

Communicating Is Crowdsourcing: Wi-Fi Indoor Localization with CSI-based Speed Estimation

Zhiping Jiang*, Jizhong Zhao*, Xiang-Yang Li†, Wei Xi*, Kun Zhao*, Shaojie Tang‡, Jinsong Han*

*School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China

†Department of Computer Science, Illinois Institute of Technology, Chicago, IL

‡Department of Computer and Information Science, Temple University, Philadelphia, PA

Email: {jiangzp.cs, weixi.cs, tangshaojie }@gmail.com, xli@cs.iit.edu, {zjz, zhaokun, hanjinsong}@mail.xjtu.edu.cn

Abstract—Numerous indoor localization techniques have been proposed recently to meet the intensive demand for location based service, and Wi-Fi fingerprint-based approaches are the most popular and inexpensive solutions. Among them, one of the main trends is to incorporate the built-in sensors of smartphone and to exploit crowdsourcing potentials. However the noisy built-in sensors and multi-tasking limitation of underline OS often hinder the effectiveness of these schemes.

In this work, we propose a *passive crowdsourcing* CSI-based indoor localization scheme, C^2IL . Our scheme C^2IL *only* requires the locating-device (e.g., a phone) to have a 802.11n wireless connection, and it does not rely on inertial sensors only existing in some smartphones. C^2IL is built upon our innovative method to accurately estimate the moving distance purely based on 802.11n Channel State Information (CSI). Our extensive evaluations show that the moving distance estimation error of our scheme is within 3% of the actual moving distance regardless of varying speeds and environment. Relying on the accurate moving distance estimation as constraints, we are able to construct a more accurate mapping between RSS fingerprints and location. To address the challenges of collecting fingerprints, a crowdsourcing-based scheme is designed to *gradually* establish the mapping and populate the fingerprints. In C^2IL , we design a trajectory clustering-based localization algorithm to provide precise real-time indoor localization and tracking. We developed and deployed a practical working system of C^2IL in a large office environment. Extensive evaluation results indicate that our scheme C^2IL provides accurate localization with error $2m$ at 80% at very complex indoor environment with minimal overhead.

I. INTRODUCTION

With the prosperity of mobile devices, especially smartphones, location based services (LBS), which use the geographic position to provide targeted services, have become pervasive to provide added value of existing services. A critical challenge of LBS is to find the accurate location of mobile devices. GPS has successfully dominated the outdoor localization. Unfortunately in indoor environment, the most facile wireless received signal strength (RSS) is neither accurate nor consistent due to the highly dynamic and complex environment. As the flourishing of smartphones and crowdsourcing computation models, numerous indoor localization techniques have been proposed to collect fractional environment features and collaboratively provide precise indoor localization.

The relatively good accuracy and simplicity of fingerprint based localization schemes has attracted massive of effort in the community. Wi-Fi fingerprint-based schemes can provide meter-level indoor localization accuracy at the expense of

explicit site-survey. Its high deployment cost and low adaptiveness to environment change hinders the practical effectiveness. Recently several novel techniques, e.g., [18], [21], [23], [25], [30], [31], have been proposed to raise the usability and accuracy. Among these approaches, a hot research trend is to incorporate crowdsourcing model and built-in sensors in today's smartphone. LiFS [31] reduces the site-survey by using the moving distance, estimated from counting the number of steps by accelerometer, as constraints for matching between the map and trace-graph, achieving average accuracy of $5.8m$. Zee [23] achieves a mean accuracy $3m$ by estimating the moving direction and moving distance by similarly leveraging the sensors in smartphone. Centuar [21] and PAL [18] both calibrate the Wi-Fi fingerprints database using acoustic ranging and they achieve 1 to $3m$ accuracy. All these approaches significantly improved the practicability of Wi-Fi based indoor localization, however, we believe there are plenty of room for improving the localization accuracy while reducing or even eliminating the dependence of site-survey and noisy inertial sensors.

In this work, we design and develop an indoor localization scheme C^2IL with even lower cost and hindrance. Our scheme exploits the channel state information (CSI) in 802.11n, for better distance estimation which in turn results in a better fingerprints matching and localization. Additionally, C^2IL , as in most recent techniques, collects the WiFi RSS fingerprints during communication between the user and WiFi APs, and populates the fingerprints database by crowdsourcing. C^2IL performs extremely well in a complex environment with rich multipath effect, while several recently developed schemes suffer from lower localization accuracy in such a complex environment.

