FIWEX: Compressive Sensing Based Cost-Efficient Indoor White Space Exploration*

Dongxin Liu†, Zhihao Wu†, Fan Wu†, Yuan Zhang§, and Guihai Chen†
†Shanghai Key Laboratory of Scalable Computing and Systems, Shanghai Jiao Tong University, China
§Department of Computer Science and Technology, Nanjing University, China
{liu-dx,sjtu_wuzhihao,wu-fan}@sjtu.edu.cn, zhangyuan@nju.edu.cn, gchen@cs.sjtu.edu.cn

ABSTRACT

Exploring the utilization of white spaces (vacant VHF and UHF TV channels) is a promising way to satisfy the rapid growth of the radio frequency (RF) demand. Although a few outdoor white space exploration methods have been proposed in the past few years, researches that focus on indoor white space exploration just emerge recently. In this paper, we propose a novel cost-efficient indoor white space exploration method by exploiting the location dependence and channel dependence of TV channels' signal strength in indoor environments. We measure the UHF TV channels in a building, and study the temporal and spatial features of indoor white spaces. Based on the extracted features, we design a cost-e**F**fficient **I**ndoor **W**hite space **EX**ploration mechanism, namely FIWEX. Furthermore, we build a prototype of FIWEX and extensively evaluate its performance in real world environments. The evaluation results show that FIWEX can identify 47.8% more indoor white spaces with 38.4% less false alarms compared with the best known existing solution.

Categories and Subject Descriptors

C.2.1 [**Computer-Communication Networks**]: Network Architecture and Design – *Wireless Communication*

Keywords

White Space, Compressive Sensing

†F. Wu is the corresponding author.

*This work was supported in part by the State Key Development Program for Basic Research of China (973 project 2014CB340303), in part by China NSF grant 61422208, 61402223, 61472252, 61272443, and 61133006, in part by CCF-Intel Young Faculty Researcher Program and CCF-Tencent Open Fund, in part by the Scientific Research Foundation for the Returned Overseas Chinese Scholars, and in part by Jiangsu Future Network Research Project No. BY2013095-1-10. The opinions, findings, conclusions, and recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies or the government.

MobiHoc'15, June 22–25, 2015, Hangzhou, China.

http://dx.doi.org/10.1145/2746285.2746298.

1. INTRODUCTION

The growth of wireless networks are currently facing an increasing shortage of available radio frequency spectrums, since the number of mobile devices and their related applications is raising rapidly. However, the amount of unlicensed spectrums that are free to use is very limited. To deal with this problem, the concept of Dynamic Spectrum Access (DSA) that aims at exploring the opportunity of sharing and utilization of licensed spectrums among both licensed users and unlicensed users is proposed and has been applied to several applications [30, 31].

In 2008, FCC (Federal Communications Commission) issued a historic ruling that allowed unlicensed devices to use the TV spectrum that is not locally occupied by licensed devices. (The unoccupied TV spectrum is often referred to as TV white spaces or simply white spaces). After that, the TV white spaces receive more and more attentions from DSA developers. Although white spaces are open for unlicensed use now, FCC also required that unlicensed white space devices should not interfere with the licensed devices (TV broadcasts). Therefore, it is essential for all user devices (especially the unlicensed ones) to find out whether a spectrum band is available before using it in communications.

To explore white spaces, there are mainly two approaches: spectrum sensing approach and geo-location database approach. The spectrum sensing approach, which is less widely used, relies on the user devices to perform spectrum sensing. Thus, it requires the devices to be equipped with proper sensing hardware and to have enough power for sensing and signal analysis. On the contrary, the geo-location database approach that is currently used by most white space user devices and standards does not need the user devices to sense and thus saves their power. Instead, a user device gets to know the availability of white spaces by querying an online geo-location database that stores a "white space availability map" indicating spectrum's availability information at different locations.

For the geo-location database approach, the map of white space availability is vital. In this paper, we study how to efficiently construct a "white space availability map", which is used by the geo-location database approach in indoor environments. Most prior works [26–28] on white space exploration only focus on outdoor scenarios. Since there is relatively few obstructions in outdoor environments, existing works use signal propagation models to infer whether a spectrum band at a specific location is available or not. However, if we directly apply these approaches to indoor environments, which generally have many man-made obstacles, it is likely that we would get results that are overly conservative.

To solve the problem of indoor white space exploration, we propose a cost-eFficient Indoor White space EXploration

mechanism, called FIWEX. Intuitively, one can always solve the problem by deploying a large number of RF-sensors that cover the entire floor in a building. However, this could incur a large amount of expense on the sensors. By strategically deploying only a limited number of sensors, FIWEX is able to reconstruct the white space availability map of the entire floor with a small number of RF sensors at high accuracy. Due to the innovative utilizations of 1) the existence of *strong channel*, 2) the *location dependence* as well as *channel dependence* of channels' signal strength in indoor environments, 3) the compressive sensing technique, and 4) a well designed *k-medoids*-based sensor deployment method, FIWEX features a promising efficiency on the total number of sensors deployed and reconstruction accuracy.

The main contributions of this paper are as follows.

- We perform indoor white space measurements in a building for two weeks. The measurement results confirm that (1) there exists strong channels whose signal strength is at least 5dB greater than the white space threshold. (2) indoor white spaces' signal strength has both location dependence and channel dependence. These two key characteristics allow us to efficiently explore indoor white spaces without making a very dense deployment of RF-sensors over the entire building.

- We propose FIWEX, a cost-efficient indoor white space exploration mechanism that does not require user devices to sense the spectrum by themselves. By taking strong channels, location dependence, and channel dependence into consideration, we propose a novel data reconstruction algorithm based on compressive sensing technique to reduce the number of sensors needed for detection. In addition, we design an innovative sensor deployment method to improve the reconstruction accuracy.

- We have built a prototype of FIWEX, and evaluated its performance with real-world experiments. First, we evaluate FIWEX's performance compared to WISER, the best indoor white space identification system so far. In average, FIWEX can identify 47.8% more indoor white spaces with 38.4% less false alarms compared to WISER. Then, we studied the performance versus the number of indoor sensors.

The remainder of the paper is organized as follows. We introduce our indoor white space measurement experiments in Section 2. We describe system model of FIWEX in Section 3. In Section 4, we provide detailed algorithms of FIWEX, including compressive sensing based data reconstruction and the sensor deployment method. In Section 5, we present evaluation results. Related work and conclusions are in Section 6 and Section 7, respectively.

2. INDOOR WHITE SPACE AVAILABILITY MEASUREMENT

We perform our indoor white space measurement in the third floor of a building. The purpose of the indoor white space measurement is exploring strong channels and location-channel dependence of indoor white spaces. These features will be utilized in designing the indoor white space exploration mechanism in Section 3.

2.1 Measurement Setup

The measurement device (Figure 1) we used consists of a USRP N210, an omni-directional receive antenna, a laptop and a uninterrupted power supply (UPS). The daughterboard of our USRP is SBX with 5-10 dBm noise figure. We

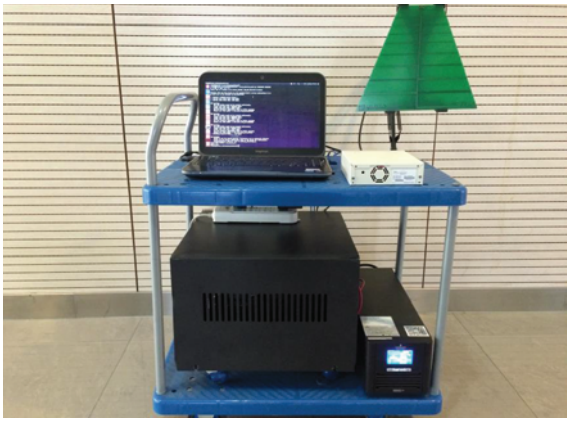


Figure 1: The measurement device.

calibrate the device using a RF signal generator to get the accurate signal strength. As suggested by FCC [1], the gain of the antenna is also taken into consideration during the calibration process.

