FitLoc: Fine-grained and Low-cost Device-free Localization for Multiple Targets over Various Areas

Liqiong Chang[†], Xiaojiang Chen[†*], Yu Wang[‡], Dingyi Fang[†], Ju Wang[†], Tianzhang Xing[†], Zhanyong Tang[†]
[†]Northwest University, [‡]University of North Carolina at Charlotte
[†] changliqiong@stumail.nwu.edu.cn, {xjchen,dyf,ju-w,xtz,zytang}@nwu.edu.cn, [‡]yu.wang@uncc.edu

***Abstract*—Device-free localization (DfL) techniques, which can localize targets without carrying any wireless devices, have attracting an increasing attentions. Most current DfL approaches, however, have two main drawbacks hindering their practical applications. First, one needs to collect large number of measurements to achieve a high localization accuracy, inevitably causing a high deployment cost, and the areas variety will further exacerbate this problem. Second, as the pre-obtained Received Signal Strength (RSS) from each location (*i.e.,* radio-map) in a specific area cannot be directly applied to new areas for localization, the calibration process of different areas will lead to the high human effort cost.**

In this paper, we propose, FitLoc, a fine-grained and low cost DfL approach that can localize multiple targets in various areas. By taking advantage of the compressive sensing (CS) theory, FitLoc decreases the deployment cost by collecting only a few of RSS measurements and performs a fine-grained localization. Further, FitLoc employs a rigorously designed transfer scheme to unify the radio-map over various areas, thus greatly reduces the human effort cost. Theoretical analysis about the effectivity of the problem formulation is provided. Extensive experimental results illustrate the effectiveness of FitLoc.

I. Introduction

Past decade has witnessed the pervasiveness and advances of the localization approaches/systems. However most methods require the targets to carry certain wireless devices [1], [2]. In practice, there are many emerging applications unable to meet this requirement, for example to localize the rare animals, the zoologists forbid equipping the giant pandas or golden monkeys with any devices, and in the intrusion detection, it is impossible to pre-install the tracking device on the intruders [3]–[5]. Therefore, the Device-free Localization (DfL) techniques [3]–[14], which don't need the targets to carry any devices, have become one of the most attractive techniques to researchers and industrials.

With the ubiquitous and low-cost Received Signal Strength (RSS) in WiFi, RFID, etc, DfL technique performs localization by utilizing the RSS measurements distorted by the targets. Particularly, current methods [13], [14] establish the model between RSS and locations, but provide a limited performance when considering both the accuracy and robustness, since the models are vulnerable to environmental noise. In contrast, radio-map methods [3]–[5], [7]–[9] could achieve a fine-grained localization accuracy by matching the real-time RSS with the pre-obtained radio-map, and have been extensively studied and become the preferred approach recent years.

However, two key enablers to make the radio-map based Device-free Localization method fully practical still remains unsolved, *the high deployment cost* and *high human effort cost* when the localization environment changes from one area to other areas. In reality, applications which need to conduct localization over various areas exist widely in outdoor environment [15].

High deployment cost. To achieve a fine-grained localization accuracy, most DfL approaches even including the state-of-the-art work [5] and [13], need to collect large number of RSS measurements from a set of dense deployed wireless transceivers, causing a high deployment cost. In addition, the transceivers always have a limited power supply in the outdoor environment [16], [17]. Hence when one conduct localization over various areas, it will not only bring more expensive deployment costs, but also result in the low scalability for localization. Therefore, the high deployment cost is one key constraint hindering the DfL technique to be fully practical.

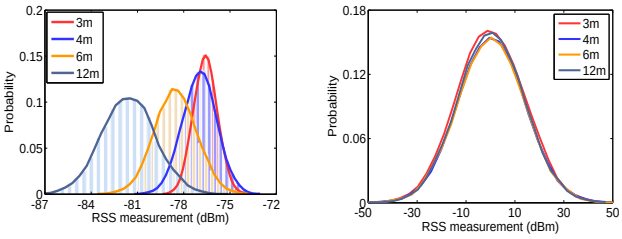
High human effort cost. When one uses the radio-map of one area to conduct localization over other areas, the radio-map of areas with different sizes are significantly different. In fact, different areas require to deploy different lengths of wireless links[1], and the RSS distributions under different link lengths are different, as shown in Fig. 1(a). Thus, real-time RSS collected in the new area would drastically deviate from the radio-map of the original area, and the localization accuracy will decrease dramatically in the new area. To deal with this problem, an intuitive solution is to repeat the manually calibration in each area, but it is a time consuming and labor intensive process[2].

Owing to the sparse recovery property of the Compressive Sensing (CS) theory [18], [19], recent work [9]–[12] could localize targets by collecting less RSS measurements. Particularly, in light of the sparse property of the localization problem, *i.e.*, the number of targets is sparse relative to the number of locations, work [9] can achieve a fine-grained localization accuracy for multiple targets, which deploys a small number of transceivers and collects only a few of RSS measurements, thus reducing the deployment cost greatly. However, this approach doesn't work out when one performs localization over various areas due to the high human effort cost caused by the repeated calibration for each area.

To overcome these limitations, *in this paper, we introduce FitLoc, the first fine-grained and low-cost device-free localization for multiple targets over various areas*. In line with the common CS based DfL, FitLoc recovers the location vector by solving an ℓ_1 optimization with the sensing matrix (radio-map) and real-time RSS, and reduces the deployment

[1]A link length is the distance between the wireless transceivers.

[2]In fact, the radio map are established through manually recording the RSS change which are distorted by the target standing in all the possible locations.



(a) before transfer (b) after transfer

Fig. 1: The radio-map differences over different areas are reflected through the RSS distributions. By collecting the RSS of two specific wireless transceivers distorted by the target at a fixed location, (a) shows that the Gaussian probability estimates of the raw RSS under 4 different link lengths (*i.e.*, 3m, 4m, 6m and 12m) are significantly different from each other. (b) shows the distribution distances are made minimum after being transferred by FitLoc.

cost benefiting from the CS theory. The challenge however is to perform localization with the sensing matrix from one area over various areas. Unlike existing approaches which don't take the area change into consideration, FitLoc aims at minimizing the calibration effort cost when the area changes. By collecting RSS of only a few locations in the new area, and integrating it with the sensing matrix of the original area, FitLoc transfers the sensing matrix and reuses it in the new area, as a result, the human effort cost is reduced greatly.

So how to design a transfer scheme that different areas can share an unified sensing matrix? To do so, one needs to project the RSS of different areas into a subspace where the distribution distances are minimized. Note that the target will distort some links when he/she locates in the area, and the RSS distribution differences are caused by not only the location differences but also the link differences. That is to say for one area, RSS distributions of the same one link but with different target locations are different, and distributions of different links even with the same target location are different. Further, two areas with different link lengths can be treated as one area scaled in size, and distributions of same target locations for a specific link in two areas are different. Consequently, the RSS distribution distances between (i) different locations for the same one link should be maximized; (ii) different links with the same length and target location should be maximized; (iii) same locations for a specific link with different link lengths should be minimized (as shown in Fig. 1(b)).

To design such a transfer scheme, our key observation is that different locations and links can be regarded as different class labels. Then inspired by the transfer method proposed in image processing which transfers the information of different domains [20], we design a transfer scheme by utilizing Fisher Linearly Discriminant Analysis (FLDA) [21] as the subspace learning method to project the RSS into a low-dimensional subspace, and use the Bregman Divergence [22] as a regularization term to measure the RSS distribution distances. Toward this end, different areas can share a unified sensing matrix. The details of the transfer scheme are presented in Section IV, and its efficiency is demonstrated in Section VI.

Based on the proposed transfer scheme, whether the CS based DfL after transfer still satisfies the CS theory to perform localization? To deal with this challenge, we provide a comprehensive analysis to show that the sensing matrix after being transferred obeys the Restricted Isometry Property (RIP) with a high probability, which justifies the viability of the transfer CS based DfL approach.

