MISEN: A Mobile Indoor White Space Exploration System

***Abstract*—With the exponential growth of wireless communication demand, spectrums have become a shortage of resources. Making use of white spaces is a promising way to tackle this problem. In this paper, we propose a mobile platform based indoor white space exploration method to profile a building's internal white space distribution. We design a Mobile Indoor Spectrum ExploratioN system, named MISEN, where we creatively consider the white space correlation in time dimension. In contrast, existing works only take location and channel correlations into consideration. We leverage a tensor completion based algorithm to recover our collected data, and utilize Alternating Direction Method of Multipliers method to solve the optimization problem. Moreover, we build a prototype of MISEN and further evaluate its performance in our verification experiment. The results illustrate that MISEN yields a superior performance over existing methods, where MISEN can detect 18.7% more white spaces with 20.0% less false rate than the best known existing solution.**

I. INTRODUCTION

With the exponential growth of demand for wireless data communications, frequency spectrums have become an increasingly scarce resources. However, the utilization of spectrums is low at many places [1]. In order to tackle this problem, researchers proposed the concept of 'Dynamic Spectrum Access', which allows unlicensed users to utilize specific spectrums without interfering licensed devices. In 2008, the Federal Communications Commission (FCC) issued a rule that allows the unlicensed use of unoccupied TV spectrums, which is also called TV White Spaces or simply White Spaces. After that, topics of white spaces became more and more popular.

Many of the existing works [4, 6, 17, 23, 30, 33] focus on outdoor white spaces. However, nearly 70% of people's spectrum demands are derived from the indoor scenarios [11, 20]. Therefore, exploration of white spaces in indoor environment is more important. In order to avoid interfering with those licensed users and devices when using white spaces, unlicensed devices need to identify which spectrums are vacant before getting access to them. Hence, our objective is to provide a method, which can profile a building's internal white space distribution. Prior literatures have proposed some excellent indoor white space exploration systems, including WISER [31] and FIWEX [21]. Both of them have been proved to get a good performance. However, there is still much room for improvement. For one thing, nearly all of the existing indoor white space exploration methods only consider location and channel correlations of white space information, and then perform data completion only basing on real-time data. While there are still white space correlations in time dimension, we can leverage both real-time data and historical data to make full use of the information we collected. For another, though FIWEX has greatly reduced the number of required detectors, the cost of these detectors is still somewhat unacceptable since RF detectors are quite expensive.

Considering above limitations, we propose a real-time spectrum detecting system, namely MISEN (**M**obile **I**ndoor **S**pectrum **E**xploratio**N**). We apply mobile detectors to detect white spaces in the whole building. By using only a few mobile detectors, we can collect data from all the locations because of their mobility. As a consequence, the number of required detectors can be greatly reduced.

However, mobile indoor spectrum exploration is not easy to implement. When deploying a small amount of mobile detectors, for example we deploy 2 detectors, we can only obtain two measurements at the same time. However, it is impossible to recover the white space distribution in the whole building from these two measurements. Previous works have proved that white space information is stable in a short period of time [21]. Motivated by this premise, we introduce a concept of time slice, where in the same time slice we think the white space information is constant. But we only have location and channel dimensions within this time slice. In order to take time dimension into consideration, we make use of the latest certain amount of time slices to perform data completion. In this way, we utilize both historical white space information and real-time white space information, and then we can further recover the undersampled data by fully exploiting white space correlations in all space, frequency and time dimensions. At the same time, when taking time dimension into consideration, the data form in our system is a three-dimension tensor. However, it is not easy to perform tensor completion, because the rank of a tensor is a quite complex concept. Here, we leverage a tensor-nuclear-norm based tensor completion method to recover our data [35]. In addition, it is also a tough problem to make a proper choice of time slices' length and the total time span of a tensor. We discuss the influence of these two factors in section V.

Here are the main contributions of this paper.

- We proposed MISEN, a Mobile Indoor White Space ExploratioN system with low cost and high efficiency. Also, we constructed a prototype of the mobile detector by mounting a spectrum detector and a localization tag on a sweeping robot. This mobile detector can move automatically while detecting white spaces.
- We designed a tensor completion based algorithm to explore indoor white spaces. In order to exploit both historical measurements and real-time measurements, we creatively take the time dimension white space correla-

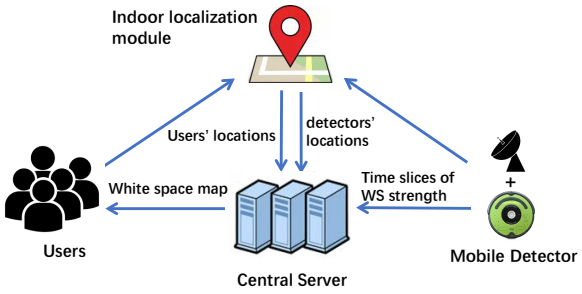


Fig. 1. System model schema of MISEN

tions into consideration. In this way, MISEN can efficiently exploit collected indoor white space information and precisely recover the undersampled data.

- We perform a proof-of-concept experiment of MISEN to evaluate its performance. We compared MISEN's performance to FIWEX and WISER using the same data set, where the results indicate that MISEN does a better job. On average, MISEN can identify 18.7% more indoor white spaces with 20.0% less false alarms than FIWEX, when they have the same information completeness rate. Moreover, MISEN can identify 46.1% more indoor white spaces with 64.9% less false alarms than WISER with the same information completeness rate.

The remainder of this paper is arranged as follows. We describe how our system model MISEN works in Section II. In Section III we present the data reconstruction algorithm and optimization algorithm of MISEN. Section IV introduces the indoor white space exploration experiment in detail. Section V shows the performance evaluations of MISEN. Then, the related works are presented in Section VI. At last, we conclude our work in Section VII.