A number of different methods were proposed for identifying the presence of signal transmissions, such as energy detection, waveform-based sensing, and matched-filtering [18]. Here, we choose energy detection, since it is the most common way with low computation and implementation complexities. In our measurements, we judge whether a TV channel is vacant by comparing the channel's signal strength with a threshold, which depends on the noise floor [19]. If the signal strength of a TV channel is greater than the threshold, we label this channel as locally occupied, otherwise the channel is labeled as vacant. We measure the digital TV channels between 470 MHz - 566 MHZ and 606 MHz - 870 MHz with 8 MHz channel bandwidth, and use the same threshold -84.5 dBm/8 MHz as [2] for digital TV signals. Due to the hardware limitations, the vacant channels determined using the above mentioned threshold may be not safe to use, but the observations drawn from the measurement are general, and our mechanism is not limited to any specific threshold. We believe that if the sensitivity of the measurement hardware can support a threshold of -114 dBm as suggested by FCC [1], our mechanism can be safely used in practice.

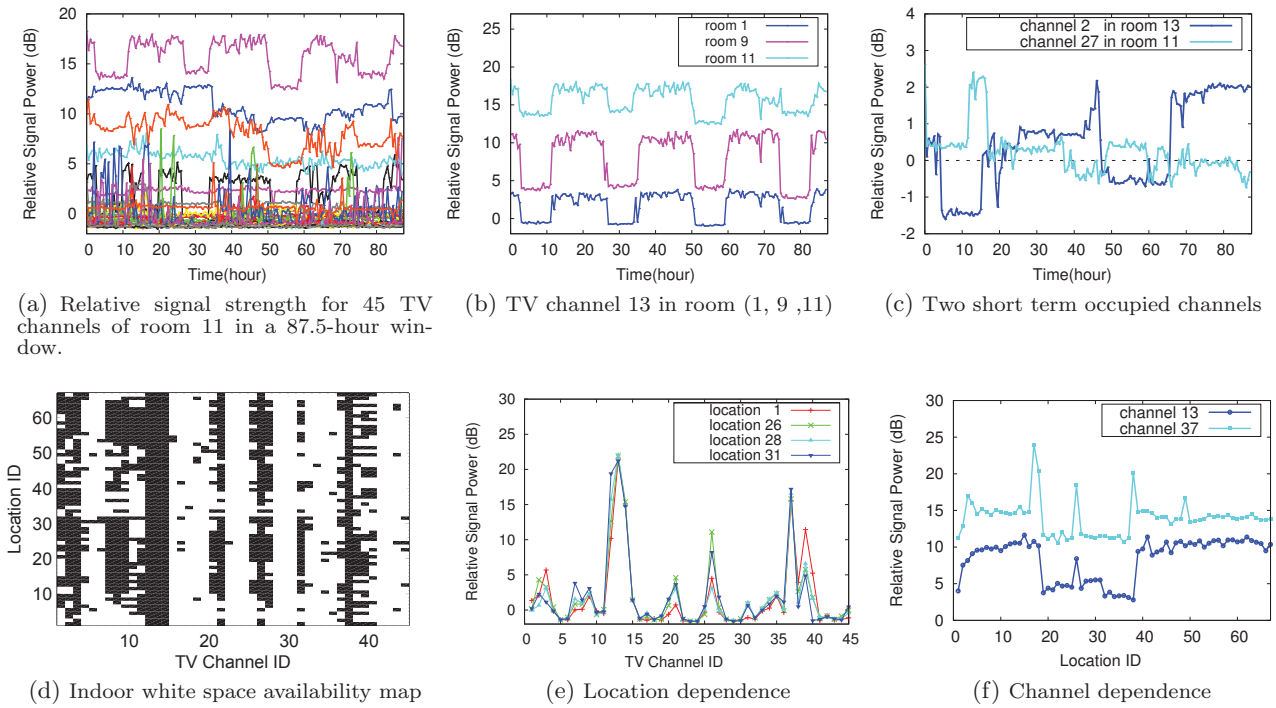
We implement the energy detector with a bin size of 1024 and sampling rate 4 MHz. We calculate the power of a channel using the average value of all corresponding bins. During the indoor white space measurement experiment, we measure 45 channels in total.

Our measurement is divided into two parts, namely *long-time sensing* and *short-time sensing*.

2.2 Long-Time Sensing

In long-time sensing, we randomly choose 20 rooms and deploy a USRP coupled with a laptop in every room. The absolute signal strength for all 45 TV channels is measured for 87.5 hours in total. For the convenience of comparisons, we convert absolute power to relative power by subtracting the white space threshold from the absolute power. Devices in different rooms are synchronized using "crontab" of Unbuntu 12.04, and the spectrum sensing programs of different computers run synchronously every half an hour.

We have the following observations from long-time sensing results. First, as shown in Figure 2(a), there exists strong channels whose relative signal strength is obviously greater than 0 during most of the time. For example, we observe 3 channels with >5dB relative signal strength during the



(a) Relative signal strength for 45 TV channels of room 11 in a 87.5-hour window.

(b) TV channel 13 in room (1, 9 ,11)

(c) Two short term occupied channels

(d) Indoor white space availability map

(e) Location dependence

(f) Channel dependence

Figure 2: Results of indoor white space measurement

whole 87.5 hours interval. That means these three channels are long time occupied in room 11.

Second, we observe that a channel may have different signal strengths in different locations. As shown in Figure 2(b), the relative signal strength of channel 13 in room 11 is greater than 5 dB during the whole measurement period; in room 9, the channel's relative signal strength moves up and down around 5 dB; in room 1 it is always below 5 dB.

Finally, short term occupied channels exist. We refer a channel which is occupied for some periods of time and vacant for the rest of time as a short term occupied channel. Figure 2(c) shows two short term occupied channels (channel 2 in room 13 and channel 27 in room 11) whose relative signal strength moves up and down around 0.

Long-time sensing experiment results show the existence of strong channels and short term occupied channels. If we find strong channels and consider them always strong in the future, then we can pay our attention to the short term occupied channels, and in this way, the resources (e.g., RF-sensors, energy) can be more efficiently utilized. Actually, there exists weak channels whose signal strength is lower than the white space threshold in our measurement experiments, but we don't consider them as always weak in order to protect the licensed user.

2.3 Short-Time Sensing

Indoor short-time sensing experiments collect the spatial-channel features of indoor white space. Our indoor white space exploration mechanism was designed based on the data collected during this process.

We choose 67 typical locations labeled as location 1 to location 67, and measure all 45 TV channels at every location and record the received signal strength using our movable device (a cart with a USRP, an antenna, a laptop, and a UPS). After each round of short-time sensing, we get a *Measurement Matrix (MM)*, which is a 67×45 matrix denoted by M. M contains the absolute signal strength for 45 channels at 67 locations, each row/column of which represents a location/channel. We perform 14 rounds of short-time sensing in a period of two weeks and get 14 measurement matrices in total. From the short-time sensing, we have the following observations.

First, Figure 2(d) shows the indoor white space availability map where white blocks represent white spaces and black blocks represent occupied channels. We observed that the signal strength of a channel at different locations could be different (a channel could be occupied at some locations while be vacant at others). This is mainly caused by the complex indoor signal attenuation patterns due to obstructions (e.g., walls). If we directly use the outdoor white space exploration approaches to indoor scenarios, we will lose a lot of white spaces.