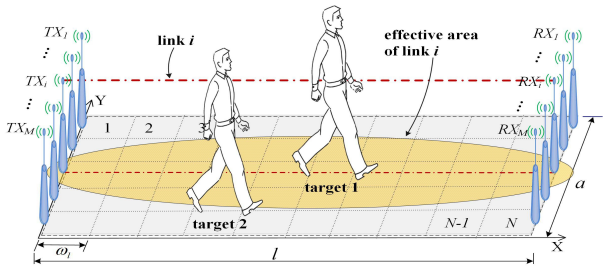


Fig. 2: Deployment of DfL System. There are two targets respectively stand in two different locations, and when the target 1 within the effective area of wireless link i which are formed by the transceivers TX_i and RX_i, the link will be distorted and the RSS measurement will change [9].

Contributions: in summary, we make the following contributions:

- We present the first fine-grained and low-cost DfL approach for multiple targets over various areas, named FitLoc, paving the practical application road of DfL in the outdoor environment.
- To our best knowledge, we are the first to design a transfer scheme that combines FLDA and Bregman Divergence to make the RSS of one area be shared by other different areas, and the basic idea can be extended to other systems.
- We formally prove that the sensing matrix after being transferred still obeys RIP with a high probability, which is the sufficient condition to enable CS based DfL.
- We perform extensive experiments to illustrate the effectiveness and robustness of FitLoc, and the results show that FitLoc achieves an average localization error of 0.89m over different areas in the outdoor environment and 1.4m in similar furnitured indoor environment.

II. Background of CS based DfL

This section introduces the Compressive Sensing (CS) based DfL approach [9], which can accurately localize multiple targets with a small number of measurements, thus reducing the deployment cost. Suppose that there are K targets randomly located in an area with size $l \times a$. The area is equally divided into N grids with edge length ω_l. By deploying M nodes at the midpoint of a grid edge on two sides of the area, there are M links formed by pairs of $\{TX_i, RX_i\}$-nodes ($i \in [1, M]$), as Fig. 2 shows. Then the locations of K targets over N grids can be denoted as a location vector

$$\Theta = [\theta_1, \cdots, \theta_j, \cdots, \theta_N]^T, \tag{1}$$

where $\theta_j \in \{0, 1\}$, and $\theta_j = 1$ when one target is at grid j; otherwise $\theta_j = 0$. Since only K elements of Θ are nonzero, Θ is a K-sparse ($K \ll N$) signal [9]. According to the CS theory, the N dimensional location vector Θ can be recovered from M dimensional RSS vector $Y_{M\times 1}$,

$$Y_{M\times 1} = \mathrm{X}_{M\times N} \cdot \Theta_{N\times 1} + \mathcal{N}, \tag{2}$$

here $Y_{M\times 1}$ is the RSS vector measured in the localization phase, $X_{M\times N}$ is the sensing matrix constructed in the pre-deployment phase, $\mathcal{N}$ is the measurement noise. In this paper, $\mathrm{X}_{M\times N} = (x_{ij})$, x_{ij} is the RSS change[3] of i-th link distorted by one target at grid j, $Y_{M\times 1} = (y_{i1})$, (y_{i1}) is the real-time

[3]RSS change is the difference between RSS measurement before and after the target distort the link.

RSS change of i-th link distorted by targets with unknown locations. It has been proved that if X satisfies the Restricted Isometry Property (RIP) and the dimension of the Y obeys $M = O(K \log(N/K))$ [9], Θ can be exactly recovered through the ℓ_1 optimization:

$$\min \|\hat{\Theta}\|_1 \quad s.t. \ \|\mathrm{X}^{\dagger}(Y - \mathrm{X}\hat{\Theta})\|_{\infty} < c\delta\sqrt{2\log M}, \tag{3}$$

where $()^{\dagger}$ is the pseudo-inverse operator, $c > 0$ is a constant, δ is a constant, and M is the number of measurements.

It is crucial to design the sensing matrix. In order to better understand the RSS, we collect Q continuous RSS as a stream for each link, *i.e.,* $\mathbf{x}_{ij} = \{x_{ij}(1), \cdots, x_{ij}(Q)\}^T$, $\mathbf{y}_i = \{y_i(1), \cdots, y_i(Q)\}^T$. Thus we have 3-dimensional sensing matrix and 2-dimensional RSS vector,

$$\mathrm{X}_{M\times N\times Q} = \begin{bmatrix} \mathbf{x}_{11} & \cdots & \mathbf{x}_{1N} \\ \vdots & \mathbf{x}_{ij} & \vdots \\ \mathbf{x}_{M1} & \cdots & \mathbf{x}_{MN} \end{bmatrix}, \tag{4}$$

$$Y_{M\times 1\times Q} = [\mathbf{y}_1, \cdots, \mathbf{y}_i, \cdots, \mathbf{y}_M]^T. \tag{5}$$

After that, to acquire the true RSS distorted by target, we set the most frequent value among Q values as the true value of one link, *i.e.,* $x_{ij} = \arg\max p(\mathbf{x}_{ij}(o))$, $y_i = \arg\max p(\mathbf{y}_i(o))$, where $1 \le o \le Q$ and $p(\cdot)$ is the probability of the Gaussian estimation. By doing so, the sensing matrix $\mathrm{X}_{M\times N\times Q}$ and RSS vector $Y_{M\times 1\times Q}$ are simplified as $\mathrm{X}_{M\times N}$ and $Y_{M\times 1}$.

III. FitLoc Overview

In view of CS based DfL, FitLoc transfers the sensing matrix of one area and performs localization over other areas with little cost. By assuming that there are two areas with different sizes, FitLoc goes through the following steps

1) We construct the sensing matrix of original area, and collect RSS of only a few randomly chosen locations in the new area. Then FitLoc solves the transfer matrix on the basis of the transfer scheme.
2) We collect the real-time RSS of unknown locations in the new area. After that, we transfer the sensing matrix and real-time RSS vector. At last, FitLoc performs localization in the new area based on the CS theory.

Now we briefly present the framework of the transfer scheme, and leave the detailed descriptions in Section IV. Consider two areas depicted in Fig.3, which are with link length l (area 1 with size $l \times a$, grid edge length ω_l) and u (area 2 with size $u \times b$, grid edge length ω_u), respectively, and suppose that $l < u$. For each area, the number of location grids is N. For the same grid j, the RSS stream distributions of the first link significantly differs from each other, thus the sensing matrix are completely different. While it requires a lot of time and labor to reconstruct the sensing matrix in area 2, we aim to find a transfer matrix $W \in \mathbb{R}^{Q\times q}$ to make the RSS distribution distances minimized in the subspace $\mathbb{R}^q$, so that we can use the sensing matrix of area 1 after being transferred to locate the targets in area 2. The framework of the transfer scheme includes two phases.

In the pre-deployment phase, we transfer the sensing matrix of area 1. By collecting RSS of n ($n \le N$) randomly choosing locations in area 2 and integrate with the sensing matrix of area 1, we can find a transfer matrix W, with which the RSS distribution distances of $\mathbf{x}_{ij}^l$ and $\mathbf{x}_{ij}^u$ (the i-th link and location j with different length l and u) are minimized, as Fig.3 shows. Then the sensing matrix of area 1 after being transferred is

$$\mathrm{Z}_{M\times N\times q} = \left(W^T\mathbf{x}_{ij}^l\right), \ i \in [1, M], j \in [1, N]. \tag{6}$$

In the localization phase, we transfer the real-time RSS stream $\mathbf{y}_i^u$ of area 2, and the RSS vector after being transferred is

$$Y'_{M\times 1\times q} = \left(W^T\mathbf{y}_i^u\right), \ i \in [1, M]. \tag{7}$$

By choosing the most frequent RSS of each stream after being transferred, $\mathrm{Z}_{M\times N\times q}$ and $Y'_{M\times 1\times q}$ can be reduced to $\mathrm{Z}_{M\times N}$ and $Y'_{M\times 1}$. And the location vector Θ in area $u \times b$ can be recovered via the following ℓ_1 optimization [9],

$$\min \|\hat{\Theta}\|_1, \ s.t. \ \|(\mathrm{Z})^{\dagger}(Y' - \mathrm{Z}\hat{\Theta})\|_{\infty} < c\delta\sqrt{2\log M}. \tag{8}$$

Note that as the grids number is fixed to N and the grid size of area 2 is bigger than area 1, the localization accuracy will decrease in area 2. To tackle this problem, we divide the grids of area 2 into subgrids, then we propose a distance based grid interpolation method to generate the RSS of subgrids.