The accurate moving distance estimation is based on the readily available CSI in IEEE 802.11n systems that use OFDM MIMO technology. Since CSI focuses more on small-scale fading, it has extraordinary advantages on capturing the mobile channel characteristics. From our preliminary testing on various scenarios when users move with constant speeds in a complex environment with rich multipath features, we find ripples-like Rayleigh deep fadings [24] with some periodicity across all subcarriers regardless of antenna configurations as shown in Fig. 1 (a) and (b). A natural question is “Are these ripples correlated with distance or speed?”.

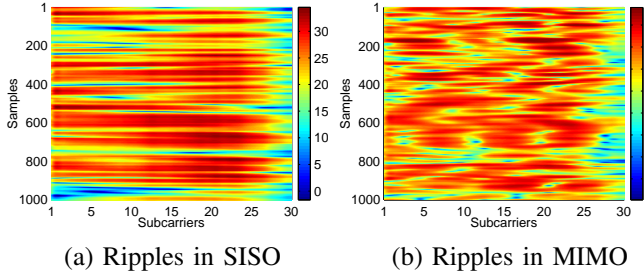


Fig. 1. (a) Ripples across nearly all subcarriers in SISO (b) The ripples in 3×3 MIMO configuration with MCS=16.

The answer is beyond a simple *yes*. In this work, we first exhibit a simple yet precise relationship between moving speed v and the frequency of ripples-like fadings f_r . Based on this important discovery, we for the first time propose a simple method to precisely estimate the moving speed (and moving distance) of a wireless device purely based on wireless traffic between the device and the WiFi APs. Since in Wi-Fi environment both AP and clients can measure the CSI and RSS, the client's moving speed and RSS can be measured remotely at AP-end *passively* and *silently*. In such a case, all the connected client devices have already become participants of the crowdsourcing system without any effort. When there is communication, they are formally contributing to the crowdsourcing, so *communication is crowdsourcing* in $\mathbf{C}^2\mathbf{IL}$.

After computing a precise speed and moving distance estimation, our localization scheme $\mathbf{C}^2\mathbf{IL}$ then applies the graph matching (GM) for several core tasks: fingerprints extraction, automatic mapping between fingerprints and floor plan, and user localization/tracking. It has been widely known that the RSS value is affected by many factors, *e.g.*, the RSS values collected at the same location using same devices with same WiFi APs could fluctuate to a few db depending on how users hold and block the signal [25]. Such fluctuation will significantly impact the fingerprint matching quality and thus impact the localization/tracking accuracy. We first identify and successfully address the directional shadowing problem for conventional threshold-based fingerprints extraction, and we propose a trajectory matching-based solution to eliminate the shadowing problem. For of fingerprints mapping task, $\mathbf{C}^2\mathbf{IL}$ supports unsupervised large scale complex indoor floor plan mapping. For localization/tracking task, a combination of trajectories matching and particle filter is proposed to achieve precise indoor localization/tracking. In summary, the contributions of $\mathbf{C}^2\mathbf{IL}$ are as follows.

- Compared with the previous approaches with good accuracy, our scheme $\mathbf{C}^2\mathbf{IL}$ does not require the localization devices to be smartphones with various inertial sensors (which are required in [23], [31]). As it only requires 802.11n, $\mathbf{C}^2\mathbf{IL}$ has the lowest barrier on devices, thus, we expect it to have better contribution from users, which is of vital importance for practical and continuously-functioning crowdsourcing.
- $\mathbf{C}^2\mathbf{IL}$ is the first practical localization system that really benefits from the multipath effect, instead of suffering

from this notoriously challenging effect. Theoretical analysis showed that it is exactly the multipath effect that enables the accurate distance estimation by CSI. Our extensive experimental evaluation indicated that in typical indoor scenarios, the distance estimation error is often within 3% regardless of moving speeds, which is much more accurate than Dead-Reckoning or pedometer based approaches.

- The adoption of graph matching and other techniques in our core design guarantees the accuracy and scalability of RSS map building and localization in very large and complex environment. We design and develop a prototype of $\mathbf{C}^2\mathbf{IL}$ in a large office environment of about $2000m^2$ with complex structure. In our extensive tests, the localization error without any historical data is within $5m$; while during tracking, thanks to the precise moving distance estimation by CSI, the tracking error could be within $1m$.