Second, in Figure 2(e), the relative signal strength of all channels is similar at four different locations (1, 26, 28, 31). Prior work [2] described this kind of relationship between different locations according to locations' similarity. They treat the correlated locations as a group and represents the white space availability of all locations in this group using only one of them. It is indeed a creative way to utilize the similarity relationship between locations. However, in our mechanism, we consider not only the similarity relationship but also the linear dependence between locations. *Location dependence* represents the linear dependence between different locations, that means any row of M can be approximately represented as a linear function of some of the other rows. Assume that M_i is ith row of M (M_i is a 1×45 row vector containing the signal strength of 45 channels at location i), we can approximate M_i as

$$M_i \approx a_0 M_0 + a_1 M_{i_1} + a_2 M_{i_2} + ... + a_k M_{i_k},$$

where M_0 is a 1×45 row vector equals to $(1, 1, ..., 1)$ and $M_{i_1}, M_{i_2}, ..., M_{i_k}$ $(i_1, i_2, ..., i_k \neq i)$ are the i_1th, i_2th, ..., i_kth rows of MM, and $a_0, ..., a_k$ are the weight parameters. Similar to location dependence, we define *channel dependence*

as the linear dependence of channels. As shown in Figure 2(f), although the similar relationship is not suitable here as the signal strength of channel 13 and channel 37 are not close, our linear dependence relationship works well since the difference of them is almost fixed in different locations. According to the definition of relative location-channel matrix X (Section 3.1), we got that location-channel dependence also exists in X, since $M - X$ is a constant matrix. In Section 4.3, we introduce a novel way to draw the location-channel dependence.

2.4 Summary

The indoor white space measurements give us a better understanding about the characteristics of indoor white spaces. Below we summarize guidelines for designing FIWEX.

- Strong channels are occupied for most of time, thus we can neglect these channels once we spot them.
- Location dependence and channel dependence allow us to infer a channel's signal strength or status based on its correlated channels or locations.

In next section, we will show how to use these guidelines to make FIWEX cost-efficient and accurate.

3. SYSTEM MODEL

FIWEX aims to accurately identify indoor white spaces with a small number of RF sensors deployed. As shown in Figure 3, FIWEX is mainly composed of two parts: central server and real time sensing module. The central server receives the real-time sensing results reported by the indoor sensors, and maintains the historical sensing records. Upon receiving a query for availability of the white space at a certain location from a user, the central server calculates the up-to-date indoor white space availability map, and make a respond.

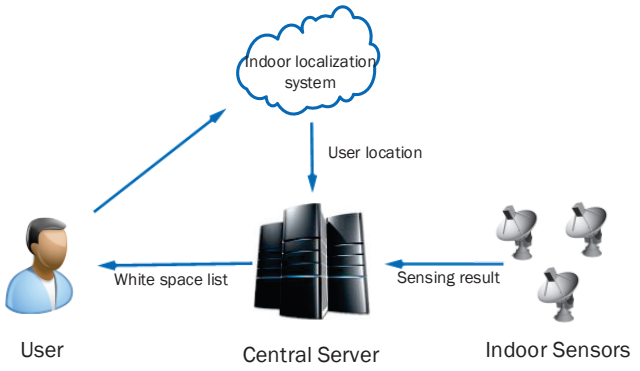


Figure 3: System architecture of FIWEX.

3.1 Real Time Sensing Module

We describe the real time sensing module using our test building with 67 locations as an example. The real time sensing module's job is choosing a part of these 67 locations, placing a sensor at each chosen location, performing a "partial sensing" (since not all 67 locations have sensors deployed), and sending the data to the central server.

If the real time sensing module deploys sensors at all 67 locations, clearly the central server can get complete short-time sensing data. However, due to the cost considerations, real time sensing module aims to use less number of sensors to get complete sensing data.

For ease of presentation, we define the following matrices:

- **Relative Location-Channel Matrix (RLCM)**: is a 67×45 matrix recording channels' relative signal strength. RLCM is denoted by X, where $X(i,j)$ is the relative signal strength of channel j at location i.

$$X(i,j) = M(i,j) - TH,$$

where TH is the white space threshold. If $X(i,j) < 0$, then channel j is vacant at location i, otherwise occupied.

- **Binary Index Matrix (BM)**: is a 67×45 matrix, which indicates where the sensors are deployed. BM is denoted by B, and for any channel j between 1 and 45

$$B(i,j) = \begin{cases} 0 & \text{if no sensor is deployed at location } i, \\ 1 & \text{otherwise.} \end{cases}$$

- **Direct Sensory Matrix (DM)**: is a 67×45 matrix, which records the relative signal strength at locations with sensors deployed. For locations without a sensor, the corresponding rows in DM contain 45 0s. DM is denoted by D:

$$D(i,j) = \begin{cases} X(i,j) & \text{if } B(i,j) = 1, \\ 0 & \text{if } B(i,j) = 0. \end{cases}$$

This means $D = B \circ X$, where $\circ$ refers to the Hadamard product. ($D = B \circ X$ means $D(i,j) = B(i,j)X(i,j)$).

Given a specific number of sensors, different deployments of these sensors lead to different performance of FIWEX. (We will discuss our indoor sensor deployment method in Section 4.5). Once a sensor deployment is determined, B is fixed. Sensors deployed at different locations collect absolute signal strength in real time and the corresponding relative ones are recorded in D. After that, real time sensing module submits matrix D to the central server at regular intervals.

3.2 Central Server:

The central server of FIWEX consists of two parts: *data reconstruction* and *white space database*.

In data reconstruction part, a complete relative location-channel matrix is reconstructed based on the direct sensory matrix (D). We define the reconstructed matrix as follows:

- **Reconstructed Matrix (RM)** : is a 67×45 matrix generated by interpolating the missing values in D. RM is denoted by $\tilde{X}$.

Data reconstruction part aims to find $\tilde{X}$ that approximates X as accurate as possible.

Strong channels can be utilized to improve the accuracy of data reconstruction. In contrast to existing works, observations in Section 2.2 (Figure 2(b)) show that different locations may have different strong channels. Based on this observation, we define the strong channel matrix as follows:

- **Strong Channel Matrix (SCM)** : is a 67×45 matrix that indicates strong channels in different locations. SCM is denoted by

$$S(i,j) = \begin{cases} 1 & \text{channel } j \text{ is a strong channel at location } i, \\ 0 & \text{otherwise.} \end{cases}$$

Given D, B and S as inputs, the central server performs data reconstruction based on compressive sensing. We will present the details of this process in Section 4.

White space database receives the result of data reconstruction $\tilde{X}$, and produces the up-to-date indoor white space availability map according to $\tilde{X}$. In order to avoid

interference with licensed users, we compare $\tilde{X}(i,j)$ with PR (instead of 0) to decide whether channel j at location i is vacant. We define the reconstructed indoor white space availability map as

$$MAP(i,j) = \begin{cases} 1 & \text{if } \tilde{X}(i,j) < PR, \\ 0 & \text{if } \tilde{X}(i,j) \geq PR, \end{cases}$$

where PR is the protection range ($PR = -0.7$ by default). Hence, MAP approximately denotes the availability (1 for vacant and 0 for occupied) of indoor white spaces with a low false alarm rate.

Users submit their indoor positions to the central server through an indoor localization system [20, 21]. Given the indoor position of a user, the database first finds one of the 67 profiled locations which is the nearest to the user, and then returns the white space list of this profiled location to the user after considering the interference with neighbors. Here for simplicity, FIWEX assumes the white space availability of a non-profiled location is the same as the availability of the nearest profiled one. We leave the case where these two availability results are not equal to our future work. We have to note that our approach proposed in the paper is to efficiently identify indoor white spaces. A closely related problem on how to utilize the identified indoor white spaces without interfering is our another future work.