IV. Transfer Scheme Across Various Areas

This section presents the detailed process of the transfer scheme. Specifically, we need to project the RSS distributions and minimize the distribution distances. However, as different RSS streams come from different distributions, the simple methods such as Principle Component Analysis (PCA) and Fisher Linear Discriminant Analysis (FLDA) [21], cannot be directly used to transfer the RSS streams to minimize the distribution distances. To do so, we add a regularization term to the basic subspace learning method to measure the distribution distances. Further, the well-known PCA cannot separate RSS distributions of different locations as it is an unsupervised method. In contrast, FLDA uses the class label information to maximize the distances between different classes and minimize the distance within the same class. On the other hand, we use the Bregman Divergence [22] to measure the distribution distances. Therefore, the transfer scheme can be formulated as a FLDA subspace learning method combined with a Bregman Divergence based regularization term.

A. Formulate the transfer scheme

We first collect a few RSS measurements of n randomly chosen locations in area 2, then integrate it with the sensing matrix of area 1. Let $\mathbf{x}^l$ be the RSS stream vector of N locations in area 1 (*i.e.,* the vector description form of the sensing matrix), and $\mathbf{x}^u$ be the RSS stream vector of n chosen locations in area 2, $n \ll N$,

$$\mathbf{x}^l = (\mathbf{x}_{11}^l, \cdots, \mathbf{x}_{M1}^l, \cdots, \mathbf{x}_{1j}^l, \cdots, \mathbf{x}_{Mj}^l, \cdots, \mathbf{x}_{1N}^l, \cdots, \mathbf{x}_{MN}^l), \tag{9}$$

$$\mathbf{x}^u = (\mathbf{x}_{11}^u, \cdots, \mathbf{x}_{M1}^u, \cdots, \mathbf{x}_{1j}^u, \cdots, \mathbf{x}_{Mj}^u, \cdots, \mathbf{x}_{1n}^u, \cdots, \mathbf{x}_{Mn}^u). \tag{10}$$

where $\mathbf{x}_{ij}^l$ and $\mathbf{x}_{ij}^u$ are the RSS streams of link i under location j with link length l and u, and the distributions of $\mathbf{x}_{ij}^l$ and $\mathbf{x}_{ij}^u$ are different. Our goal is to find a transfer matrix W, with which one can project $\mathbf{x}_{ij}^l$ and $\mathbf{x}_{ij}^u$ into a subspace, where with the same link length, i) the RSS distribution distances between different links but under the same location are maximized, $p(\mathbf{x}_{ij}) \ne p(\mathbf{x}_{i'j})$, $i \ne i' \in [1, M]$; ii) distribution distances between same link but under different locations are maximized,

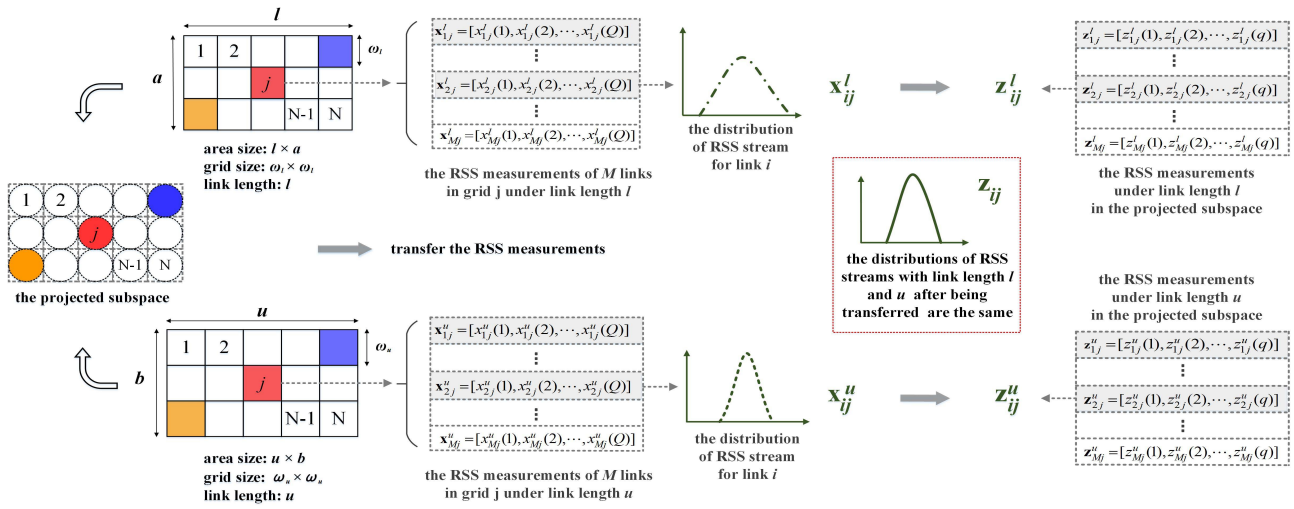


Fig. 3: **Illustration of Transfer Scheme**. For two different monitoring areas with sizes $l \times a$ and $u \times b$, the link lengths are l and u, respectively. We divide each area into N location grids, and when the target locates in grid j, he will distort some of M links, and for the same link i, the distributions of Q continuous RSS with link length l and u are different. By transfer the RSS streams $\mathbf{x}_{ij}^{l}$ and $\mathbf{x}_{ij}^{u}$ into a projected subspace $\mathbb{R}^q$ as $\mathbf{z}_{ij}^{l}$ and $\mathbf{z}_{ij}^{u}$, the distributions are made as close as possible, *i.e.*, the RSS stream distribution distances for the same locations are minimized. Afterwards, the 3 dimensional sensing matrix of area $l \times a$ after being transferred as $\mathrm{Z}_{M\times N\times q}$ can be applied to perform localization in area $u \times b$.

$p(\mathbf{x}_{ij}) \neq p(\mathbf{x}_{ij'})$, $j \neq j' \in [1, N]$; and with different link lengths, the RSS distribution distances between the same links and locations are minimized, $p(\mathbf{x}_{ij}^{l}) = p(\mathbf{x}_{i'j}^{u})$, $i = i', j = j'$.

Let $\mathbf{x}^l$ and $\mathbf{x}^u$ after being transferred are $\mathbf{z}^l = W^T\mathbf{x}^l$ and $\mathbf{z}^u = W^T\mathbf{x}^u$, the distributions of $\mathbf{z}^l$ and $\mathbf{z}^u$ are $p_l(\mathbf{z})$ and $p_u(\mathbf{z})$, respectively, where $\mathbf{z} = (\mathbf{z}^l, \mathbf{z}^u)$, $\mathbf{x} = (\mathbf{x}^l, \mathbf{x}^u)$. Then the general framework can be denoted as

$$\mathbf{z} = W^T\mathbf{x}, \tag{11}$$

$$W = \underset{W\in\mathbb{R}^{Q\times q}}{\arg\min}\{F(W) + \lambda D_W(p_l||p_u)\}, \tag{12}$$

with respect to $W^TW = I$. $F(W) + \lambda D_W(p_l||p_u)$ is the merit function, $F(W)$ is the FLDA subspace learning method which projects the RSS streams of $\mathbf{x}^l$ or $\mathbf{x}^u$ into the subspace $\mathbb{R}^q$, $q \leq Q$, where the data classification loss are minimized, $D_W(p_l||p_u)$ is the Bregman Divergence regularization that measures the distances between $p_l(\mathbf{z})$ and $p_u(\mathbf{z})$, λ is the regularization parameter that controls the trade-off between $F(W)$ and $D_W(p_l||p_u)$. In fact, FLDA gives a good initial solution to the regularization term.