II. SYSTEM MODEL

In this section, we will introduce the system structure of MISEN, which aims at profiling a building's internal white space distribution precisely with only a few mobile detectors. When deploying mobile detectors, in addition to the white space information, we also need to know the routes of mobile detectors. In our system, we adopt a 'random moving and locating' method to follow the track of mobile detectors. As shown in Fig.1, MISEN mainly consists of three parts: central host, indoor localization module, and mobile detecting module.

In the mobile detecting submodule, our mobile detector can collect white space information while sweeping the room, and then transmit the information to the central host at regular interval. At the same time, the indoor localization submodule locates our mobile detectors in real-time and sends their locations to the central host. The central host matches the locations with the information collected by the mobile detectors, and then constructs the white space availability map basing on them.

When users need the white space availability at their locations, they send their positions to the central host. The central host, as a response, transmits the corresponding white space availability to the users.

A. Indoor Localization Module

In MISEN, indoor locating module is used to acquire users' locations and record detectors' tracks, then sends detectors' location list L_1 to the central host. Indoor localization has been developed to a mature technology in recent years [14, 28]. In order to accurately locate our mobile detector in real time, MISEN takes a commercial localization product, which utilizes Ultra-wideband (UWB) signals to perform localization. Note that the operating frequency of UWB localization submodule is 3.8GHz - 5.8GHz, which means these radio signals do not interfere with the frequency spectrums of white space.

B. Mobile detecting Module

Mobile detecting module is applied to perform a 'partial' indoor white space detecting with several mobile detectors, we call it partial since only a few locations are detected at the same time. As what we mentioned in section I, nearly all the pre-existing white space detecting system deploy static detectors. However, the methods with static detectors sometimes may not be efficient enough, because white space information has redundancies in time dimension [21]. Hence, we do not need to fix detectors at some selected locations all the time. A natural consideration is that deploying mobile detectors is more efficient. For one thing, they are able to collect data from different locations and different channels with time passing by. For another, they can largely reduce the number of detectors required which greatly saves cost. In recent years, sweeping robots have become an important part of people's daily lives. So, it would be a convenient way to mount the spectrum detectors on the sweeping robots, and then collect the indoor white space information in a mobile way.

Fig.4 shows the picture of our mobile detector prototype. We attached a detector and a localization tag on the sweeping robot. When this submodule is working, the sweeping robot takes the detector to traverse the indoor environment based on its sweeping route plan algorithm. At the same time, the detector detects frequency spectrums at the same time, and then periodically sends its spectrum information list L_2 to the central host. Additionally, the localization module locates these mobile detectors' tracks through the localization tags and transmits them to the central host.

It is conceivable that the spectrum information collected from different locations is not synchronous because the detector is moving. However, according to previous work, indoor white spaces are stable in a short period of time [21]. So it is reasonable to assume that those frequency spectrums of concern are constant in a short period of time, such as 5 minutes. And it has been experimentally proved that the mobile detecting module could do a good job.

C. Central Host

The location list L_1 and spectrum information list L_2 are transmitted to the central host periodically. After that, our

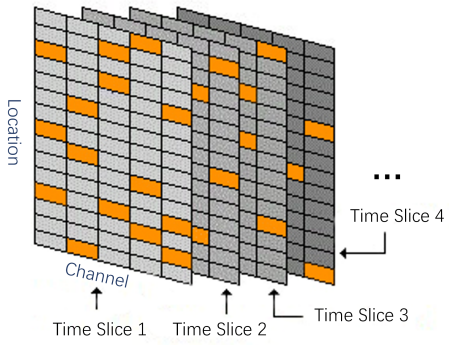


Fig. 2. Incomplete tensor

central host parses these lists into an incomplete location-channel matrix at the same time slice, which records channels' signal strengths, and stores this matrix in memory. Here we introduce a concept of time slice, which is defined as follows:

Definition 1: **Time slice.** A time slice here is the minimum resolution in time dimension in our scenario, where in the same time slice we think white space information is constant. In the following part of this paper, we also use 'time slice' to represent the location-channel signal strength matrix at this time slice.

Central host joints a series of time slices to an incomplete location-channel-time tensor to make use of historical white space information, and then reconstructs it into a complete tenser. Fig.2 shows the schematic diagram of an incomplete tensor, where the orange blocks represent the detected entries. Central host takes the first frontal time slice as the real-time white space availability map. When users start a query, central host returns a list of corresponding white space list. In summary, the central host mainly performs two functions: processing data and serving as a white space availability database.

The data processing mainly contains two aspects: raw data parsing and spectrum data recovering. To facilitate the explanation of this process, we will describe the data processing part by taking our experiment as an example. In our experiment, we select 40 profiled locations in the target building to discretize the plane. The channels of concern are Chinese digital TV bands of 470MHz-806MHz, where the bandwidth of a channel is 8MHz. So, there are 42 TV channels being considered. Moreover, in time dimension we utilize the latest 50 minutes' data to form a tensor. Here, we consider one time slice as 5 minutes, and hence we have 10 time slices in a tensor. Thus, we have the following definitions.

Definition 2: **Ground truth tensor** (GT) is a $40 \times 42 \times 10$ tensor, which shows the ground truth information of signal strengths. GT is denoted by $\mathcal{X}$, where $\mathcal{X}(i,j,k)$ denotes the signal strengths of channel j at location i and time slice k.