4. DATA RECONSTRUCTION AND SENSOR DEPLOYMENT

In this section, we present our compressive sensing based indoor white space reconstruction algorithm, and propose a cost-efficient sensor deployment method.

Compressive sensing is a generic data reconstruction technique based on the structure and redundancy of real-world signals or datasets [3–6]. So far, compressive sensing has been widely applied to different realms [7–11]. However, indoor white spaces have unique strong channels and location-channel dependence. Consequently, we cannot directly apply the traditional compressive sensing technique to the relative location-channel matrix recovery. To deal with this issue, we introduce strong channels and location-channel dependence to compressive sensing, and solve the problem using alternative least square method. Compressive sensing based data reconstruction algorithms perform very differently given different data loss patterns [8], and thus different sensor deployments may lead to different performance of FIWEX. In order to find appropriate locations to deploy sensors, we propose a *k-medoids* [13] based sensor deployment method.

4.1 Compressive Sensing

Given D, B and S, compressive sensing can help us to approximately reconstruct X. According to the theory of compressive sensing, matrices with low-rank feature can be reconstructed with a high accuracy. When the vector containing all singular values of a matrix is sparse, the matrix is low-rank. In Figure 4, we illustrate the distribution of singular values in 7 of the 14 relative location-channel matrices we got in the short time sensing experiments. The X-axis presents the i-th singular values, and the Y-axis presents the normalized values of i-th singular value. This figure suggests that the energy is always contributed by the top several singular values in X. In average, the top 25% singular values contribute most of the energy in this graph. This phenomenon reveals that X exhibits approximate low-rank structure. The followings are the details about our compressive sensing algorithm.

We begin with the Singular Value Decomposition (SVD) using the similar methodology as [7]. SVD is a kind of factorization of a matrix, and it is a useful tool for creating low-rank matrix approximation. Here, we generalize the discussion of the relative location-channel matrix (X) to an $m \times n$ matrix. For the $m \times n$ matrix X, there exists a factorization of the form

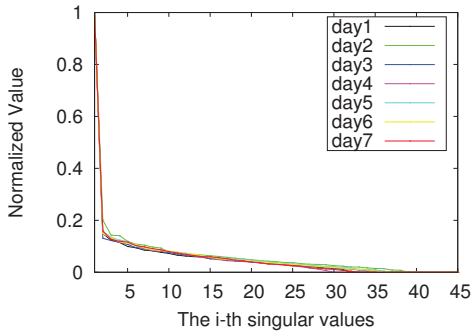


Figure 4: Low-rank feature of X.

$$X = U\Sigma V^T = \sum_{i=1}^{min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (1)$$

where U is an $m \times m$ unitary matrix (*i.e.*, $UU^T = U^T U = I$), V^T is the transpose of an $n \times n$ unitary matrix, Σ is an $m \times n$ diagonal matrix with the singular value σ_i of X on the main diagonal, where $\sigma_i \geq \sigma_{i+1}$. Here $\mathbf{u}_i$ and $\mathbf{v}_i$ are the ith columns of U and V, respectively. As we mentioned before, the top 25% singular values of X contribute most of the sum of all singular values, and this means that

$$\sum_{i=1}^{r} \sigma_i \approx \sum_{i=1}^{min(m,n)} \sigma_i \quad (2)$$

and

$$\sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \approx \sum_{i=1}^{min(m,n)} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (3)$$

where $r \ll min(m,n)$. Hence we can approximately represent X as

$$\tilde{X} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad (4)$$

where r is the rank of $\tilde{X}$. Actually, $\tilde{X}$ is the best r-rank approximation that minimize the Frobenius norm $\|\cdot\|_F$ between X and $\tilde{X}$. That is, $\tilde{X}$ is the solution to:

$$\begin{aligned} \text{Minimize} \quad & \|X - \tilde{X}\|_F, \\ \text{Subject to} \quad & \text{rank}(\tilde{X}) \leq \text{r}. \end{aligned} \quad (5)$$

In the indoor white space exploration, we are given D and required to estimate X. Then, we can judge whether a channel is busy or not at a given location. It is impossible to directly solve (5), since we do not know the value of original matrix and the proper rank. Instead, we use $B \circ X = D$ as the constraint. Considering the low rank property of X, we can alternatively solve the following rank minimization problem:

$$\begin{aligned} \text{Minimize} \quad & rank(X), \\ \text{Subject to} \quad & B \circ X = D. \end{aligned} \quad (6)$$

However, it is also difficult to solve this rank minimization problem as it is non-convex. So, we transform it to the

nuclear norm minimization problem in help of the SVD-like factorization of X:

$$X = U\Sigma V^T = LR^T, \tag{7}$$

where $L = U\Sigma^{1/2}$, $R = V\Sigma^{1/2}$.

According to the compressive sensing theory [3, 4], we can perform the rank minimization by solving the nuclear norm minimization problem for the low rank matrix LR^T, if the isometry property [4] holds on binary index matrix B. Therefore we just need to minimize the sum of L and R's Frobenius norms:

$$\begin{aligned} \text{Minimize} \quad & \|L\|_F^2 + \|R\|_F^2, \\ \text{Subject to} \quad & B \circ (LR^T) = D. \end{aligned} \tag{8}$$

It is usually difficult to find L and R that strictly satisfy (8), because (i) the direct sensory matrix D contains noise, (ii) X is just approximately low-rank in the real world, and the rank of X may be largely different in different indoor environments. Thus, we use Lagrange multiplier method to solve (8):

$$\text{Minimize} \quad \|B \circ (LR^T) - D\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2). \tag{9}$$

In this way, the constraint $B \circ (LR^T) = D$ is not strictly enforced, and we use the Lagrange multiplier λ to control the tradeoff between the precise fit to the measurement and the rank minimization.

The compressive sensing approach (9) finds the global low-rank structure in matrix X. We can further improve the accuracy in two ways: strong channel improvement and location-channel dependence improvement.

4.2 Introducing Strong Channel

Performance of the compressive sensing based indoor white space reconstruction can be improved by taking strong channels into consideration. In contrast to most prior works, which just define a single set of strong channels shared by all indoor locations, we argue that different locations have different strong channels as shown in Section 2. Based on this observation, we define strong channel matrix S in Section 3.2 to describe the spatial variance of strong channels.

In short-time sensing experiment, we obtain 14 short-time sensing data (MMs) in total and we use 7 of them as the training set while the other 7 are used for evaluation. If the signal strength of channel j at location i are at least 5dB higher than the white space threshold in all 7 training MMs, then channel j is a strong channel at location i and $S(i,j) = 1$; otherwise, $S(i,j) = 0$. We simply consider strong channels as always busy because channels that carry strong signals are normally used for TV broadcast and TV broadcasting arrangement is normally in long term (e.g., years). Hence, S can be used in FIWEX.

To combine compressive sensing and strong channels' feature, we define new Binary Index Matrix B_s and Direct Sensory Matrix D_s considering the influence of S:

$$B_s(i,j) = \begin{cases} 1 & \text{if } S(i,j) = 1, \\ B(i,j) & \text{otherwise.} \end{cases}$$

$$D_s(i,j) = \begin{cases} ave_{ij} & \text{if } S(i,j) = 1, \\ D(i,j) & \text{otherwise.} \end{cases}$$

Here ave_{ij} is the average relative signal strength of channel j at location i in the training set. In this way, formula (9) can be rewritten as

$$\text{Minimize} \quad \|B_s \circ (LR^T) - D_s\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2). \tag{10}$$

4.3 Introducing Location-Channel Dependence

The channel dependence and location dependence (Section 2) represent the channel/location structure of indoor white spaces, and can be utilized to improve the performance of FIWEX. Taking compressive sensing, strong channels, channel dependence and location dependence into consideration, we expand (10) as follows:

$$\begin{aligned} \text{Minimize } & \|B_s \circ (LR^T) - D_s\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2) \\ & + \|P(LR^T) - P_0\|_F^2 + \|(LR^T)C - C_0\|_F^2, \end{aligned} \tag{11}$$

where P and P_0 are the location dependence constraint matrices, while C and C_0 represent the channel dependence constraint matrices. Since different choices of P, P_0, C, and C_0 yield different performance of FIWEX, we discuss how to choose these matrices in the following section.