B. Obtain the transfer matrix

We introduce the solving process of the transfer scheme to obtain the transfer matrix W. There are many optimization algorithms to solve the transfer matrix, and we choose the simplest Gradient Descent algorithm [21] to obtain the optimal solution iteratively. By taking steps proportional to the negative of the gradient of the function at the current point, we have

$$W_{k+1} = W_k - \eta(k)\Big(\frac{\partial F(W)}{\partial W} + \lambda \cdot \frac{\partial D_W(p_l||p_u)}{\partial W}\Big), \tag{13}$$

where η_k is the learning rate for the k-th iteration which controls the gradient step size, and we set $\eta_k = \eta_0/k$ to achieve an effective iteration. Then W can be obtained when the derivative of $\partial F(W)$ and $D_W(p_l||p_u)$ are known to us.

First, we compute $\partial F(W)$ in (13). FLDA aims to find a subspace where RSS streams in the high-dimensional space $\mathbb{R}^Q$ with different classes are separated as far as possible, while the streams in the same class are compressed as compact as possible. Here the class labels are decided by the locations. That is to say, the RSS streams for the same location but with different link length belong to the same class, and RSS streams of different locations even with the same link length belong to different classes. Let S_B be the between-class covariance matrix which denotes the separation between different streams, and S_W be the within-class covariance matrix which denotes the separation between RSS streams around their respective stream center,

$$S_W = \sum_{i=1}^{MN}\sum_{j=1}^{n_i}(\mathbf{x}_j - \mathbf{m}_i)(\mathbf{x}_j - \mathbf{m}_i)^T \tag{14}$$

$$S_B = \sum_{i=1}^{MN} n_i(\mathbf{m}_i - \bar{\mathbf{m}})(\mathbf{m}_i - \bar{\mathbf{m}})^T \tag{15}$$

$$\mathbf{m}_i = \frac{1}{n_i}\sum_{j=1}^{n_i}\mathbf{x}_j, \quad \bar{\mathbf{m}} = \frac{1}{MNn_i}\sum_{j=1}^{MNn_i}\mathbf{x}_j, \tag{16}$$

here MN is the class number and n_i is the number of RSS streams in the i-th class, in addition $\mathbf{m}_i$ and $\bar{\mathbf{m}}$ are the mean of n_i RSS streams and mean of all RSS streams, respectively. FLDA maximizes the trace ratio between S_B and S_W, and $F(W)$ becomes

$$F(W) = tr(W^TS_BW)^{-1}tr(W^TS_WW), \tag{17}$$

and its derivative with respect to W is

$$\frac{\partial F(W)}{\partial W} = 2tr(W^TS_BW)^{-1}tr(S_WW) - 2tr[(W^TS_BW)^{-1}]^2tr(W^TS_WW)tr(S_BW). \tag{18}$$

Second, to give a computational tractable realization of $D_W(p_l||p_u)$, we choose the simplest convex function $\Phi(\mathbf{z}) = \mathbf{z}^2$, and the Bregman Divergence which measures the distance between distributions $p_l(\mathbf{z})$ and $p_u(\mathbf{z})$ is

$$D_W(p_l||p_u) = \int p_l^2(\mathbf{z}) - 2p_l(\mathbf{z})\,p_u(\mathbf{z}) + p_u^2(\mathbf{z})\,d\mathbf{z}. \tag{19}$$

To take the noisiness of individual samples into account and mitigates the sample variations caused by the environment, we

use the Kernel Density Estimation (KDE) [21] which estimates the probability function as a sum of kernels between the variable and each of the other samples, and the probability functions $p_l(\mathbf{z})$ and $p_u(\mathbf{z})$ are

$$p_l(\mathbf{z}) = \frac{1}{MN \cdot \sigma_l} \sum_{i=1}^{MN} G_{\sum_l}\left(\frac{\mathbf{z}-\mathbf{z}_i^l}{\sigma_l}\right), \tag{20}$$

$$p_u(\mathbf{z}) = \frac{1}{Mn \cdot \sigma_u} \sum_{i=1}^{Mn} G_{\sum_u}\left(\frac{\mathbf{z}-\mathbf{z}_i^u}{\sigma_u}\right), \tag{21}$$

where $\mathbf{z}_i^l$ and $\mathbf{z}_i^u$ are samples of variable $\mathbf{z}^l$ and $\mathbf{z}^u$, MN and Mn are the variable lengths, σ_l and σ_u are the kernel widths, $G_{\sum_l}$ and $G_{\sum_u}$ are the Gaussian kernel functions with the covariance matrix $\sum_l$ and $\sum_u$. Here we use the Gaussian kernel function since most of the RSS streams fit the log-normal distributions [9]. Then (19) becomes

$$\begin{aligned} D_W(p_l||p_u) = & \frac{1}{M^2N^2\sigma_l^2} \sum_{i=1}^{MN}\sum_{i'=1}^{MN} G_{2\sigma_l^2\sum_l}(\mathbf{z}_{i'}^l - \mathbf{z}_i^l) \\ & + \frac{1}{M^2n^2\sigma_u^2} \sum_{i=1}^{Mn}\sum_{i'=1}^{Mn} G_{2\sigma_u^2\sum_u}(\mathbf{z}_{i'}^u - \mathbf{z}_i^u) \\ & - \frac{2}{M^2Nn\sigma_u\sigma_l} \sum_{i=1}^{MN}\sum_{i'=1}^{Mn} G_{\sigma_l^2\sum_l+\sigma_u^2\sum_u}(\mathbf{z}_{i'}^u - \mathbf{z}_i^l). \end{aligned} \tag{22}$$

Note that $D_W(p_l||p_u)$ is related to the transferred RSS streams $\mathbf{z}_i^l$ and $\mathbf{z}_i^u$. For (22), we have

$$\frac{\partial}{\partial \mathbf{z}_i^l} G_{\sum_l+\sum_u}(\mathbf{z}_i^l - \mathbf{z}_i^u) = (\mathbf{z}_i^u - \mathbf{z}_i^l)(\textstyle\sum_l+\sum_u)^{-1} G_{\sum_l+\sum_u}(\mathbf{z}_i^l - \mathbf{z}_i^u),$$

And the derivative of $D_W(p_l||p_u)$ can be obtained as follows

$$\begin{aligned} \frac{\partial D_W(p_l||p_u)}{\partial W} &= \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}} \cdot \frac{\partial \mathbf{z}}{\partial W} \\ &= \sum_{i=1}^{MN} \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^l} \cdot \frac{\partial \mathbf{z}_i^l}{\partial W} + \sum_{i=1}^{Mn} \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^u} \cdot \frac{\partial \mathbf{z}_i^u}{\partial W} \\ &= \sum_{i=1}^{MN} \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^l} \mathbf{x}_i^l + \sum_{i=1}^{Mn} \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^u} \mathbf{x}_i^u, \end{aligned} \tag{23}$$

with which

$$\begin{aligned} \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^l} &= \frac{(\sum_l)^{-1}}{M^2N^2\sigma_l^4} \sum_{i'=1}^{MN} (\mathbf{z}_{i'}^l - \mathbf{z}_i^l) G_{2\sigma_l^2\sum_l}(\mathbf{z}_i^l - \mathbf{z}_{i'}^l) \\ &- \frac{2(\sigma_l^2\sum_l + \sigma_u^2\sum_u)^{-1}}{M^2Nn\sigma_u\sigma_l} \sum_{i'=1}^{Mn} (\mathbf{z}_{i'}^u - \mathbf{z}_i^l) G_{\sigma_l^2\sum_l+\sigma_u^2\sum_u}(\mathbf{z}_i^l - \mathbf{z}_{i'}^u), \\ \frac{\partial D_W(p_l||p_u)}{\partial \mathbf{z}_i^u} &= \frac{(\sum_u)^{-1}}{M^2n^2\sigma_u^4} \sum_{i'=1}^{Mn} (\mathbf{z}_{i'}^u - \mathbf{z}_i^u) G_{2\sigma_u^2\sum_u}(\mathbf{z}_i^u - \mathbf{z}_{i'}^u) \\ &- \frac{2(\sigma_l^2\sum_l + \sigma_u^2\sum_u)^{-1}}{M^2Nn\sigma_u\sigma_l} \sum_{i'=1}^{MN} (\mathbf{z}_i^u - \mathbf{z}_{i'}^l) G_{\sigma_l^2\sum_l+\sigma_u^2\sum_u}(\mathbf{z}_{i'}^l - \mathbf{z}_i^u). \end{aligned}$$

The kernel width σ_l and σ_u are unknown and critical to solving (22), and we test the best values in our experiments (Section VI-B).