The rest of the paper is organized as follows. We review related techniques in Section II, present $\mathbf{C}^2\mathbf{IL}$ system overview in Section III and our innovative distance estimation in Section IV. Fingerprint extraction and complete crowdsourcing based localization scheme are introduced in Section V and Section VI respectively. We present the fine-grained indoor tracking in Section VII, report our extensive performance evaluation of $\mathbf{C}^2\mathbf{IL}$ in Section VIII. We conclude the paper in Section IX.

II. RELATED WORK

A. Indoor Localization Schemes

The localization problem are applied in two main scenarios: outdoor and indoor. The most popular outdoor localization method is GPS [19]. Other existing techniques of both outdoor and indoor localization mainly fall into two categories: Fingerprint-based and Modeling-based. Fingerprints are utilized in many literatures to assist positioning, and the most widely used one is WiFi signal. In indoor environment, fingerprint based methods (*e.g.*, Radar [4], Horus [32], Surround-Sense [3], PinLoc [26],) first collect fingerprint of WiFi signal (or cellular, or FM, or other sensors such as light) in advance at known locations inside a building, and then identify the user's location by matching the fingerprint of this user with the fingerprint stored in database. Dead-Reckoning is another stream of techniques (*e.g.*, [10], [12]) proposed in the literature for localization.

The most used fingerprint is the RSS value. LiFS [31] proposed a crowdsourcing based indoor localization, which exploits the possibility of automatically establishing the mapping between fingerprint set F and position set P . Acoustic ranging (AR) assisted Wi-Fi positioning was recently developed to provide distance estimation between two users (*e.g.*, [18], [21]). These schemes leverage the accurate AR and are able to provide high localization accuracy using the mapping of fingerprints with some additional distance constraint.

CSI has potential for accurate indoor localization since the CSI tool [13] has been released to public on off-the-shelf hardware. CSI is not a simple extension of RSS on physical subcarriers but it reveals totally different information

on frequency selective fading process. SpinLoc [25] system proposed a rotation based indoor localization system that leveraged the human bodies' strong fading to Line-of-Sight (LoS) components. PinLoc [27] proposed a CSI fingerprint-based localization system which can achieve meter-level precise indoor point localization. FILA [30] proposed a precise indoor ranging system by eliminating the non-LoS components in CSI information. Compared with these schemes, our scheme **C²IL** provides an accurate moving distance estimation of a single user in a complex indoor environment.

B. Estimate Moving Distance by CFR

There is a long history of estimating moving velocity of a mobile station according to wireless signal [2], [20]. Most of them focused on fast moving stations, *i.e.*, a mobile station in cars or trains. However, the algorithm for estimating the maximum Doppler frequency f_d , on which most of these methods based, is not suitable for estimating human walking speed. The maximum Doppler frequency in 2.4G or 5G Wi-Fi environment could be almost totally ignored, *i.e.*, for a moving station with a velocity $1.5m/s$, f_d is merely 12Hz/25Hz compared to the carrier frequency 2.4Ghz or 5.2Ghz.

To the best of our knowledge, [22] is the only previous work that implemented an indoor speed estimation system, which is based on DVB-T signal working at 746Mhz. This work used the relationship $v = \xi \frac{\lambda}{T_c}$ to estimate velocity, where $\xi = 0.423$ a pre-defined constant, λ the wavelength, and T_c is the channel coherent time. However, for Wi-Fi signal with small covering range, which causes non-uniformly Rayleigh fading, a constant ξ is not appropriated, and it is extremely difficult and challenging to precisely estimate the ξ in a dynamic indoor environment.

C. Graph Matching via Relaxation

Graph matching (GM) is a widely used technique to find the best correspondence between two graphs, documents, or images. In **C²IL**, we use GM to extract RSS fingerprints and automatically establish the mapping between RSS fingerprints and floor plan.