4.3.1 Choice of location-channel dependence constraint matrices

Choice of $P_{(67\times67)}$ and $P_{0(67\times45)}$: Matrix P and P_0 represent the location dependence in relative location-channel matrix (X) and express the linear relationship between different rows of $X_{(67\times45)}$. We propose an innovative way to find appropriate P and P_0. For each row X_i of X, location dependence means that we can approximate X_i as a linear function of other K most correlated rows $X_{i_k}(k = 1, 2, ..., K$ and $i_k \neq i)$ of X_i. We will discuss the proper value of K in Section 4.3.2. The correlation between different rows is measured by the sum of *Pearson product-moment correlation coefficient* [32]. Then we perform multivariate linear regression to find a set of weights w_{i_k}, such that X_i can be best approximated by a linear function of X_{i_k}:

$$X_i \approx w_{i_0} X_0 + \sum_{k=1}^{K} w_{i_k} X_{i_k}, \tag{12}$$

where X_0 is a row vector (1×45) equals to $(1, 1, 1, ..., 1)$. Initially, every element of P is set to be 0. For the i-th row of P, we set $P(i,i) = 1$, $P(i,i_k) = -w_{i_k}$, for $k = 1, 2, ..., K$. We also set $P_0(i,j) = w_{i_0}$ for $j = 1, 2, ..., 45$. In this way,

$$P(LR^T) - P_0 = \begin{bmatrix} X_1 - \sum_{k=1}^{K} w_{1_k} X_{1_k} - w_{1_0} X_0 \\ X_2 - \sum_{k=1}^{K} w_{2_k} X_{2_k} - w_{2_0} X_0 \\ \cdots \\ \cdots \\ X_{67} - \sum_{k=1}^{K} w_{67_k} X_{67_k} - w_{67_0} X_0 \end{bmatrix},$$

and it describes the difference between each row and its linear representation. Since $X_i \approx w_{i_0} X_0 + \sum_{k=1}^{K} w_{i_k} X_{i_k}$, the value of $\|P(LR^T) - P_0\|_F^2$ is expected to be small. Therefore, the existence of $\|P(LR^T) - P_0\|_F^2$ in minimizing (11) guarantees the location dependence in X and this really improve the reconstruction accuracy.

Choice of $C_{(45\times45)}$ and $C_{0(67\times45)}$: Matrix C and C_0 represent the channel dependence in relative location-channel matrix (X), and express the linear relationship between different columns of $X_{(67\times45)}$. Actually, the way to find C and C_0 is almost the same as that of finding P and P_0, except that C and C_0 consider the columns of X instead of rows. In this way, the channel dependence in X is guaranteed by $\|(LR^T)C - C_0\|_F^2$ in (11).

4.3.2 Stability of location-channel dependence

In this part, we study the stability of the location-channel dependence based on the short-time indoor sensing results, and then find the proper value of K. In Section 2.3, a total of 14 short-time sensing data sets were collected and we can get 14 relative location-channel matrices (RLCM) $X^{(1)}$,

$X^{(2)}$, ..., $X^{(14)}$. We calculate P, P_0, C, and C_0 based on one of these 14 RLCMs. If the location dependence and channel dependence are stable over time, then P, P_0, C, and C_0 we get can approximately represent the location dependence and channel dependence of other 13 RLCMs.

We calculate P, P_0, C, and C_0 based on $X^{(1)}$ in the way described in Section 4.3.1. As we mentioned before, the value of $\|PX^{(1)} - P_0\|_F^2$ and $\|X^{(1)}C - C_0\|_F^2$ is expected to be small. If the value of $\|PX^{(i)} - P_0\|_F^2(\|X^{(i)}C - C_0\|_F^2)$ $(i = 2, 3, ..., 14)$ is close to $\|PX^{(1)} - P_0\|_F^2(\|X^{(1)}C - C_0\|_F^2)$, then the four matrices (P, P_0, C, C_0) can approximately represent the location-channel dependence in $X^{(i)}$ $(i = 2, 3, ..., 14)$. As a result, the location dependence and channel dependence are stable over time.

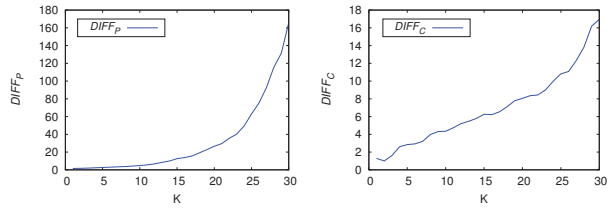
We compare $\|PX^{(1)} - P_0\|_F^2$ and the average of $\|PX^{(i)} - P_0\|_F^2$ $(i = 2, 3, ..., 14)$, and set

$$DIFF_P = \frac{[\frac{1}{13}\sum_{i=2}^{14}\|PX^{(i)} - P_0\|_F^2] - \|PX^{(1)} - P_0\|_F^2}{\|PX^{(1)} - P_0\|_F^2}.$$

Similarly, we set

$$DIFF_C = \frac{[\frac{1}{13}\sum_{i=2}^{14}\|X^{(i)}C - C_0\|_F^2] - \|X^{(1)}C - C_0\|_F^2}{\|X^{(1)}C - C_0\|_F^2}.$$

If the values of $DIFF_P$ and $DIFF_C$ are small, then the location dependence and channel dependence are stable. Figure 5 shows the value of $DIFF_P$ and $DIFF_C$, when K varies from 1 to 30. $DIFF_P$ gets its minimum value (1.54) when $K = 1$ and $DIFF_C$ gets its minimum value (1.28) when $K = 2$. It means that we would obtain very similar P, P_0, C, and C_0 using different $X^{(i)}$ as input, indicating that the location-channel dependence is consistent across different days. Since $DIFF_P$ and $DIFF_C$ get their minimum value when $K = 1$ and $K = 2$, respectively, we set $K = 1$ when calculating P and P_0 and $K = 2$ when calculating C and C_0.



(a) Location dependence (b) Channel dependence

Figure 5: Stability of location dependence and channel dependence

4.4 Data Reconstruction Algorithm

Data reconstruction part aims to find $\tilde{X} = \tilde{L}\tilde{R}^T$ that optimize (11). We design an algorithm to get $\tilde{L}$ and $\tilde{R}$ based on alternating least squares method. Before solving (11), we should scale P, P_0, C, and C_0 so that $\|P(LR^T) - P_0\|_F^2$, $\|(LR^T)C - C_0\|_F^2$ and $\|B_s \circ (LR^T) - D_s\|_F^2$ in (11) are of similar order of magnitude, otherwise they may overshadow each other during the optimization process [7].

As shown in Algorithm 1, L is randomly initialized. Then we fix L and find proper R to optimize (11) by finding its derivative with respect to R^T and making the derivative be equal to zero. Next we fix R and find the proper L in a similar way. We repeat this process until convergence. In Algorithm 1, we repeat for t times ($t = 100$ by default).