C. Grid interpolation to improve the localization accuracy

After we get the transfer matrix, the sensing matrix of area 1 and real-time RSS with unknown targets' locations of area 2 can be projected into a subspace, then the targets' locations in area 2 can be estimated by solving (8). Owe to the fact that grid size of area 2 is bigger than area 1, the localization accuracy in area 2 will decrease.

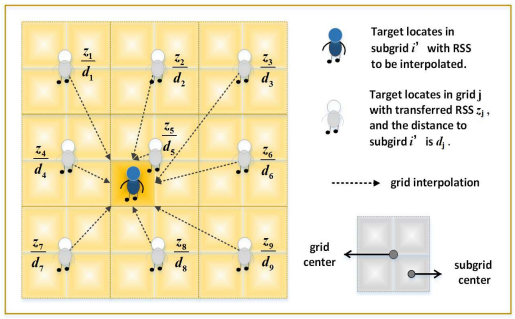


Fig. 4: **Distance based grid interpolation**. It shows the interpolation case when the subgrid is in the middle region of the monitoring area. The distorted RSS with target locates in this subgrid is interpolated by RSS with target locates in its 9 neighbor grids.

To deal with this problem, we divide the grid into subgrids and use the interpolation method to generate the RSS of subgrids. However, traditional methods, such as Newton or Hermite interpolation, don't consider the inherent relations between RSS of different locations, having a limited accuracy. As a matter of fact, RSS fingerprints are similar when the target is at neighbor locations. Therefore, by using RSS of neighbor locations, we propose a distance based gird interpolation method to generate the RSS of subgrids.

According to Section IV-A, the number of grids is fixed to be N when the link length changes from l to u, and the gird edge length is ω_l and ω_u, respectively. The size of the j-th grid in two areas[4] satisfies $w_l^2/w_u^2 = l^2/u^2$. In order to increase the number of grids in area 2, we divide each grid into $\lceil u/l \rceil^2$ subgrids with edge length $\frac{\omega_u}{\lceil u/l \rceil}$, and the total number of subgrids in area 2 is $N\lceil u/l \rceil^2$.

For grid i in area 2, we choose its 8 neighboring grids. Fig.4 shows that each grid is divided into four subgrids. When the grid i locates at the middle of the area, we choose the grids which locate at the square region around grid i as the neighbors. Then RSS of subgrid i' can be generated by

$$z_{i'} = \sum_{j=1}^{9} \frac{z_j}{d_j D}, \tag{24}$$

$$D = \sum_{j=1}^{9} \frac{1}{d_j}, \tag{25}$$

where $i' \in [1, \lceil u/l \rceil^2]$, $z_{i'}$ and z_j are the RSS of subgrid i' and grid j, d_j is the Euclidean distance between i' and j. After that, the localization accuracy of area 2 can be improved.

V. Theocratical Analysis of FitLoc

In this section, we prove that the transferred sensing matrix satisfies the RIP , which is the sufficient condition to enable CS based DfL. Particularly, we present a theorem that the transferred sensing matrix Z satisfies RIP with high probability tending to 1, and this ensures high recovery accuracy.

Theorem 1: If the RSS of each row in the transferred sensing matrix Z is subject to Gaussian distribution and $M = O(K\log(N/K))$, then the probability that Z satisfies $(1-\delta) \le \frac{\|Z\Theta\|_2^2}{\|\Theta\|_2^2} \le (1+\delta)$ for all N-dimensional K-sparse vector Θ tends to 1, where $\delta \in (0,1)$.

Proof: It is easy to verify that each row of the transferred sensing matrix Z follows the Gaussian distribution [9]. After

[4]We suppose $l/u = a/b$ to simplify the size scale of the location grid.

that, we normalize Z as $\sqrt{\frac{1}{\sigma M}}[\mathrm{Z}_1, \mathrm{Z}_2, \cdots, \mathrm{Z}_M]^T$ to simplify the proof, and assume $\mathrm{E}(\mathrm{Z}_{ij}) = \mu$, $\mathrm{Var}(\mathrm{Z}_{ij}) = \mathrm{E}((\mathrm{Z}_{ij})^2) = \sigma$. And the mean and variance of the product $\langle \sqrt{\frac{1}{\sigma M}}\mathrm{Z}_i, \Theta \rangle$ are

$$\mathrm{E}\big(\langle \frac{1}{\sqrt{\sigma M}}\mathrm{Z}_i, \Theta \rangle\big) = \frac{1}{\sqrt{\sigma M}} \sum\nolimits_{j=1}^{N} \mathrm{E}(\mathrm{Z}_{ij})\theta_j = \frac{K\mu}{\sqrt{\sigma M}}, \tag{26}$$

$$\mathrm{Var}\big(\langle \frac{1}{\sqrt{\sigma M}}\mathrm{Z}_i, \Theta \rangle\big) = \frac{1}{\sigma M} \sum\nolimits_{j=1}^{N} \mathrm{Var}(\mathrm{Z}_{ij})\theta_j^2 = \frac{\|\Theta\|_2^2}{M}. \tag{27}$$

Then we can obtain the mean of $\|\mathrm{Z}\Theta\|_2^2$ as

$$\mathrm{E}\big(|\mathrm{Z}\Theta\|_2^2\big) = \sum\nolimits_{i=1}^{M} \mathrm{Var}\big(\langle \sqrt{\frac{1}{\sigma M}}\mathrm{Z}_i, \Theta \rangle\big) = \|\Theta\|_2^2. \tag{28}$$

Based on [23], $p\Big(\big|\frac{\|\mathrm{Z}\Theta\|_2^2}{\|\Theta\|_2^2} - 1\big| \geq \delta\Big) \leq 2\exp\big(\frac{-M\delta^2}{c}\big)$, where c is a constant number. Then for $C_N^K \leq \big(\frac{eN}{K}\big)^K$ possible K-dimensional subspace of Z, the probability of K-sparse Θ which satisfies $\big|\frac{\|\mathrm{Z}\Theta\|_2^2}{\|\Theta\|_2^2} - 1\big| \geq \delta$ is

$$\big(\frac{eN}{K}\big)^K \cdot 2\exp(\frac{-M\delta^2}{c}) = 2\exp\big(\frac{-M\delta^2}{c} + K\log\big(\frac{N}{K}\big) + 1\big). \tag{29}$$

With this and $M = O(K\log(N/K))$, the probability that Z satisfies $(1-\delta) \leq \frac{\|\mathrm{Z}\Theta\|_2^2}{\|\Theta\|_2^2} \leq (1+\delta)$ for all N-dimensional K-sparse vector Θ tends to 1. ■

Based on the analysis in Theorem 1, the targets' locations can be recovered by the CS theory as the transferred sensing matrix satisfies the RIP with high probability.

