

WhiFind: A Matrix Completion Approach for Indoor White Space Identification

Xiao Hu, Fan Wu, Xiaofeng Gao and Guihai Chen
 Shanghai Key Laboratory of Scalable Computing and Systems
 Department of Computer Science and Engineering
 Shanghai Jiao Tong University, China
 Email: hux226@sjtu.edu.cn, {fwu, gao-xf, gchen}@cs.sjtu.edu.cn

Abstract—In this paper, we introduce WhiFind, which is an indoor white space identification system. It leverages matrix completion techniques to provide accurate indoor white space availability information with only a small number of spectrum sensors deployed.

I. INTRODUCTION

Utilizing white spaces is a promising way to fulfill the growing demand for available spectrum resources. One of the fundamental issues in utilizing white spaces is to accurately recognize occupied and vacant channels. While most existing works focus on outdoor scenarios, only a limited number of works deal with indoor white spaces. Among them, Ying et al. [1] measured white spaces in a typical metropolis Hong Kong, and revealed that more than 70% of the TV spectrum bands are white spaces in indoor scenarios. They proposed an indoor white space determination system, called WISER, which increases the number of TV channels identified for indoor white space networking by building a more accurate white space database locally. However, it simply uses the anchor sensors' readings to represent those of other locations, which is likely to lead to errors and loss of white spaces.

In this poster, we carry out intensive measurements and perform comprehensive analyses on indoor white space features. We observe that indoor white spaces have several interesting and important features, which are very useful to help us to design a novel matrix completion approach to reduce the sensing overhead and to improve the identification accuracy.

II. UNDERSTANDING INDOOR WHITE SPACE FEATURES

In this section, we motivate the design of WhiFind by demonstrating several major features of indoor white spaces, including low-rank structure, temporal stability, spatial and spectral correlations.

We carry out measurements on the third floor of SEIEE buildings in Shanghai Jiao Tong University. A total of 67 indoor locations are selected for the measurements. At each location, we record the received signal strength of 45 TV

channels. In practical measurements, we profile the locations one by one with the aid of a movable handcart. At each location we scan all channels automatically by running a Python script. Then we can obtain a 45×67 matrix, called Received Signal Strength Matrix (RSS Matrix). We collected 20 RSS Matrices like this in total over two months.

We mine the low-rank structure feature in above measurement data by using Principal Component Analysis (PCA). Any $m \times n$ matrix X can be decomposed into three matrices:

$$X = U\Sigma V^T = \sum_{i=1}^{\min(m,n)} \sigma_i u_i v_i^T, \text{ where } U \text{ is a } m \times m$$

unitary matrix, V is a $n \times n$ unitary matrix, and Σ is a $m \times n$ diagonal matrix containing the singular value σ_i of X , which indicates the energy of i -th principal component according to PCA. We analyze the distribution of singular values of RSS Matrices, and find that the energy is always contributed by the top 5% singular values. It means that the RSS Matrix can be approximated by a matrix of low rank. This low-rank structure feature provides us the prerequisite for using matrix completion.

We analyze the distribution of RSS on time domain to reveal the temporal feature. We measure the temporal stability of channel i at location j by computing the normalized difference values between adjacent time slots. We observe that more than 80% values are very small, which indicates that the temporal stability feature does exist for indoor white spaces.

There are another two important observations from the measurement results. First, we observe that there are strong correlations between multiple locations in their spectrum usage patterns. We refer this observation as the spatial correlation feature. Second, for multiple channels, there are strong correlations in their RSS across different locations. We refer this observation as the spectral correlation feature. These two features may be caused by the geographical approach of locations and sharing the same TV tower. It suggests that we can calculate a channel RSS at a location from its multiple correlated channels at other locations.

III. WHIFIND DESIGN

A. System Overview

WhiFind is an indoor white space identification system based on matrix completion. It consists of one white space database and a limited number of indoor sensors. With WhiFind, every node detects the channel utilizations and send them to the database. The database employs a novel matrix

This work was supported in part by the State Key Development Program for Basic Research of China (973 project 2012CB316201), in part by China NSF grant 61422208, 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.

F. Wu is the corresponding author.

completion method to calculate the channel RSS at locations without sensor deployed. To realize this functionality, we first conduct spectrum profiling at all locations. Then we perform k -medoids clustering on the locations, and choose the center of each cluster to deploy our sensors. The database receive real-time readings periodically from the sensors and then perform a matrix completion approach to calculate the unknown entries in RSS Matrix.

B. Matrix Completion Approach

If we take time domain into consideration, the RSS Matrix is actually a 3-dimension array (e.g., $X(i, j, k)$ denotes the received signal strength of channel i at location j at time slot t). In order to apply the matrix completion, we convert it into a 2-dimension matrix by collapsing the first two dimensions of X into one dimension like $X(k, t) = X(i + (j - 1)n, t)$.