Definition 3: **Loss pattern** (LP)is a $40 \times 42 \times 10$ tensor, indicating which entries of the tensor have been detected by the mobile detecting module. We denote LP as Ω. For channel j, location i and time slice k:

$$\Omega(i,j,k) = \begin{cases} 1 & \text{if this entry is observed} \\ 0 & \text{if this entry is not observed} \end{cases} \quad (1)$$

Definition 4: **Measurement tensor** (MT) is a $40 \times 42 \times 10$ tensor, which indicates the signal strengths at those observed channels in each location and time slice, and at those unobserved channels, locations and time slice the entry equals to 0. We denote MT as $\mathcal{M}$:

$$\mathcal{M}(i,j,k) = \begin{cases} \mathcal{X}(i,j,k) & \text{if this entry is observed} \\ 0 & \text{if this entry is not observed} \end{cases} \quad (2)$$

Hence, it is conceivable that $\mathcal{M} = \Omega \circ \mathcal{X}$, where $\circ$ refers to Hadamard product. ($\mathcal{M} = \Omega \circ \mathcal{X}$ means $\mathcal{M}(i,j,k) = \Omega(i,j,k) \times \mathcal{X}(i,j,k)$)

Definition 5: **Reconstructed tensor** (RT) is a $40 \times 42 \times 10$ tensor, which is generated by reconstructing the measurement matrix. And we denote it as $\mathcal{A}$.

Thus, raw data parsing refers to the process that central host receives raw data list L_1, L_2 periodically and parses them into tensors Ω and $\mathcal{M}$. Then spectrum data recovering means utilizing tensor completion method to obtain the reconstructed tensor $\mathcal{A}$ based on Ω and $\mathcal{M}$.

As for the white space database function, its data are obtained from the latest frontal slice $\mathcal{A}(:,:,1)$ of tensor $\mathcal{A}$. Central host calculates the real-time indoor white space availability by comparing the entries of $\mathcal{A}(:,:,1)$ to a presupposed threshold. In order to protect the licensed user from being interfered, we add a protection range PR here. We define the real-time **W**hite **S**pace **A**vailability **M**ap as

$$MAP(i,j) = \begin{cases} 1 & \text{if } \mathcal{A}(i,j,1) \geq threshold - PR \\ 0 & \text{if } \mathcal{A}(i,j,1) < threshold - PR \end{cases} \quad (3)$$

In this matrix, $MAP(i,j) = 1$ means location i channel j is occupied while $MAP(i,j) = 0$ means location i channel j is vacant. Users submit their indoor positions to the central host. Upon receiving these users' positions, the central host first finds the nearest profiled location, then outputs a list of vacant channels at this location according to MAP.

III. DATA RECONSTRUCTION ALGORITHM

In this section, we will present the tensor completion based indoor white space reconstruction algorithm. Data completion from partial measurements has been an important part of research. Specially, tensor completion, as a general form of data completion, has been widely studied [15, 19], and utilized to different realms, such as computer vision [22, 35].

In our scenario, we consider the time correlation in the data measured by the mobile spectrum detectors. As a result, our data form is a three-dimension tensor, including space, frequency and time. Furthermore, we cannot deploy so many detectors to detect an entire frequency spectrum information due to the cost limitation. Hence, we need to perform data reconstruction from partial measurements.

TABLE I
NOTATIONS USED IN THIS PAPER

Variables	Meanings
L_1	Time-location list(TLL)
L_2	Time-spectrum list(TSL)
$\mathcal{X}$	Complete ground truth tensor (GT)
$\hat{\mathcal{X}}$	the Flourier transform along the third dimension of $\mathcal{X}$
$\mathcal{A}$	Reconstructed tensor (RT)
Ω	Observation tensor (OT)
$\mathcal{M}$	Measurement tensor (MT)
WSM	White space availability map

A. Notation and Definition

In this paper, we leverage a tensor-nuclear-norm [35] based tensor completion method to recover the white space three-dimension tensors. That is because this method has been proved to be more efficient for multilinear data completion. White space data exactly has strong linear dependence on time-varying fibers [21]. The time-varying fiber here is a vector defined by fixing location and channel indices.

We will introduce some important definitions on our tensor completion method and describe the notations used in this section. Table.I illustrates the notations we use in this paper. Here we use 'Euclid Math One' font like $\mathcal{A}$ to represent tensors and use normal letters to represent matrixes and vectors. For a tensor $\mathcal{A}$, we use $\mathcal{A}(:,:,k)$ or $\mathcal{A}^{(k)}$ to represent its k_{th} frontal slice, and use $\mathcal{A}(i,j,:)$ to denote the $(i,j)_{th}$ time-varying fiber. In addition, we define $blkdiag(\mathcal{A})$ as follows:

$$blkdiag(\mathcal{A}) = \begin{pmatrix} \mathcal{A}^{(1)} & & & \\ & \mathcal{A}^{(2)} & & \\ & & \ddots & \\ & & & \mathcal{A}^{(n_3)} \end{pmatrix} \tag{4}$$

The tensor complexity can be denoted by tensor's rank. Here we use the tensor multi-rank [19] to represent our tensor complexity. It is defined as follows:

Definition 6: **Tensor multi-rank.** The multi-rank of a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ means a column vector with the length of n_3. The i_{th} element of this vector is the rank of the i_{th} frontal slice of $\hat{\mathcal{A}}$, where $\hat{\mathcal{A}}$ can be obtained by computing the Fourier transformation along each time-varying fibers of $\mathcal{A}$.

We need to perform adequate convex relaxations when applying the tensor multi-rank in practice. Hence, for tensor multi-rank we have the following definition:

Definition 7: **Tensor-nuclear-norm.** The tensor-nuclear-norm (TNN) of tensor $\mathcal{A}$ is the sum of all singular values of the frontal slices of $\hat{\mathcal{A}}$. And we denote the tensor-nuclear-norm as $\|\mathcal{A}\|_{TNN}$.

It has been proved that tensor-nuclear-norm is the tightest convex relaxation to the ℓ_1 norm of the tensor multi-rank [35]. In the following parts, we will present the detailed realization of the tensor completion algorithm.