VI. Implementation and Experimental Results

A. Experimental setup

We attempt to perform localization when the areas are 6m×6m and 12m×12m under the condition that knowing the sensing matrix of area 3m×3m and 4m×4m, then the transfer work includes three cases: doubling the link length from l=3m to u=6m, tripling the link length from l=4m to u=12m, and quadrupling the link length from l=3m to u=12m. We add our transfer scheme into two state-of-the-art algorithms RASS [5] and RTI [13], referred as RASS w/ Trans. and RTI w/ Trans. for a fair comparison. And we also compare with the traditional CS based localization method with the sensing matrix of link length u as ground truth, referred as CS w/o Trans. The performance of FitLoc are investigated by considering the following parameters: (i) K: the number of targets, (ii) M: the number of chosen links, (iii) n: the number of chosen locations in the new area, (iv) λ: the regularization parameter, (v) σ: the kernel width, and (vi) $\eta(0)$: the initial learning rate, and the default values are shown in Table I.

We perform extensive experiments in an open-space depicted in Fig. 13. Based on the work in [9], we set the grid edge length as ω=0.5m when the link length is 4m[5], thus the grid number is 64. We use the MICAZ [24] nodes that work on 2.4G as the transceivers, and put them on the height of 0.95m to obtain the best propagation property [9], each link records 100 measurements.

[5]In fact, when the gird size is other values, the performance of FitLoc are consistent through our extensive experiments. To save the space, we only report the experiments with ω=0.5m.

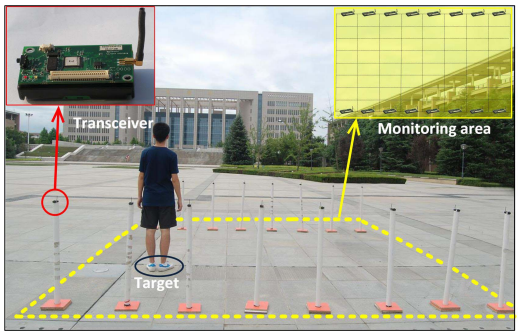


Fig. 13: Experiment scene with area size of 4m×4m.

TABLE I: DEFAULT VALUES OF PARAMETERS.

Parameters	Default Values
The number of targets K	2
The number of chosen links M	4
The number of chosen locations n	4
The kernel width σ	3.5dBm
The regularization parameter λ	0.5
The initial learning rate $\eta(0)$	10
The transfer case	from 4m to 12m

B. Impact of parameters

We discuss the parameters K, M, n, σ, λ and $\eta(0)$ under the default transfer case, and the values are the same to the other two transfer cases.

Target Number K: we discuss how many targets can be accurately localized by FitLoc, and the other parameters are set as the default values. When the targets number increases from 1 to 6, Fig. 5 shows that FitLoc can localize 3 targets with the average localization error of 0.89m, and the localization error of RTI w/ Trans. and RASS w/ Trans. are much bigger than FitLoc. According to the CS theory, FitLoc satisfies $M > K(log(N/K)) = 3.98$, thus FitLoc can accurately localize multiple targets and performs better than RTI w/ Trans. and RASS w/ Trans. algorithms. It also illustrates that RTI w/o Trans. and RASS w/o Trans. performs worse when comparing with RTI w/ Trans. and RASS w/ Trans., which demonstrates the effectiveness of our transfer scheme.

Link Number M: we inspect the number of links used in localization and increase the link number M from 1 to 6. Fig. 6 shows the average localization error under different M. It can be seen that the localization errors of RTI and RASS without transfer (RTI w/o Trans. and RASS w/o Trans.) are large and random. On the contrary, the localization errors of FitLoc, RTI and RASS with transfer (RTI w/ Trans. and RASS w/ Trans.) decrease significantly when the link number increases. FitLoc performs almost as well as the method which directly uses the CS algorithm with the ground truth sensing matrix (CS w/o Trans.), illustrating the feasibility and effectiveness of the transfer scheme. In addition, as $M = 4$ is a sparse deployment, CS w/o Trans. and FitLoc satisfy $M = 4 > K(log(N/K)) \approx 3$, so they outperform RTI w/Trans. and RASS w/Trans. which need a dense deployment to collect enough data.

Location Number n: we attempt to test the best location number n to solve the transfer matrix, and the average localization errors under different number of randomly chosen locations are shown in Fig. 7. When n increases from 0

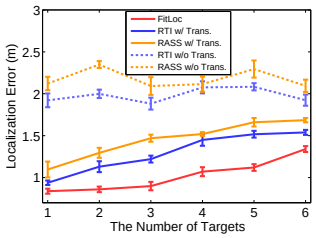


Fig. 5: Impact of target number.

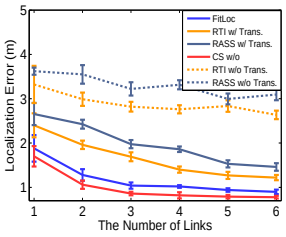


Fig. 6: Impact of link number.

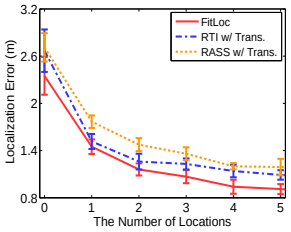


Fig. 7: Impact of location number.

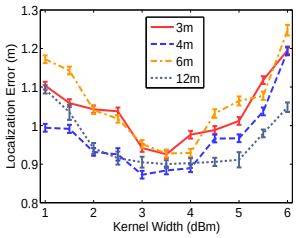


Fig. 8: Impact of kernel width.

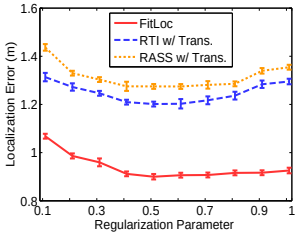


Fig. 9: Regularization Parameter.

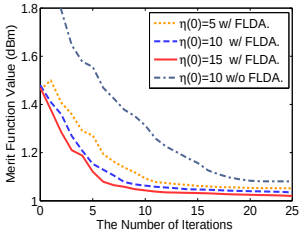


Fig. 10: Performance of learning rate.

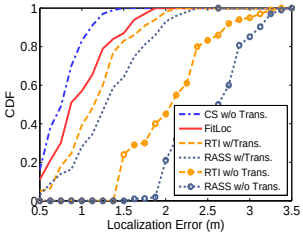


Fig. 11: Outdoor performance.

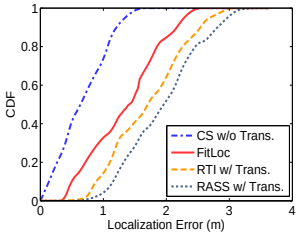


Fig. 12: Indoor performance.

to 1, the localization error for all three methods decreases dramatically, which suggest that the localization accuracy can be greatly improved by randomly choosing a location within the effective area of a link. And as the location number increases continuously, it is easily to understand that the big location number is, the more precise of the transfer matrix is, and the lower the localization error will be. We set n=4 as the most appropriate location number since it will increase the collection effort if we use too many locations.

Kernel Width σ and Regularization Parameter λ: we investigate the most appropriate kernel width which used for estimating the RSS distributions under link length 3m, 4m, 6m and 12m. The average localization errors under different kernel width are illustrated in Fig. 8. It shows that the appropriate kernel widths for link length 3m, 4m and 6m and 12m are between 3dBm and 4dBm. And we find that the most appropriate kernel widths are almost consistent with the RSS noise, and we set σ=3.5. We also inspect the localization performance under different regularization parameter λ. Fig. 9 shows that the most appropriate value of λ is 0.5. It is not difficult to understand that when $\lambda = 0.5$, the regularization part is the same important to the FLDA algorithm and makes the transfer scheme most effective.

Initial Learning Rate $\eta(0)$: we study the critical parameter $\eta(0)$ that used in the gradient descent to achieve an optimal solution for the transfer matrix, which with initial solutions generated by the FLDA algorithm, referred as w/ FLDA with different $\eta(0)$ in Fig. 10. It can be seen that the values of the merit function $F(W)+D_W(p_l||p_u)$ converge to the minimum under three different $\eta(0)$ after about 20 iterations, and the convergence rate have little difference from each other. We set $\eta(0) = 10$ as the default value since the convergence curve is more smooth. What's more, in order to study the efficacy of the FLDA algorithm, we directly use the gradient descent to solve the transfer matrix, referred as $\eta(0) = 10$ w/o FLDA. The convergence curve falls more slowly and more iterations are needed when compared with $\eta(0) = 10$ w/ FLDA., which illustrates that the FLDA algorithm provides a good initialization to $D_W(p_l||p_u)$.