As mentioned previously, X can be approximated by a low-rank matrix \hat{X} . We propose to obtain \hat{X} by solving an optimization program according to matrix completion theory. This program finds matrix with minimum rank that fits the data:

$$\begin{aligned} & \text{minimize} \quad \text{rank}(\hat{X}) \\ & \text{subject to} \quad M.*\hat{X} = X, \end{aligned} \quad (1)$$

where M is a binary matrix that indicates where a entry in X is unknown, and $*$ is the element-wise production. As this optimization is NP-hard, we present a SVD-like factorization: $\hat{X} = U\Sigma V^T = LR^T$, where $L = U\Sigma^{1/2}$ and $R = V\Sigma^{1/2}$. Specifically, if RIP holds, minimization the nuclear norm can result to rank minimization. Hereby, we just need to find L and R that minimize the summation of their Frobenius norms:

$$\begin{aligned} & \text{minimize} \quad \|L\|_F^2 + \|R\|_F^2, \\ & \text{subject to} \quad M.*(LR^T) = X. \end{aligned} \quad (2)$$

Instead of solving Eqn. (2) directly, we solve the following equation using the Lagrange multiplier method:

$$\text{minimize} \quad \|M.*(LR^T) - X\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2), \quad (3)$$

where the Lagrange multiplier λ is a tunable parameter.

To exploit the observations of indoor white space features, we complete the matrix completion approach by developing Eqn. (3) as follows:

$$\begin{aligned} & \text{minimize} \quad \|M.*(LR^T) - X\|_F^2 + \lambda(\|L\|_F^2 + \|R\|_F^2) \\ & \quad + \|S(LR^T)\|_F^2 + \|(LR^T)T^T\|_F^2, \end{aligned} \quad (4)$$

where S is the spatial and spectral constraint matrix, and T is the temporal constraint matrix respectively. A simple choice for the temporal constraint matrix is $T = \text{Toeplitz}(0, 1, -1)$, which denotes the Toeplitz matrix with central diagonal given by ones, and the first upper diagonal given by negative ones. And we choose S based on k-nearest neighbor algorithm (KNN) and linear regression. For each row i of X , we find K most correlated rows $j_k \neq i$ ($k = 1, \dots, K$). We perform linear regression to find a set of weights $w(k)$ such that the linear combination of rows j_k best approximates row i : $X(i, *) \approx \sum_{k=1}^K w(k)X(j_k, *)$. Assuming that $\sum_{k=1}^K w(k)X(j_k, *)$ approximates $X(i, *)$ well, we then set

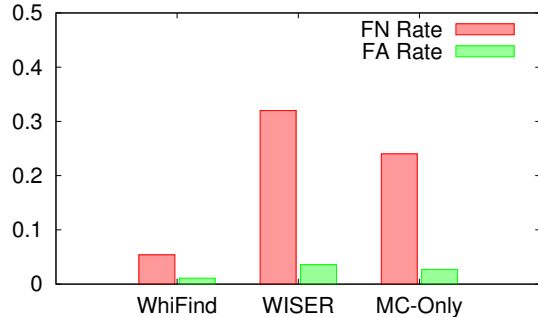


Fig. 1. The average FA and FN rates for WhiFind, WISER and MC-Only.

$S(i, i) = 1$ and $S(i, j_k) = -w(k)$ for $k = 1, 2, \dots, K$. The rest entries of S matrix is given by zeros.

There are many different methods to solve this optimization problem [2] [3]. In this paper, the alternating least squares procedure is adopted.

IV. EVALUATION

In this section, we build a WhiFind prototype and conduct real-world experiments to evaluate its performance. The prototype consists of one white space database and 20 indoor sensors. We borrow the metrics FA and FN rates from [1] to quantify the performance. We compare it with other alternative solutions, WISER and MC-Only. MC-Only recovers RSS Matrix without taking indoor white space features into consideration. The comparison results are depicted in Fig. 1. With WhiFind, the average FA rate is 1.1% and the average FN rate is 5.4%. The rates are 3.6% and 32% with WISER, 2.7% and 24% with MC-Only respectively. We can see that, WISER and MC-Only have similar average FA rates, which are both higher than that of WhiFind. However, the average FN rate of WhiFind is much smaller than those of WISER and MC-Only. Specifically, WhiFind can identify 26.6% more white spaces correctly.

V. CONCLUSION

In this paper, we have carried out a measurement-driven approach to exploring indoor white space identification. We have revealed that the indoor white spaces have several interesting and important features. Motivated by that, we then proposed a system called WhiFind, which is able to utilize the features to identify indoor white spaces. In particularly, the evaluation results show that our system outperform the best known existing solution WISER and can identify 26.6% more indoor white spaces.

REFERENCES

- [1] X. Ying, J. Zhang, L. Yan, G. Zhang, M. Chen, and R. Chandra, "Exploring indoor white spaces in metropolises," in *Proceedings of MOBIKOM*. ACM, 2013, pp. 255–266.
- [2] B. Recht, M. Fazel, and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization," *SIAM Review*, vol. 52, no. 3, pp. 471–501, 2010.
- [3] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, "Spatio-temporal compressive sensing and internet traffic matrices," in *Proceedings of SIGCOMM*.