We now explain the alternating least squares method in detail. First, we assume that L is fixed and find proper R to minimize (11). It seems that we should set the derivative of (11) with respect to R^T to 0, but it's difficult to directly find this derivative. Thus, we try another way to solve this problem. We alternatively use (10) as the objective function and $\|P(LR^T) - P_0\|_F^2$, $\|(LR^T)C - C_0\|_F^2$ as constraints just like [8]. That means in every iteration, we find the L and R through making the derivative of (10) to be 0 (line 3-4 in Algorithm 1), then we compare the new value of (11) (v) to the current smallest value ($\tilde{v}$). If $v < \tilde{v}$, we update $\tilde{L}$ and $\tilde{R}$, otherwise, we do nothing (line 5-8 in Algorithm 1). The evaluation results in Section 5 show that Algorithm 1 really performs well.

Algorithm 1 Compressive sensing

Input:
D_s: DM with strong channels, B_s: BM with strong channels, λ: Lagrange multiplier, t: iteration times, r: rank bound, $\tilde{v}$: initially a sufficiently large number.
Output:
$\tilde{X}$: estimation matrix

1: $L \leftarrow random\_matrix(67, r)$
2: **for** 1 to t **do**
3: $R \leftarrow \text{findDer}(B_s, D_s, L, \lambda, r)$;
4: $L \leftarrow \text{findDer}(B_s^T, D_s^T, R, \lambda, r)$;
5: $v \leftarrow \|B_s \circ (LR^T) - D_s\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2)$ $+\|P(LR^T) - P_0\|_F^2 + \|(LR^T)C - C_0\|_F^2$
6: **if** $v < \tilde{v}$ **do**
7: $\tilde{L} \leftarrow L; \tilde{R} \leftarrow R; \tilde{v} \leftarrow v$
8: **end if**
9: **end for**
10: $\tilde{X} \leftarrow \tilde{L}\tilde{R}^T$
11: return $\tilde{X}$;

Algorithm 2 describes the detailed derivation process. Let

$$f(L, R^T) = \|B_s \circ (LR^T) - D_s\|_F^2 + \lambda(\|L\|_F^2 + \|R^T\|_F^2),$$

and L be fixed. The value of R can be obtained by solving: $\frac{\partial f}{\partial R^T} = \mathbf{0}$. Let $\Theta = R^T$, we have

$$\begin{aligned} &\frac{\partial f(L,\Theta)}{\partial \Theta} \\ =\ &\frac{\partial(\|B_s \circ (L\Theta) - D_s\|_F^2 + \lambda(\|L\|_F^2 + \|\Theta\|_F^2))}{\partial \Theta} \\ =\ &\frac{\partial(\sum_{j=1}^{45}\|Diag(\boldsymbol{b_j})L\boldsymbol{\theta_j} - \boldsymbol{d_j}\|_F^2 + \lambda\|\boldsymbol{\theta_j}\|_F^2)}{\partial \Theta} \end{aligned} \tag{13}$$

where $\boldsymbol{b_j}$, $\boldsymbol{d_j}$, and $\boldsymbol{\theta_j}$ represent the jth column of B_s, D_s and Θ, respectively, and $Diag(\boldsymbol{b_j})$ refers to a diagonal matrix with elements of $\boldsymbol{b_j}$ on its main diagonal. $\frac{\partial f(L,\Theta)}{\partial \Theta} = \mathbf{0}$ is equal to that for every j from 1 to 45

$$\frac{\partial f(L, \Theta)}{\partial \theta_j} = \mathbf{0},$$

which means

$$\frac{\partial(\|Diag(\boldsymbol{b_j})L\boldsymbol{\theta_j} - \boldsymbol{d_j}\|_F^2 + \lambda\|\boldsymbol{\theta_j}\|_F^2)}{\partial \boldsymbol{\theta_j}} = \mathbf{0}. \tag{14}$$

Equation (14) is a linear least squares problem and we have

$$\boldsymbol{\theta_j} = (Q_j^T Q_j + \lambda I_r)^{-1}(Q_j^T \boldsymbol{d_j}),$$

where $Q_j = Diag(\boldsymbol{b_j})L$. At last, the value of R can be calculated by $R = \Theta^T$. When R is fixed, L can be found in a similar way.

4.5 Sensor Deployment

As mentioned before, different deployments of sensors lead to different performance of FIWEX. Although brute-force

Algorithm 2 Derivation

Θ^T=**findDer**(B_s,D_s,L,λ,r)
1: $[m, n]$ =size(B_s);
2: **for** $j = 1$ to n **do**;
3: $Q_j \leftarrow Diag(\boldsymbol{b_j})L$;
4: $\boldsymbol{\theta_j} = (Q_j^T Q_j + \lambda I_r)^{-1}(Q_j^T \boldsymbol{d_j})$;
5: **end for**
6: **return** Θ^T;

enumeration of indoor locations can find the best deployment solution, the computation overhead is unacceptable ($C_{67}^N - 1$ comparisons are needed given N sensors). In this section, we propose a novel sensor deployment method based on location dependence and the clustering technique.

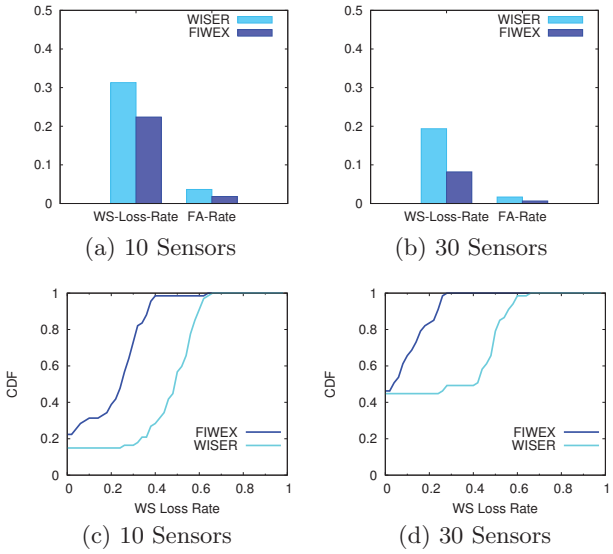
As we mentioned in Section 2, there exists linear dependences between different locations. In section 4.3, we utilize the location dependence to improve the performance of FIWEX by representing the channels strength at one location in form of the linear combination of those at other K most correlated locations. Given N sensors, FIWEX can collect the channels strength of N locations, and then reconstruct channels strength of all locations based on them. Furthermore, if the channels strength in every location can be approximated by the combination of these N locations, the reconstruction accuracy would be high. Hence, deploying sensors at "independent" locations is a good choice as every one of them can represents a set of correlated locations.

Clustering is a useful method to deploy sensors at "independent" locations. Our sensor deployment algorithm can be briefly divided into two steps: (i) Cluster all locations into N groups (N is the number of sensors here). (ii) Deploy one sensor in each center of these N groups. In particular, we choose *k-mediod* [13] technology to do the clustering and use the sum of Pearson product-moment correlation coefficients of one row as the similarity metric. In Algorithm 3 line 2, kmedoids(X, N) divides the rows of X into N clusters, the centers of which are stored in SL. v stores the sum of all distances between every row and the center of cluster it belongs to, and a smaller v means a better clustering result. Considering the truth that *k-medoids* clustering algorithm may find the local optimization, we run kmedoids for 100 times and select the best result. Then we deploy N sensors in the N cluster centers.

At last, we consider the stability of the clustering results. In Section 4.3.2, we show that the location dependence is stable over time, as a result, the clustering algorithm based on location dependence is also stable. That means, if we deploy sensors according to different relative location-channel matrices, the results would be similar.

Algorithm 3 Sensor deployment

Input:
X: relative location-channel matrix
N: the number of given sensors
v_{opt}: initially a sufficiently large number.
Output:
SL_{opt}: a list of sensor locations
1: **for** 1 to 100 **do**
2: $[SL, v]$ = kmedoids(X, N);
3: **if** $v < v_{opt}$ **do**
4: $SL_{opt} = SL; v_{opt} = v$;
5: **end if**
6: **end for**
7: return SL_{opt};



(a) 10 Sensors (b) 30 Sensors
(c) 10 Sensors (d) 30 Sensors

Figure 6: Average FA rate, WS Loss Rate and CDF curves.