C. Effectiveness of FitLoc

We attempt to discuss the effectiveness of FitLoc from the following three aspects: localization performance, human effort cost, and energy consumption.

Localization Performance. Fig. 11 illustrates the CDF of localization error under the transfer from 4m to 12m in outdoor environment. It shows that the performance of FitLoc is approximate to the CS w/o Trans., and the localization errors for three transfer based methods do not change obviously. FitLoc performs best with 50^{th} and 80^{th} percentile error of 0.89m and 1.23m, respectively. When compared with RTI w/o Trans. and RASS w/o Trans., RTI w/ Trans. and RASS w/ Trans. improve 58% and 66% for 80^{th} percentile error, respectively, which shows that the transfer scheme significantly improves the performance of RTI and RASS algorithms. In addition, under the transfer from 3m to 6m and 3m to 12m, the localization error for RTI and RASS respectively improves 54%, 59% and 76%, 85% with the transfer scheme. This indicates that after transfer, the pre-obtained sensing matrix can be adaptively reused in the new area.

On the other hand, when the deployment areas are in-door environment with the similar presence of multipath (i.e., the areas are similar furnitured rooms), the localization performance of FitLoc degrades and is showed in Fig.12. It can be seen that the 50^{th} localization error of FitLoc is 1.4m and 52% higher than the CS w/o Trans. algorithm, which illustrates that the transfer scheme performs bad in the in-door environment, due to the fact that different indoor areas suffer from severe multipath and noisy environment. In spite of this, FitLoc performs better than RTI w/ Trans. and RASS w/ Trans., which shows that CS based localization is more robust to the environment noise.

Human Effort Cost. In general, the RSS of a new area are obtained manually. We use the time-cost of the pre-deployment to examine the human effort. The area is divided into grids with edge length 0.5m, and 100 continuous RSS changes are collected at each grid with 1.5 seconds each measurement. Thus the construction time-cost of the sensing matrix for area 4m×4m and 12m×12m are at least $\frac{100\times1.5\times(4/0.5)^2}{3600} \approx 2.67$

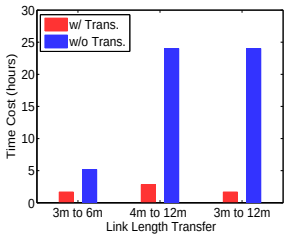


Fig. 14: Time cost.

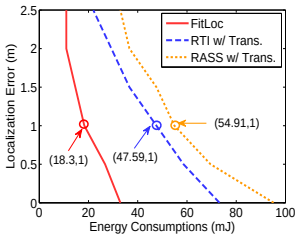


Fig. 15: Energy consumption.

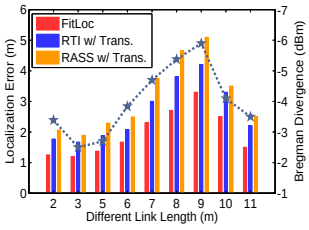


Fig. 16: Robustness.

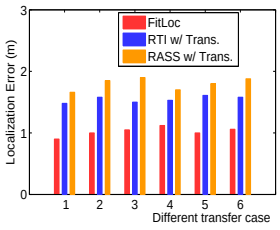


Fig. 17: Scalability.

and $\frac{100\times1.5\times(12/0.5)^2}{3600} = 24$ man-hours, respectively. And when we use the transfer scheme, in addition to the time-cost for the 4m×4m area, only 4 grids' human effort are needed for the 12m×12m area, thus the time-cost is $\frac{100\times1.5\times(4/0.5+4)}{3600} \approx 2.83$. Fig. 14 shows the time-cost of three transfer cases with and without our transfer scheme, and the human effort decreases are 41% from 3m to 6m, 88% from 4m to 12m, 93% from 3m to 12m.

Energy Consumption. We compare the energy consumption of FitLoc, RASS w/ Trans. and RTI w/ Trans. by increasing the number of links until the localization accuracy reaches the given value, and then calculate the energy consumption. According to the first order radio model [25], the energy consumption for each packet of the link is calculated by $E_{radio} = e_l B b^2 + 2BE_{elc}$, where B is the size of a packet in bits, b is the link length, e_l=100pJ/(bit/m^2) and E_{elc}=50nJ/bit. We set B=320bits, b=12m and send 100 packets each time. Thus, the energy consumption for one algorithm with M links is $M\times$3.66mJ. Note that for each algorithm, the number of links required for achieving the same localization error is different. Fig. 15 shows the energy consumptions under different localization errors. When the given localization error is 1m, the energy consumption for FitLoc, RTI w/ Trans. and RASS w/ Trans. are 18.3mJ, 47.59mJ, and 54.91mJ, respectively. It illustrates that RTI and RASS require more measurements for an accurate localization, while FitLoc can accurately localize targets even with a small number of measurements, highly reducing the energy consumption.

D. Robustness and Scalability of FitLoc

Robustness. In order to investigate the robustness of FitLoc, we evaluate the localization performance when the link length of a new area is not used for modeling the transfer function. A new link length means a new Bregman Divergence over the existing link lengths, and how do the Bregman Divergence differences among links affect the localization performance? By randomly choosing some locations and recording the RSS with different new link length, the average localization error and the Bregman Divergence are depicted in Fig. 16. It shows that the localization error is bound to 2.3m when the link length smaller than 6m or bigger than 10m, but become large for the other link lengths. Another observation is that when the absolute value of the Bregman Divergence larger than 4dBm, the localization error increases distinctly. In other words, the robustness of FitLoc is good when the noise lower than 4dBm.

Scalability. We also study the scalability of FitLoc when there are more than two kinds of link length in a new area, which means that the RSS to solving the transfer matrix corresponding to diverse link length in the new area. We evaluate the localization performance in the new area under 6 different transfer cases, including l = 3m to u={4m,6m}, u={6m,12m}, u={4m,6m,12m} and l=4m to u={3m,6m}, u={6m,12m}, u={3m,6m,12m}, and the results are illustrated in Fig. 17. It can be seen that the average localization error of FitLoc, RTI w/ Trans. and RASS w/ Trans. keep stable, since the transfer scheme is used for the coverage area of a specific link, the transfer scheme performs well when there are multiple link lengths in the new area.

E. Discussions

- To achieve the best localization performance, it needs to set a proper grid size in the pre-deployment phase. On one hand, a small grid size provides a fine-grained resolution of localization accuracy, and vice versa. On the other hand, a big grid size results in an unreliable RSS interpolation. In view of these, the proper grid size should be determined by the specific application accuracy requirement. And we choose a grid size of 0.5m×0.5m through the extensive experiments.
- For areas with similar indoor environments (*i.e.*, the deployment of furniture, wireless interference, etc.), the performance of FitLoc degrades to an average localization error of 1.4m from 0.89m in the outdoor environment, according to Section VI. This indicates that FitLoc is more capable of areas with relatively similar environment. Even so, we suggest that FitLoc satisfies the daily indoor localization requirement with a room-level accuracy without the need of re-calibration.
- Since the communication range of wireless transceivers is limited, we divide the large area into small subareas [5], [9]. Then when the area scales up, the cost can be greatly reduced by applying the transfer scheme to a number of subareas. Note that for the irregular deployment of transceivers with different link length in an area (it is not common in most settings actually), the transfer can be done for the effective area of each link.

VII. Related Work

DfL has received much more attention with the needless of target-attached devices [3]–[5], [7]–[14]. Compared with video-based [26] and ultrasonic-based DfL [27], one main advantage of RF based DfL is that the RSS measuring are ubiquitous in existing wireless infrastructures and without requiring additional devices. It can be generally divided into two categories. The first one is Radio Topology Imagine model based approaches [10]–[14] which have a limited performance as the model is vulnerable to environment noise. The other one is the radio-map based methods [3]–[5], [7]–[9] which could achieve a fine-grained localization accuracy by comparing the real-time RSS with the radio-map.