B. Tensor Completion Details

In our indoor white space exploration scenarios, we aim to get an approximation of $\mathcal{X}$ given the measurements $\mathcal{M}$. Previous works have proved that signal strengths of white space have plenty of inherent location correlations and channel correlations [21]. Moreover, the white space location-channel matrix, which records signal strengths, is sparse with a quite low rank. Also, it has been proved that white space information is stable in a short period of time [21]. Basing on these premises, it is conceivable that our white space strength tensor is a sparse tensor with a low rank. So, we start our tensor completion from the minimizing tensor complexity problem. As we have defined in last subsection, the tensor multi-rank is a measure of tensor complexity. Furthermore, we know that tensor-nuclear-norm $\|\cdot\|_{TNN}$ is the tightest convex relaxation to ℓ_1 norm of the tensor multi-rank [35]. So we use tensor-nuclear-norm $\|\cdot\|_{TNN}$ to approximate the complexity of our tensor, and then get the following convex optimization problem:

$$\begin{aligned} &\text{Minimize} \quad \|\mathcal{A}\|_{TNN}, \\ &\text{Subject to} \quad P_\Omega(\mathcal{A}) = P_\Omega(\mathcal{X}). \end{aligned} \tag{5}$$

where $\mathcal{A}$ represents our recovered tensor and $\mathcal{A} \in \mathbb{R}^{40\times 42\times 10}$ in our experiments. Moreover, P_Ω is an operator, which vanishes those corresponding entries who have an 0 value in Ω. So the $(i,j,k)_{th}$ component of $P_\Omega(\mathcal{A})$ is equal to $\mathcal{X}_{ijk}$ if $\Omega_{ijk} = 1$ and zero otherwise.

As we have said in last subsection: $\mathcal{M} = P_\Omega(X)$. We can define $\mathcal{G} = \mathcal{F}_3 P_\Omega \mathcal{F}_3^{-1}$, where $\mathcal{F}_3$ and $\mathcal{F}_3^{-1}$ are Fourier and inverse Fourier transform operators, which represent to perform Fourier and inverse Fourier transform along the third dimension of our tensors. So, we can get $\hat{\mathcal{M}} = \mathcal{G}(\hat{\mathcal{X}})$, where $\hat{\mathcal{M}}$ and $\hat{\mathcal{X}}$ represent $\mathcal{F}_3(\mathcal{M})$ and $\mathcal{F}_3(\mathcal{X})$ respectively. Then according to the definition of tensor-nuclear-norm, we can transform 5 to the following problem:

$$\begin{aligned} &\text{Minimize} \quad \|blkdiag(\hat{\mathcal{A}})\|_*, \\ &\text{Subject to} \quad \hat{\mathcal{M}} = \mathcal{G}(\hat{\mathcal{A}}). \end{aligned} \tag{6}$$

where $blkdiag(\hat{\mathcal{A}}$ is defined in 4 and $\|blkdiag(\hat{\mathcal{A}})\|_*$ is the trace norm of matrix $blkdiag(\hat{\mathcal{A}}$, where $\|\mathcal{A}\|_{TNN} = \|blkdiag(\hat{\mathcal{A}})\|_*$. To solve the problem in 6, we can rewrite 6 as its equivalently form:

$$\begin{aligned} &\text{Minimize} \quad \|blkdiag(\hat{\mathcal{Z}})\|_* + \mathbf{1}_{\hat{\mathcal{M}}=\mathcal{G}(\hat{\mathcal{A}})}, \\ &\text{Subject to} \quad \hat{\mathcal{A}} - \hat{\mathcal{Z}} = 0. \end{aligned} \tag{7}$$

where $\mathbf{1}$ is the indicator function. Thus, we can leverage the general framework of Alternating Direction Method of Multipliers (ADMM) [8] to solve this optimization problem. We can iteratively update $\hat{\mathcal{A}}, \hat{\mathcal{Z}}and\mathcal{Q}$ as follows:

$$\begin{aligned} &\mathcal{A}^{k+1} \\ &= arg\min_{\mathcal{A}} \{\mathbf{1}_{\mathcal{M}=P_\Omega(\mathcal{A})} + \mathcal{A}(:)^T \mathcal{Q}^k(:) + \frac{1}{2}\|\mathcal{A} - \mathcal{Z}^k\|_F^2\} \\ &= arg \min_{\mathcal{A}:\mathcal{M}=P_\Omega(\mathcal{A})} \{\|\mathcal{A} - (\mathcal{Z}^k - \mathcal{Q}^k)\|_F^2\} \end{aligned} \tag{8}$$

$$\hat{\mathcal{Z}}^{k+1} = arg\min_{\hat{\mathcal{Z}}}\{\frac{1}{\rho}\|blkdiag(\hat{\mathcal{Z}})\|_* + \frac{1}{2}\|\hat{\mathcal{Z}} - (\hat{\mathcal{A}}^{k+1} + \hat{\mathcal{Q}}^k)\|_F^2\} \tag{9}$$

$$\mathcal{Q}^{k+1} = \mathcal{Q}^k + (\mathcal{X}^{k+1} - \mathcal{Z}^{k+1}) \tag{10}$$

where Equation 8 represents the least-squares projection on the constraint in 7. We can solve Equation 9 by the singular value thresholding method [9, 27]. In addition, $\mathcal{X}(:)$ and $\mathcal{Q}^k(:)$ are tensor vectorizing forms, which are MATLAB notations.

IV. EXPERIMENT SETUP

As we have mentioned before, mobile detectors can help MISEN explore spectrums' strengths from both space dimension and time dimension. We built a proof-of-concept prototype of MISEN and conducted actual experiments to evaluate its performance. In this section, we will introduce the details of the experiments.

A. Equipment

Fig. 3. Localization module

Our equipment can be grouped into three parts: localization module, mobile detecting module and fixed detecting module. As shown in Fig.3 the equipment in localization module consists of beacons, tags, a laptop computer and several portable power sources. We utilized those beacons to locate the tags. A main beacon was connected to the laptop computer, and all the locating data were transmitted to this laptop computer. Before this localization module comes into working, we measured 25 selected locations by both a standard tape measure and our localization module, and then we did a linear regression to these measurements to calibrate the distance records of localization module.