5. PERFORMANCE EVALUATION

In this section, we perform experiments to evaluate the performance of FIWEX. To the best of our knowledge, WISER is the most efficient indoor white space exploration system till now, and we examine the cost-efficiency of FIWEX by comparing it to WISER. First, we evaluate the performance of FIWEX and WISER in the scenarios of 10 sensors and 30 sensors. Then we compare the FA Rates and WS Loss Rates of FIWEX and WISER in different number of sensors, and we calculate the ratio of extra white spaces identified by FIWEX. The evaluation results show that FIWEX is better than WISER.

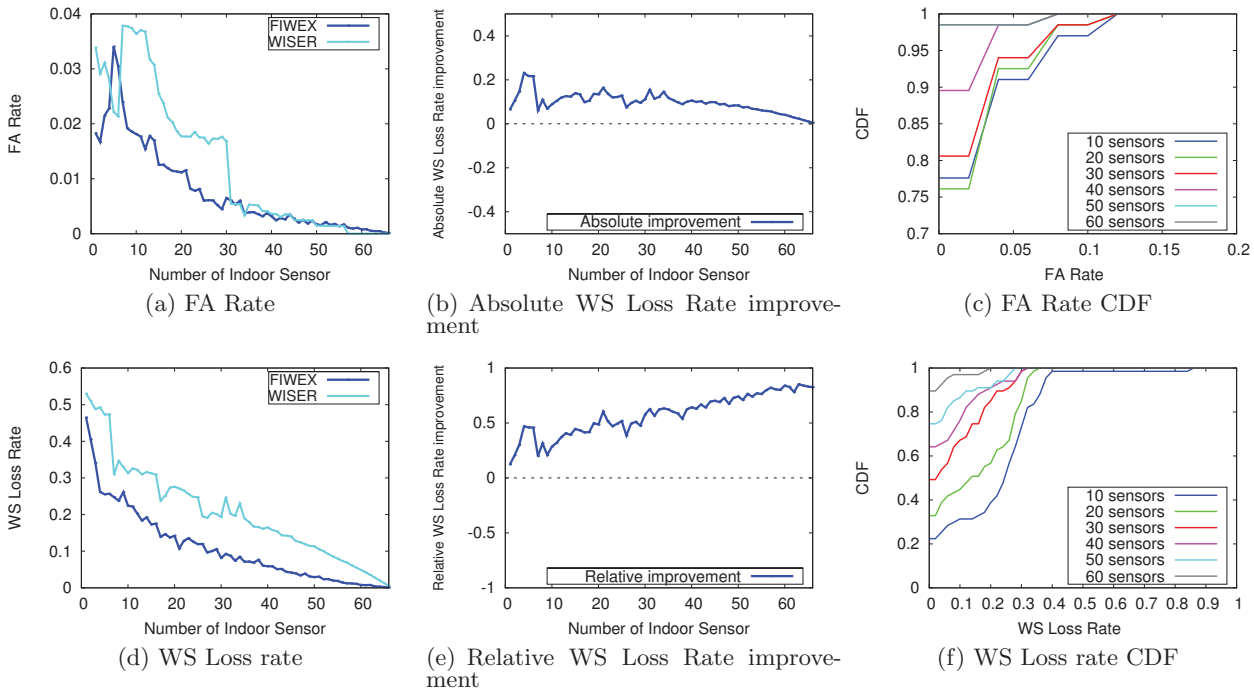
5.1 Methodology

To evaluate FIWEX's performance, we build a prototype of it on the third floor of a building. We set the white space threshold to be -84.5dBm/8 MHz as WISER does, and a total of 14 short-time sensing data sets were collected. We divide these 14 data sets into two parts: one part of 7 data sets for training FIWEX, the other part for evaluations. From the training set, we get 7 relative location-channel matrices, and we use the average of these 7 relative location-channel matrices to find P, P_0, C, C_0 and to deploy sensors. Then we evaluate the performance of FIWEX based on the other 7 datasets. We choose FA Rate and WS Loss Rate as the metrics of system's performance. Their definitions are as follows.

- **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 identified by the system.
- **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.

Since we use 7 datasets for evaluation, we get 7 sets of FA Rates and WS Loss Rate, then we use their average in the following part of evaluation.

5.2 Performance Comparison With WISER



(a) FA Rate (b) Absolute WS Loss Rate improvement (c) FA Rate CDF

(d) WS Loss rate (e) Relative WS Loss Rate improvement (f) WS Loss rate CDF

Figure 7: Evaluation results on the number of indoor sensors

We first run FIWEX and WISER using two different numbers of sensors: 10 and 30, and compare their FA Rates and WS Loss Rates. As shown in Figure 6(a), when there are 10 sensors, the FA Rate and WS Loss rate of FIWEX are 1.8% and 22.3% while WISER performs an FA rate 3.6% and a WS Loss Rate 31.3%. The result shows that FIWEX has lower or better FA Rate and WS Loss Rate compared with WISER when 10 sensors are used. We also depict the CDF (Cumulative Distribution Function) curves of WS Loss Rate after calculating the WS Loss Rate for each location. Figure 6(c) shows the CDF curves of WS Loss Rate of FIWEX and WISER with 10 sensors. In WISER, 83.6% locations suffer a WS Loss Rate that is higher than 30%. The number is 26.9% in FIWEX. In WISER, 71.6% locations suffer a WS Loss Rate which is higher than 40% while the number is only 1.5% in FIWEX. Figure 6(b) shows the FA Rates and WS Loss Rates in the scenario of 30 sensors and the corresponding CDF curves are shown in Figure 6(d). These two figures shows that FIWEX also outperforms WISER in 30 sensors scenario.

5.3 Performance on The Number of Indoor Sensors

In a general indoor scenario, there exists a tradeoff between system performance and the number of indoor sensors that are deployed. To understand this tradeoff, we vary the number of sensors from 1 to 66 (when there are 67 sensors, FA Rate and WS Loss Rate of FIWEX and WISER are both 0), and evaluate the performance of FIWEX and WISER. In this experiment, we define the following four metrics to describe the FIWEX's performance improvement compared to WISER.

- **Absolute FA Rate improvement:** the difference between FA Rates of WISER and FIWEX.
- **Relative FA Rate improvement:** the ratio between absolute FA Rate improvement and the corresponding FA rate of WISER. (If the FA Rate of WISER is 0 while the absolute FA Rate improvement is not 0, we set the relative FA Rate improvement to be -1.)
- **Absolute WS Loss Rate improvement:** the difference between WS Loss Rates of WISER and FIWEX.
- **Relative WS Loss Rate improvement:** the ratio between absolute WS Loss Rate improvement and the corresponding WS Loss rate of WISER. (If the WS Loss Rate of WISER is 0 while the absolute WS Loss Rate improvement is not 0, we set the relative WS Loss Rate improvement to be -1.)

In Figure 7(a), we compare the FA Rates. It is clear that the FA Rates of FIWEX are lower than those of WISER in most cases. Actually, the average FA rate (from 1 sensor to 66 sensors) of WISER is 1.25% while this number is 0.77% for FIWEX. FIWEX has an average absolute FA Rate improvement of 0.48% and an average relative FA Rate improvement of 38.4% compared with WISER.

Figure 7(d) shows the WS Loss Rate of FIWEX and WISER with different number of indoor sensors. The average WS Loss Rate (from 1 sensor to 66 sensors) of WISER is 20.3% and for FIWEX, this number is 10.6%. FIWEX has an average absolute WS Loss Rate improvement of 9.7% and an average relative WS Loss Rate improvement of 47.8% compared to WISER. Figure 7(b) and Figure 7(e) show the absolute and relative WS Loss Rate improvement versus the number of indoor sensors, respectively, and they demonstrate the amount of extra white spaces FIWEX can identify compared with WISER.

