*, 2008.

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

[20] M. Islam, C. Koh, S. Oh, X. Qing, Y. Lai, C. Wang, Y. Liang, B. Toh, F. Chin, G. Tan, and W. Toh. Spectrum survey in singapore: Occupancy measurements and analyses. In *Cognitive Radio Oriented Wireless Networks and Communications, CrownCom 2008.*, 2008.

[21] V.D. Jaap, R. Janne, A. Andreas, and M.Petri. Uhf white space in europe: a quantitative study into the potential of the 470-790 mhz band. In *Proc. IEEE DySPAN*. IEEE, 2011.

[22] P. Jallon. An algorithm for detection of dvb-t signals based on their second-order statistics. *EURASIP Journal on Wireless Communications and Networking*, 2008:28, 2008.

[23] J. Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963.

[24] N. Klepeis, W.Nelson, and W.Ott. The national human activity pattern survey (nhaps): A resource for assessing exposure to environmental pollutants. *Journal of Exposure Analysis and Environmental Epidemiology*, 11:231–252, 2001.

[25] D. Makris, G. Gardikis, and A. Kourtis. Quantifying tv white space capacity: A geolocation-based approach. In *IEEE Communications Magazine: Topics in Radio Communications*. IEEE, 2012.

[26] M. McHenry, P. Tenhula, D. McCloskey, D. Roberson, and C. Hood. Chicago spectrum occupancy measurements & analysis and a long-term studies proposal. In *Proc. of the first international workshop on Technology and policy for accessing spectrum*. ACM, 2006.

[27] E. Meshkova, J. Ansari, D. Denkovski, J. Riihijarvi, J.Nasreddine, M.Pavloski, L.Gavrilovska, and P.Mahonen. Experimental spectrum sensor testbed for constructing indoor radio environmental maps. In *Proc. of IEEE DySPAN*, 2011.

[28] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl. Senseless: A database-driven white spaces network. In *Proc. of IEEE DySPAN*, 2011.

[29] M. Nekovee. Quantifying the availability of tv white spaces for cognitive radio operation in the uk. In *Proc. IEEE ICC Workshop: Cognitive Wireless Networks and Systems*. IEEE, 2009.

[30] E. Obregon and J. Zander. Short range white space utilization in broadcast systems for indoor environment. In *Proc. of IEEE DySPAN*, 2010.

[31] R. Obregon, L. Shi, J. Ferrer, and J. Zander. Experimental verification of indoor tv white space opportunity prediction model. In *IEEE CROWNCOM*. IEEE, 2010.

[32] Office of the Communications Authority. Hong kong table of freqency allocaitons. Technical report, Office of the Communications Authority, 2012.

[33] W.M. Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical Association*, 66(336):846–850, 1971.

[34] T. Taher, R. Bacchus, K. Zdunek, and D. Roberson. Long-term spectral occupancy findings in chicago. In *Proc. IEEE DySPAN*, pages 100–107. IEEE, 2011.

[35] M. Wellens, J. Wu, and P. Mahonen. Evaluation of spectrum occupancy in indoor and outdoor scenario in the context of cognitive radio. In *Cognitive Radio Oriented Wireless Networks and Communications*, 2007.

[36] L. Yan and M. Chen. Exploring white spaces in metropolises: A measurement study in hong kong. Technical report, The Chinese University of Hong Kong, 2011.

[37] L. Yang, W. Hou, L. Cao, B. Zhao, and H. Zheng. Supporting demanding wireless applications with frequency-agile radios. In *Proc. of ACM NSDI*, 2010.