Fig.4 shows the components of our mobile detecting module. We can see that our measurement devices consist of an iRobot Create 2, a USRP N210 [3], a portable battery bank and a log periodic PCB antenna (400-1000 MHz). We use the USRP board combined with a SBX daughter board and

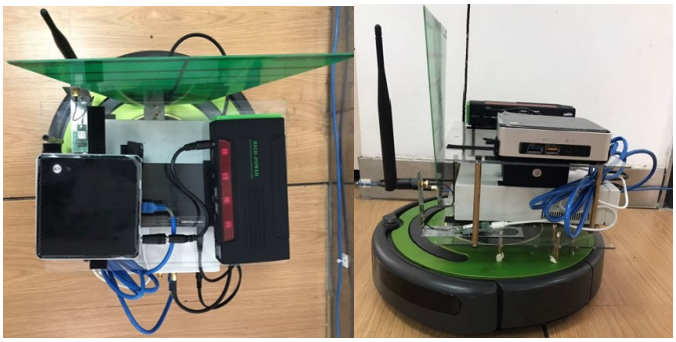


Fig. 4. Details of our mobile detector

a GNU Radio platform to build a spectrum detector. Our daughter board is able to get access to a wide spectrum bands in 400MHz - 4400MHz and has a typical noise figure of 5dB [2]. And we deploy this spectrum detector, a mini host, and another portable battery bank on iRobot Create 2 to construct a mobile detector. Here we leverage a RF signal generator generating signals with different powers to calibrate our signal strength measurements. In this experiment, we utilized two mobile detectors in total, which can help us collect the totally synchronized signal strengths measurements.

As for the fixed detecting module, we also deploy another ten hosts and ten spectrum detectors to acquire a set of ground truth data, which are utilized to estimate the accuracy of our reconstructed data.

B. Scenario and Settings

We set up our spectrum measurement on the 3rd floor of our Lab building. We performed indoor white space signal strength measurements 5 times (May 12th - May 16th, 2017). We selected 40 profiled locations as datum points, where 10 datum points were deployed with fixed detectors to collect ground truth data. We also deployed 4 localization beacons at the 4 corners of our experimental site.

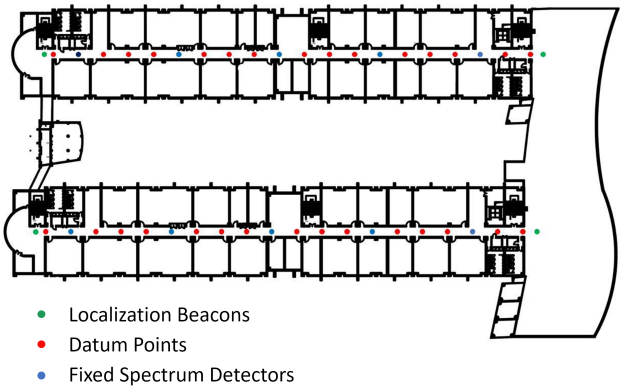


Fig. 5. Experimental site

Fig.5 shows the map of our experimental site, where the red points are datum points, blue points represent fixed spectrum

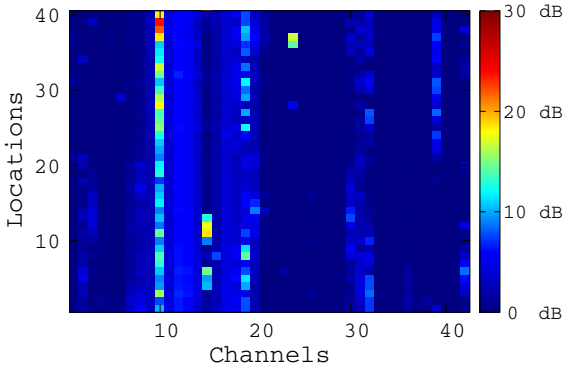


Fig. 6. Indoor white space signal strengths distribution

detectors and we deploy localization beacons on green points. The profiled points are only set on corridors since we can not get access to those professors' personal offices.

Fig.6 shows the signal strengths distribution of the datum locations, where the dark blocks mean those channels with large signal strengths, while the bright blocks mean those channels with low strengths. We can conclude that the white spaces' signal strengths vary quite a bit in our experiment scenario.

The spectrum bands of concern are 470MHz-806MHz, which are the Chinese digital TV bands with 8MHz bandwidth of each channel. Here we leverage a commonly used way, energy detection, to detect the 42 TV channels. Hence, whether a channel is considered to be vacant is totally decided by comparing its energy to a threshold. In our experiment, we set the threshold as -83.5 dBm/8 MHz. Furthermore, we set the gain of the antenna as 0dBi and adopt a 1024 bins FFT with sampling rate of 2MHz in the GNU Radio platform.

In addition, we also conducted another experiment to compare the performance of MISEN with previous systems, where we deploy 22 static detectors in 22 locations. These detectors perform the white space signal strengths measurement once every 5 minutes for 2 weeks.

C. Reliability of Locating System

In our mobile detecting module, acquiring the tracks of mobile detectors play a very important role, because our loss pattern tensors are built basing on these tracks. Here we leverage the localization module to obtain mobile detectors' tracks. Therefore, the reliability of our localization has a great influence on MISEN's performance. In this subsection, we will discuss the reliability of our localization module from space domain and time domain.

As we have said before, we measured a set of points using both a tape measure and our localization module in order to calibrate this localization module. Then we calculated the mean absolute errors at each point basing on these measurements. The results illustrate that the mean absolute errors are no more than 20 centimeters in most of the selected points, and the maximum mean absolute error among those points is around 30 centimeters. While in our experiments, the average distance between neighboring points is 2.5 meters, so the influence of locating errors can be neglected. Consequently, we can conclude that this locating system is reliable in space domain.