GM is essentially an integer quadratic programming (IQP) problem, and it is NP-hard [8]. Given two graphs $G^P = (\mathcal{V}^P, \mathcal{E}^P)$ and $G^Q = (\mathcal{V}^Q, \mathcal{E}^Q)$, the goal of GM is to find the best correspondence between two graphs. Let $C^{P \times Q}$ represent the possible matching candidates set, the affinities between all candidates are recorded in an adjacency matrix $\mathcal{M}^{C^{P \times Q}}$ based on applications. Let $\mathcal{X} \in \{0, 1\}^{n_{\mathcal{M}} \times C^{P \times Q}}$ be a column-wise binary vector which indicates the selected correspondences, the graph matching problem can be expressed as finding the best indicator vector \mathcal{X}^* that maximizes a score function $S(\mathcal{X}) = \mathcal{X}^T \mathcal{M}^{C^{P \times Q}} \mathcal{X}$, *i.e.*, $\mathcal{X}^* = \arg\max S(\mathcal{X})$.

Various relaxation based methods have been proposed. Most of them relax the integer constraint of \mathcal{X} . After obtaining the optimal \mathcal{X}^* in real number domain with different insights of \mathcal{M} , the discretization of \mathcal{X}^* will make the best approximated solution to the underlined IQP problem. In **C²IL**, two major graph matching algorithms, spectral matching (SM) [17] and

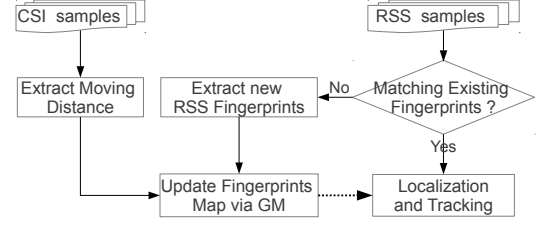


Fig. 2. **C²IL** System Architecture.

RRWM algorithm [8] are used in different stages. SM is used in extracting RSS fingerprints graph; while RRWM is used in mapping between RSS fingerprints and floor plan.

III. ARCHITECTURE OVERVIEW

Since CSI and RSS can be estimated by both AP and clients, **C²IL** can be deployed in either AP-end or client-end. Whenever it is deployed at either end, the core of the system remains. The positions of APs are not required to be known in either case.

Fig. 2 demonstrates our system architecture. When a user enters the building, we assume that the device held by the user will build wireless connections with the APs inside the building. The APs will record the RSS values and the CSI values of the received signal from the client device and send the data to a localization server. Based on the sequence of the CSI values taken during a time-window, the server will then quickly estimate the moving distance of the client in this time-window. The RSS values will be used to build a matching with the fingerprint RSS values stored in the fingerprint database, which itself is populated using crowdsourcing techniques. The estimated moving distance, together with the estimated geodesic distance of different fingerprint locations in the map, will be used to further improve the quality of the matching, and thus, the accuracy of the localization.

IV. ESTIMATING MOVING SPEED BY CSI

In this section, we mainly focus on the techniques of estimating moving speed and thus moving distance by CSI. Theoretical basis is presented first, and then we present our algorithm implementation.

A. The Electromagnetic Standing Wave Field

Wireless radio propagation in compact environment could be modelled as a superposition of large-scale path-loss, middle-scale shadowing, and small-scale multipath fading [24]. For the multipath fading, it is usually fitted to a statistical model called Rayleigh or Rician (Rayleigh fading plus strong LoS components) distribution. The ripples-like deep fading shown in Fig.1 (b) and (c) are typical Rayleigh fading pattern.

Previous speed estimation methods, are based on some statistical properties of Rayleigh distribution, *e.g.* level crossing rate (LCR) or coherent time T_c . Although it has been experimentally validated that the distance between two adjacent ripples (deep fadings) is about $\lambda/2$ (λ is the carrier wavelength) even in large-scale multipath environment (like

Manhattan city) [28], no previous works explicitly exploits such $\lambda/2$ fluctuation, since such fluctuation is encapsulated and blurred in a too general model.

However, some detailed studies of radio propagation [5], [6], [34] have indicated that, in a complex multipath environment the constructive or destructive interferences of the large sum of reflected and scattered waves will generate a *standing waves field*, and the environment becomes a weak *Electromagnetic Cavity Resonator* (ECR) [14] which hold standing waves in a very short time. According to basic physics of wave propagation, the distance between two adjacent *antinodes* (position with maximum amplitude), *towards any direction*, is $\lambda/2$, thereby the experimentally observed $\lambda/2$ fluctuation. Therefore, when a antenna traverse the indoor space with a speed v , a periodically ripples-like pattern with a frequency $f_o = 2\frac{v}{\lambda}$ appears. Such simple relationship inspires us that the moving speed v could be precisely estimated purely from the CSI, if we could precisely estimate f_o .