Most radio-map based DfL approaches require a high deployment and human effort cost to perform localization

over various areas. In this paper, we propose FitLoc, which employs the CS based DfL proposed in [9] to reduce the deployment cost and combines with the rigorously designed transfer scheme to reduce the calibration cost. Work [10]–[12] are related to FitLoc which take advantage of the CS theory, but they are model based and need dense deployment. Work [20] uses the similar transfer method to FitLoc, while it transfers the information of different domains in image processing, and doesn't consider the noise influence. To transfers RSS of one target to another, work [28] designs a linear transfer model by utilizing the Maximum Mean Discrepancy to measure the distribution distance, but it can not be used as an regularization term in FitLoc. Another related work [15] transfers the learned model from one spatial area to another for indoor WiFi localization, but it follows the simple premise that the two areas must share some common devices and the target needs to carry a wireless device. In summary, FitLoc equips the target without any wireless devices, and enables fine-grained multiple targets localization over various areas with little cost, thus more practical for deployment setups.

VIII. Conclusion

This paper presents the first fine-grained multiple targets DfL approach over various areas with little cost. By a novel transfer scheme, which projects the RSS into a subspace where the distribution distances over different areas are minimized, the radio-map of one area can be reused by various areas. Thus, the calibration effort is greatly reduced. On the other hand, based on the CS theory, FitLoc reduces the deployment cost by deploying a small number of transceivers and collecting only a few of RSS. We have also evaluated the effectiveness of FitLoc through both theoretical analyses and extensive experiments

IX. Acknowledgements

This work is partially supported by the US National Science Foundation under Grant No. CNS-1319915 and CNS-1343355, and the National Natural Science Foundation of China under Grant No. 61170218, 61272461, 61428203, 61572347, 61373177 and 61572402. Xiaojiang Chen is the corresponding author.

References

[1] Chenshu Wu, Zheng Yang, Chaowei Xiao, Chaofan Yang, Yunhao Liu, and Mingyan Liu. Static power of mobile devices: Self-updating radio maps for wireless indoor localization. In *IEEE INFOCOM*, 2015.

[2] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. Tagoram: Real-time tracking of mobile rfid tags to high precision using cots devices. In *ACM MobiCom*, pages 237–248, 2014.

[3] Moustafa Youssef, Matthew Mah, and Ashok Agrawala. Challenges: device-free passive localization for wireless environments. In *ACM MobiCom*, pages 222–229, 2007.

[4] M. Seifeldin, A. Saeed, et al. Nuzzer: A large-scale device-free passive localization system for wireless environments. *IEEE Trans. on Mobile Computing*, 12(7):1321–1334, 2013.

[5] Dian Zhang, Yunhuai Liu, et al. RASS: A real-time, accurate, and scalable system for tracking transceiver-free objects. *IEEE Trans. on Parallel and Distributed Systems*, 24(5):996–1008, 2013.

[6] Zimu Zhou, Chenshu Wu, Zheng Yang, and Yunhao Liu. Sensorless Sensing with WiFi. *ACM Computing Surveys*, 20(1):1–6, 1 2015.

[7] D. Zhang, K. Lu, R. Mao, Y. Feng, Y. Liu, Z Ming, and L. M Ni. Fine-grained localization for multiple transceiver-free objects by using RF-based technologies. *IEEE Trans. on Parallel and Distributed Systems*, 25(6):1464–1475, 2014.

[8] Chenren Xu, Bernhard Firner, Robert S Moore, Yanyong Zhang, Wade Trappe, Richard Howard, Feixiong Zhang, and Ning An. SCPL: indoor device-free multi-subject counting and localization using radio signal strength. In *IEEE IPSN*, pages 79–90, 2013.

[9] Ju Wang, Dingyi Fang, Xiaojiang Chen, Zhe Yang, Tianzhang Xing, and Lin Cai. LCS: Compressive sensing based device-free localization for multiple targets in sensor networks. In *IEEE INFOCOM*, pages 145–149, 2013.

[10] M A Kanso and M G Rabbat. Compressed RF tomography for wireless sensor networks: Centralized and decentralized approaches. In *IEEE DCOSS*, pages 173–186. 2009.

[11] J. Wang, Q. Gao, X. Zhang, and H. Wang. Device-free localisation with wireless networks based on compressive sensing. *IET Communications*, 6(15):2395–2403, 2012.

[12] J. Wang, Q. Gao, H. Wang, P. Cheng, and K. Xin. Device-free localization with multidimensional wireless link information. *IEEE Trans. on Vehicular Technology*, 64(1):356 – 366, 2015.

[13] Joey Wilson and Neal Patwari. See-through walls: Motion tracking using variance-based radio tomography networks. *IEEE Trans. on Mobile Computing*, 10(5):612–621, 2011.

[14] Joey Wilson and Neal Patwari. Radio tomographic imaging with wireless networks. *IEEE Trans. on Mobile Computing*, 9(5):621–632, 2010.

[15] S. J. Pan, D. Shen, Q. Yang, et al. Transferring localization models across space. In *AAAI*, pages 1383–1388, 2008.

[16] Yongmin Zhang, Shibo He, and Jiming Chen. Data gathering optimization by dynamic sensing and routing in rechargeable sensor networks. In *IEEE/ACM Trans. on Networking*. To appear, 2015.

[17] Xiaolong Zheng, Zhichao Cao, Jiliang Wang, Yuan He, and Yunhao Liu. Zisense: towards interference resilient duty cycling in wireless sensor networks. In *ACM Conference on Embedded Network Sensor Systems*, pages 119–133, 2014.

[18] Haifeng Zheng, Shilin Xiao, Xinbing Wang, and Xiaohua Tian. Energy and latency analysis for in-network computation with compressive sensing in wireless sensor networks. In *IEEE INFOCOM*, pages 2811–2815, 2012.

[19] Haifeng Zheng, Feng Yang, Xiaohua Tian, Xiaoying Gan, Xinbing Wang, and Shilin Xiao. Data gathering with compressive sensing in wireless sensor networks: A random walk based approach. *IEEE Trans. on Parallel and Distributed Systems*, 26(1):35–44, 2015.

[20] Si Si, Dacheng Tao, and Bo Geng. Bregman divergence-based regularization for transfer subspace learning. *IEEE Trans. on Knowledge and Data Engineering*, 22(7):929–942, 2010.

[21] Christopher M Bishop et al. *Pattern recognition and machine learning*, volume 4. springer New York, 2006.

[22] Noboru Murata, Takashi Takenouchi, Takafumi Kanamori, and Shinto Eguchi. Information geometry of u-boost and bregman divergence. *Neural Computation*, 16(7):1437–1481, 2004.

[23] M. A Davenport. *Random observations on random observations: Sparse signal acquisition and processing*. PhD thesis.

[24] TelosB Datasheet. Crossbow Inc. http://www.xbow.com.

[25] Wendi Rabiner Heinzelman, Anantha Chandrakasan, and Hari Balakrishnan. Energy-efficient communication protocol for wireless microsensor networks. In *Proc. International Conference on System Sciences, 2000*, pages 10–20, 2000.

[26] Huadong Ma, Chengbin Zeng, and Charles X Ling. A reliable people counting system via multiple cameras. *ACM Trans. on Intelligent Systems and Technology*, 3(2):31, 2012.

[27] Peng Cheng, Fan Zhang, Jiming Chen, Youxian Sun, and Xuemin Shen. A distributed tdma scheduling algorithm for target tracking in ultrasonic sensor networks. *IEEE Trans. on Industrial Electronics*, 60(9):3836–3845, 2013.

[28] Ju Wang, Xiaojiang Chen, Dingyi Fang, Chase Qishi Wu, Zhe Yang, and Tianzhang Xing. Transferring compressive-sensing-based device-free localization across target diversity. *IEEE Trans. on Industrial Electronics*, 62(4):2397–2409, 2015.