Ideally our localization module refreshes 5 times per second. However, in practical application there are some occasions that localization module refreshes once per several seconds or sometimes even tens of seconds due to some sudden obstructions breaking in that may block the signal, like people. We set a threshold 10s here to deal with these cases. If the refreshing interval is smaller than or equal to 10s, we simply think our mobile detector goes along a straight line with a uniform velocity. However, if the refreshing interval is larger than 10s, we would discard the corresponding spectrum data in this interval, since we don't know the exact tracks of our mobile detector.

V. PERFORMANCE EVALUATION

We will discuss the performance evaluation of MISEN in this section. At first, we focus on the influence of the time span of a tensor, which includes different choice of time slices' length and different choice of the amount of time slices. And then, we compare the performance of MISEN with existing systems, including WISER and FIWEX.

A. Methodology

We only considered 40 points in our experiment, because of the limitations of localization accuracy. Although they may not cover the whole building, the results are still a good indicator of MISEN's performance. In addition, we set the threshold of white space as -83.5dBm/8MHz, which is higher but safer than the threshold set in [31]. The evaluation results show that MISEN outperforms WISER and FIWEX in both cost and accuracy.

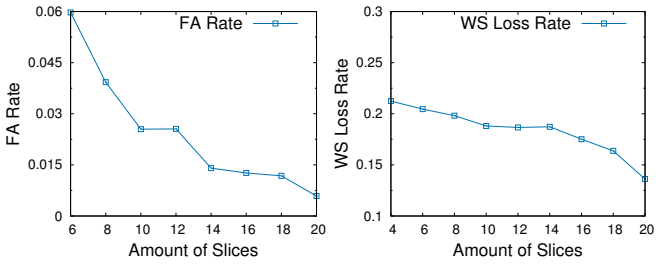
Here we utilize three metrics, including FA Rate and WS Loss rate to evaluate the system performance. Previous literatures have defined them as follows [31]:

- **False Alarm Rate (FA Rate)**: the ratio between the number of channels that a system misidentifies 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 misidentifies as occupied and the total number of actually vacant channels.

B. the Influence of Tensors' Time Span

In our system, a major factor that influences the performance of MISEN is a tensor's total time span T. It is conceivable that the longer the time span T is, the more white space information is contained in a tensor. However, we cannot increase the span of tensor without bound to increase the system performance. That is because the larger span of the tensor means the more data processing time. When the data processing time becomes larger, our central host may fail to output white space map in real-time. Furthermore, according

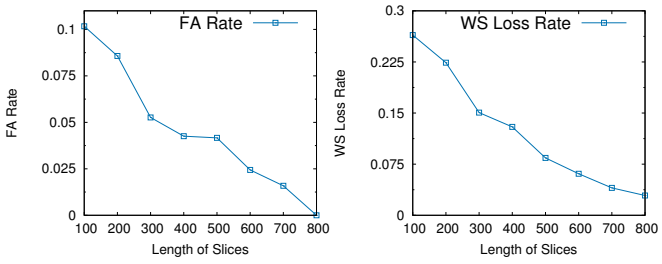
to the concept of '*Marginal utility*' the subsequent unit of utilization of time span yields less utility than the prior units, with a continuing reduction for greater amounts. So, in this subsection, we will discuss the influence of the total time span T on MISEN's performance.



(a) FA Rate (b) WS Loss Rate

Fig. 7. System performance of different amount of time slices

It is evident that the total time span T is affected by two aspects: the amount of time slices, and the time slice's length. Hence, we study the influences of T from these two aspects. Firstly, we fixed the amount of time slices and then changed the time slice's length. Fig.7(a) and Fig.7(b) show the FA Rate and WS Loss Rate of MISEN in different amount of time slices, where we fixed the time slice length as 200 seconds. In Fig.7(a), we can see that the FA Rate curve goes down rapidly when the amount of time slices is small. Then the pace of this curve's decline gradually slowed down with the increasing of the amount of time slices. We can find that the results conform to '*Marginal utility*', which requires a continuing increasing amount of time slices for a lower FA Rate. While Fig.7(b) illustrates a slow decline throughout the whole figure. Considering the characteristics of both FA Rate and WS Loss Rate, we can draw the conclusion that 10 pieces of time slices is a good choice in our scenario, which equals to a total time span of 33 minutes.

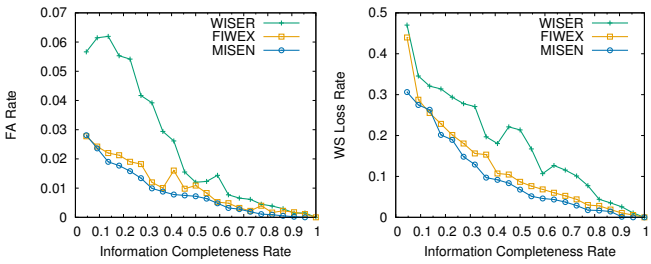


(a) FA Rate (b) WS Loss Rate

Fig. 8. System performance of different time slices' length

Secondly, we fixed the length of the time slice and then changed the number of time slices. And Fig.8(a) and Fig.8(b) illustrate the results of this part, where we fixed the amount of time slices as 4. We can see that both FA Rate and WS Loss Rate go down steadily with the increasing of time slice length. This observation proves that a larger span of the tensor leads to a better system performance.