We depict the CDF curves of FIWEX regarding to the FA Rates and WS Loss Rates of all locations. Figure 7(c) shows the 6 CDF curves (10,20,...,60 sensors) about FA rate, and considering the attributes of FA rate (the value of FA rate is small), we change the range of axes to [0:0.2] and [0.7:1]. The CDF curve is "higher" when more sensors are used, and this means that the number of locations with high FA Rates

decreases as the increasing number of sensors. In addition, 50 sensors and 60 sensors share the same CDF curve, so we can only find 5 CDF curves in Figure 7(c). The CDF curves of WS Loss rate of different number of sensors are shown in Figure 7(f).

6. RELATED WORK

Most of the existing works on TV white space focus on outdoor scenarios. For example, in [24], the authors design a Wi-Fi like AP system constructed on top of UHF white spaces and a dynamic spectrum allocation network system is proposed in [25]. Spectrum sensing and geo-location database are two main approaches for white space exploration, and have been widely studied these years [26, 27, 29].

The outdoor white space exploration methods cannot be directly applied to indoor scenarios because of the complicated indoor environment. In [2], the authors propose the first indoor white space identification system, called WISER. WISER utilized the channel-location clustering based algorithm to explore indoor white spaces, and obtained great improvement compared to the baseline approaches. However, the property that different locations have different strong channels and the linear dependence of locations and channels are not considered by the authors. Based on these observations, we present FIWEX, a compressive sensing based cost-efficient indoor white space exploration mechanism.

These years, compressive sensing theory has been widely studied and utilized in a lot of fields. For instance, [6] proposed the robust network compressive sensing and proved its efficiency based on a large amount of evaluation, and compressive sensing technology has been utilized for network traffic estimation [7], localization in mobile networks [9], soil moisture sensing [10] and data gathering [11, 12]. In FIWEX, we combine compressive sensing with indoor white space exploration in an innovative way and efficiently explore indoor white spaces with a high accuracy.

7. CONCLUSIONS

In this paper, we performed indoor white space measurements in a real building to study the characteristics of indoor white spaces. The measurement results confirmed the existence of strong channels and location-channel dependence. Motivated by these observations, we proposed a cost-efficient indoor white space exploration mechanism, called FIWEX. Given the same number of sensors, FIWEX can identify more indoor white spaces with an average less false alarms compared to the existing indoor white space exploration systems.

References

[1] Second Memorandum Option And Order. https://apps.fcc.gov/edocs_public/attachmatch/FCC-10-174A1.pdf

[2] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra, "Exploring indoor white spaces in metropolises," In *Proc. of ACM MOBICOM*, 2013.

[3] D. Donoho, "Compressed sensing," *IEEE Trans. on Information Theory*, 52(4): 1289-1306, 2006.

[4] B. Recht, M. Fazel, and P. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization," *SIAM Review*, 52(3): 1289-1306, 2007.

[5] E. Candes, J. Romberg, and T. Tao, "Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information," *IEEE Trans. on Information Theory*, 52(2): 489-509, 2006.

[6] Y. Chen, L. Qiu, Y. Zhang, G. Xue and Z. Hu, "Robust Network Compressive Sensing," In *Proc. of ACM MOBICOM*, 2014.

[7] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, "Spatio-Temporal Compressive Sensing and Internet Traffic Matrices," In *Proc. of ACM SIGCOMM*, 2009.

[8] L. Kong, M. Xia, X. Liu, M. Wu and X. Liu, "Data Loss and Reconstruction in Sensor Networks," In *Proc. of IEEE INFOCOM*, 2013.

[9] S. Rallapalli, L. Qiu, Y. Zhang, and Y. Chen, "Exploiting Temporal Stability and Low-Rank Structure for Localization in Mobile Networks," In *Proc. of ACM MOBICOM*, 2010.

[10] X. Wu and M. Liu, "In-situ Soil Moisture Sensing: Measurement Scheduling and Estimation using Compressive Sensing," In *Proc. of ACM IPSN* 2012.

[11] H. Zheng, S. Xiao, X. Wang and X. Tian, "Energy and Latency Analysis for In-network Computation with Compressive Sensing in Wireless Sensor Networks," In *Proc. of IEEE INFOCOM*, 2012.

[12] C. Luo, F. Wu, J. Sun, and C. Chen, "Compressive data gathering for large-scale wireless sensor networks," In *Proc. of ACM MOBICOM* 2009.

[13] H.S. Park and C.H. Jun, "A simple and fast algorithm for K-medoids clustering," *Expert Systems with Applications*, 36(2): 3336-3341, 2009.

[14] M. McHenry, P. Tenhula and D. McCloskey, "Chicago Spectrum Occupancy Measurements & Analysis and A Long-term Studies Proposal" In *Proc. of ACM TAPAS*, 2006.

[15] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner and J. Grosspietsch, "Geo-location Database Techniques for Incumbent Protection in the TV White Space", In *Proc. of IEEE DySPAN*, 2008.

[16] N. Klepeis, W. Nelson, W. Ott, J.Robinson, A. Tsang, P. Switzer, J. Behar, S. Hern and W. Engelmann, "The national human activity pattern survey (nhaps): A resource for assessing exposure to environmental pollutants", *Journal of Exposure Analysis and Environmental Epidemiology*, 11(3): 231-252, 2001.

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

[18] Yucek, Tevfik, and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys & Tutorials, IEEE*, 11(1): 116-130, 2009.

[19] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, 55(4): 523-531, 1967.

[20] M. Vossiek, L. Wiebking, P. Gulden, J. Wiehardt, C. Hoffmann, and P. Heide, "Wireless Local Positioning", *IEEE Microwave Magazine*, 4(4): 77-86 2003.

[21] S. Kumar, S. Gil, D. Katabi and D. Rus, "Accurate Indoor Localization With Zero Start-up Cost," In *Proc. of ACM MOBICOM*, 2014.

[22] E. Candes and B.Recht, "Exact Matrix Completion via Convex Optimization," *Foundations of Computational mathematics* 9(6): 717-772, 2009.

[23] B. Recht, W. Xu, and B. Hassibi, "Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization," In *Proc. of IEEE CDC*, 2008.

[24] 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.

[25] S. Deb, V. Srinivasan, and R. Maheshwari, "Dynamic spectrum access in DTV whitespaces: design rules, architecture and algorithms," In *Proc. of ACM MOBICOM*, 2009.

[26] 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.

[27] T. Zhang, N. Leng and S. Banerjee, "A Vehicle-based Measurement Framework for Enhancing Whitespace Spectrum Databases," In *Proc. of ACM MOBICOM*, 2014.

[28] R. Ahuja, R. Corke, and A. Bok, "Cognitive radio system using IEEE 802.11a over UHF TVWS," In *Proc. of IEEE DySPAN*, 2008.

[29] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner and J. Grosspieetsch, "Geo-location database techniques for incumbent protection in the TV white space," In *Proc. of IEEE DySPAN*, 2008.

[30] Ian F. Akyildiz, W. Lee, et al, "NeXt generation/ dynamic spectrum access/cognitive radio wireless networks: a survey," *Computer Networks* 50(13): 2127-2159, 2006.

[31] Y. Xing, R. chandramouli, S. Mangold and S. Shankar, "Dynamic spectrum access in open spectrum wireless networks," *Selected Areas in Communications, IEEE Journal* 24(3): 626-637, 2006.

[32] J. Rodgers, Joseph, and W. Alan Nicewander, "Thirteen ways to look at the correlation coefficient," *The American Statistician* 42(1): 59-66, 1988.