C. Comparison with FIWEX and WISER



(a) FA Rate (b) WS Loss Rate

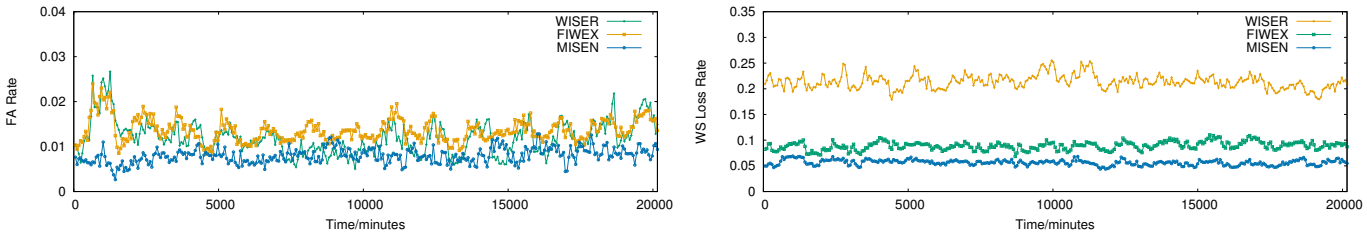
Fig. 9. Comparison with WISER and FIWEX

We will describe MISEN's performance comparison with WISER and FIWEX in this subsection. WISER and FIWEX deploy static detectors, where system performance is mainly determined by the amount of detectors. While MISEN deploys mobile detectors, where system performance is mainly determined by the length of time slices. We noticed that both MISEN's time slice length and FIWEX's detector number correlate to the completeness of collected white space information. Hence, to compare them fairly, we introduce a concept of information completeness rate, which is defined by the number of detected entries divided by total entries of a matrix (or a time slice). We compare the performance of different systems on FA Rate and WS Loss Rate by controlling this information completeness rate. And the results are illustrated in Fig.9. We can observe from 9(a) and 9(b) that MISEN can detect 18.7% more indoor white spaces with 20.0% less false alarms than FIWEX, when they have the same information completeness rate. Furthermore, MISEN can identify 46.1% more indoor white spaces with 64.9% less false alarms than WISER with the same information completeness rate. We also studied the time varying performance of these three systems in Fig.10, where we can conclude that MISEN has a stable lower FA Rate and lower WS Loss Rate with time passing by.

VI. RELATED WORK

Since FCC allowed unlicensed devices to get access to vacant TV channels, TV white spaces have gained a wide attention. For instance, [5] studied the white space characteristics on spatial variation, spectrum fragmentation and temporal variation. In addition, researchers have performed a lot of outdoor white space measurements in metropolitan cities like Chicago [23], Guangzhou [30], Singapore [17] and Hong Kong [31]. There are also plenty of works focusing on the applications for TV white spaces. Bahl et al. proposed the first white space Wi-Fi like wireless network, named WhiteFi in [5]. A scalable sensor network architecture: Sensor Network Over White Spaces (SNOW) is proposed by [25]. Researchers also make efforts to utilize active TV channels, including [34] designs a system to enable WiFi transmission in active TV channels (WATCH).

Besides, as people say, spectrum detecting is a primary task in cognitive radio. To make use of white spaces, we need to



(a) FA Rate

(b) WS Loss Rate

Fig. 10. Time varying performance of MISEN, FIWEX and WISER

identify white spaces first. Bansal et al. proposed a metric to evaluate users' sensing accuracy and designed an algorithm maximizing the sensing accuracy [7]. However, sensing the spectrums every time when unlicensed users want to utilize white spaces is sometimes a little tedious. Thus, we want to build a query type database for users to acquire white space availabilities, where the task of spectrum detecting is turned over to a specific sensing module. Many previous works focus on improving outdoor white space databases, which are built based on signal propagation model, including [10, 24].

There are also some researchers study white space exploration methods in indoor scenarios. Ying et al. proposed the first indoor white space identification system WISER, which utilizes the channel-location clustering based algorithm to explore indoor white spaces [32]. Then, [21] presented FIWEX, a cost-efficient indoor white space exploration mechanism basing on compressive sensing algorithm. By taking the white space linear dependence of locations and channels into consideration, FIWEX obtained great improvement compared to WISER. However, both of these two works do not consider the time correlation of indoor white space characteristics. Motivated by this observation, we proposed our mobile indoor spectrum exploration mechanism, MISEN.

We used tensor completion method to recover frequency spectrum data. Tensor completion is a widely used signal processing technique which helps acquire and reconstruct signals in an efficient way. The theory of tensor completion has been widely studied in the past few years [13, 15, 19, 29]. As the development of the tensor completion theory, people begin to apply it to a lot of other fields, including computer vision [22, 35] and machine learning for online prediction of ratings [16]. Specially, Zhang et al. [35] put forward novel methods for tensor completion based on tensor-Singular Value Decomposition (t-SVD) and a related tensor nuclear norm proposed by [19]. Furthermore, Zhang et al. apply this algorithm to video recovery, which shows a superior performance over previous methods.

The route plan design is an important part of the mobile detecting technique. In our system, we just apply a random route. But it is obvious that a proper route can lead to a good performance. Ji et al. put forward the concept of Bayesian compressive sensing and designed an adaptive projection selection method using differential entropy [18]. Adaptive compressive sensing, which has been widely used in wireless sensor network [12, 26] and other realms, can help us to design a more efficient route plan in white space detecting scenario, which is one of our future works.

VII. CONCLUSION

In this paper, we creatively utilize mobile detectors to explore indoor white spaces, and then further exploit collected white space information by correlations of location, channel, and time dimensions. Based on these motivations, we proposed a mobile indoor white space exploration system, namely MISEN. We also built a proof-of-concept prototype of MISEN and conducted actual experiments to evaluate its performance. The results illustrate that MISEN yields a superior performance over existing methods, where MISEN can detect 18.7% more white spaces with 20.0% less false rate than the best known existing solution when they have the same information completeness rate.

REFERENCES

[1] Average spectrum occupancy by band in chicago and new york. http://defenseelectronicsmag.com/.

[2] Sbx 400-4400 mhz rx/tx (40 mhz). https://www.ettus.com/product/details/SBX.

[3] Universal software radio peripheral. https://www.ettus.com.

[4] R. Ahuja, R. Corke, and A. Bok. Cognitive radio system using IEEE 802.11 a over UHF TVWS. In *Proceedings of DySPAN*, pages 1–9. IEEE, 2008.

[5] P. Bahl, R. Chandra, T. Moscibroda, R. Murty, and M. Welsh. White space networking with wi-fi like connectivity. In *Proceedings of SIGCOMM*, pages 27–38. ACM, 2009.

[6] R. Balamurthi, H. Joshi, C. Nguyen, A. K. Sadek, S. J. Shellhammer, and C. Shen. A TV white space spectrum sensing prototype. In *Proceedings of DySPAN*, pages 297–307. IEEE, 2011.

[7] T. Bansal, B. Chen, and P. Sinha. Discern: Cooperative whitespace scanning in practical environments. In *INFOCOM, 2013 Proceedings IEEE*, pages 719–727. IEEE, 2013.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.

[9] J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

[10] A. Chakraborty and S. R. Das. Measurement-augmented spectrum databases for white space spectrum. In *Proceedings of the 10th ACM International on Conference on emerging Networking Experiments and Technologies*, pages 67–74. ACM, 2014.

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

[12] C. Chou, R. Rana, and W. Hu. Energy efficient information collection in wireless sensor networks using adaptive compressive sensing. In *Proceedings of LCM*, pages 443–450. IEEE, 2009.

[13] G. Ely, S. Aeron, N. Hao, and M. E. Kilmer. 5d and 4d pre-stack seismic data completion using tensor nuclear norm (tnn). In *SEG Technical Program Expanded Abstracts 2013*, pages 3639–3644. Society of Exploration Geophysicists, 2013.

[14] Z. Farid, R. Nordin, and M. Ismail. Recent advances in wireless indoor localization techniques and system. *Journal of Computer Networks and Communications*, 2013, 2013.

[15] S. Gandy, B. Recht, and I. Yamada. Tensor completion and low-n-rank tensor recovery via convex optimization. *Inverse Problems*, 27(2):025010, 2011.

[16] E. Hazan, S. Kale, and S. Shalev-Shwartz. Near-optimal algorithms for online matrix prediction. In *Conference on Learning Theory*, pages 38–1, 2012.

[17] M. H. Islam, C. L. Koh, S. W. Oh, X. Qing, Y. Y. Lai, C. Wang, Y. Liang, B. E. Toh, F. Chin, G. L. Tan, and W. Toh. Spectrum survey in singapore: Occupancy measurements and analyses. In *Proceeding of CrownCom*, pages 1–7. IEEE, 2008.

[18] S. Ji, Y. Xue, and L. Carin. Bayesian compressive sensing. *IEEE Transactions on Signal Processing*, 56(6):2346–2356, 2008.

[19] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 34(1):148–172, 2013.

[20] N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, W. H. Engelmann, et al. 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.

[21] D. Liu, Z. Wu, F. Wu, Y. Zhang, and G. Chen. Fiwex: Compressive sensing based cost-efficient indoor white space exploration. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, pages 17–26. ACM, 2015.

[22] J. Liu, P. Musialski, P. Wonka, and J. Ye. Tensor completion for estimating missing values in visual data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):208–220, 2013.

[23] M. A. McHenry, P. A. Tenhula, D. McCloskey, D. A. Roberson, and C. S. Hood. Chicago spectrum occupancy measurements & analysis and a long-term studies proposal. In *Proceedings of the first international workshop on Technology and policy for accessing spectrum*, page 1. ACM, 2006.

[24] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl. Senseless: A database-driven white spaces network. *IEEE Transactions on Mobile Computing*, 11(2):189–203, 2012.

[25] A. Saifullah, M. Rahman, D. Ismail, C. Lu, R. Chandra, and J. Liu. Snow: Sensor network over white spaces. In *Proceedings of the International Conference on Embedded Networked Sensor Systems (ACM SenSys)*, 2016.

[26] J. Wang, S. Tang, B. Yin, and X. Li. Data gathering in wireless sensor networks through intelligent compressive sensing. In *Proceedings of INFOCOM*, pages 603–611. IEEE, 2012.

[27] G. A. Watson. Characterization of the subdifferential of some matrix norms. *Linear algebra and its applications*, 170:33–45, 1992.

[28] C. Wu, Z. Yang, Y. Liu, and W. Xi. Will: Wireless indoor localization without site survey. *IEEE Transactions on Parallel and Distributed Systems*, 24(4):839–848, 2013.

[29] Y. Xu and W. Yin. A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. *SIAM Journal on imaging sciences*, 6(3):1758–1789, 2013.

[30] S. Yin, D. Chen, Q. Zhang, M. Liu, and S. Li. Mining spectrum usage data: a large-scale spectrum measurement study. *IEEE Transactions on Mobile Computing*, 11(6):1033–1046, 2012.

[31] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra. Exploring indoor white spaces in metropolises. In *Proceedings of the 19th annual international conference on Mobile computing & networking*, pages 255–266. ACM, 2013.

[32] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra. Exploring indoor white spaces in metropolises. In *Proceedings of MOBICOM*, pages 255–266. ACM, 2013.

[33] T. Zhang, N. Leng, and S. Banerjee. A vehicle-based measurement framework for enhancing whitespace spectrum databases. In *Proceedings of MOBICOM*, pages 17–28. ACM, 2014.

[34] X. Zhang and E. W. Knightly. Watch: Wifi in active tv channels. *IEEE Transactions on Cognitive Communications and Networking*, 2(4):330–342, 2016.

[35] Z. Zhang, G. Ely, S. Aeron, N. Hao, and M. Kilmer. Novel methods for multilinear data completion and denoising based on tensor-svd. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3842–3849, 2014.