B. Theoretical Basis

Wireless signal propagation in indoor environment can be well modelled as Rician fading channel. In such a model, the CFR at the i -th subcarrier is described as [1]

$$H_{rice}(i) = \sqrt{\frac{1}{K+1}}H_m(i) + \sqrt{\frac{K}{K+1}}H_{LoS}(i) \quad (1)$$

where the $H_{LoS}(i)$ represents the deterministic Line-of-Sight (LoS) component, $H_m(i)$ represents the random multipath component. The Rician factor K determines the power ratio between these two components.

To simplify the system model, we will first focus on a simplified model, Rayleigh fading channel, which is a specialized form of Rician fading channel when $K = 0$. Generalized solution under Rician channel will be discussed later.

1) *In Rayleigh Fading Channel*: In a multipath environment, there are lots of objects in the environment that scatters the wireless signal. The received fading envelop at an antenna will be the superposition of a large number of these reflected and scattered waves. Since different position has different constructive or destructive interference pattern among these waves, the received signals amplitude at different locations become a random variable. When a mobile antenna passes through the environment, ripples-like deep fading appears in the instantaneous CFR as shown in Figure 1 (b) and (c).

Assuming Wide Sense Stationary Uncorrelated Scattering (WSSUS) environment and the uniformly distributed angle of arrival (AoA) of multipath components, Clark [9] has derived the auto-correlation of Rayleigh faded CIR $h(t)$ with motion at a scalar velocity v **towards any direction** is a *zeroth-order Bessel function of the first kind* that

$$E\{h(t) * h(t + t_0)\} = J_0(2\pi f_d t_0) \quad (2)$$

where $J_0(2\pi f_d t_0)$ is the bessel function with delay t_0 when the maximum Doppler shift is f_d . A general definition of the

Bessel function, for integer values of n could be represented in an integral form:

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(nt_0 - x \sin t_0) dt_0 \quad (3)$$

For $J_0(x)$, the function can be further simplified to

$$J_0(x) \approx \sqrt{\frac{2}{\pi x}} \cos(x - \frac{\pi}{4}) \quad (4)$$

There is a large positive value ϵ , when $x > \epsilon$, Eq. (4) can be approximated by a periodical function that

$$J_0(x) = J_0(x + 2\pi), x > \epsilon, \epsilon \gg 0 \quad (5)$$

Substituting Eq. (2) into Equation Eq. (5), we get:

$$J_0(2\pi f_d t_0) = J_0(2\pi f_d (t_0 + \frac{1}{f_d})) \quad (6)$$

Substituting the maxium Doppler frequency $f_d = \frac{v}{\lambda}$ into Eq. (6), we obtain:

$$J_0(2\pi \frac{v}{\lambda} t_0) = J_0(2\pi \frac{v}{\lambda} (t_0 + t_\lambda)) \quad (7)$$

where $t_\lambda = \frac{\lambda}{v}$.

Apparently, $t_\lambda \times v = \lambda$, and Eq. (7) tells a very elegant result that, **iff** $\int_{t_1}^{t_2} v dt = \lambda/2$,

$$J_0(2\pi f_d t_1) = J_0(2\pi f_d t_2) \quad (8)$$

The periodicity of the auto-correlation function of $h(t)$ will undoubtedly reflect on the original CIR $h(t)$, and CFR $H(t)$. Thereby we could say, in a typical Rayleigh fading dominant isotropically scattering environment, the distance d_f between two adjacent deep fading(s) is a half-wavelength $\lambda/2$. For example, In a 2.4G channel the wavelength $\lambda_{2.4G} \approx 12.5cm$, and in 5G channel the wavelength $\lambda_{5G} \approx 5.4cm$.

Thus, for the purpose of estimating moving distance, if there is a method to accurately count the number of deep fading(s) N_f during a time interval Δt , we are able to precisely estimate the moving distance S by

$$N_f \times \lambda = S \quad (9)$$

Noticing that the error caused by Doppler effect does exist in the estimated distance S . However since Doppler frequency caused by human walking f_d is in tenth level, which is too small compared to the channel frequency. Thus, the error caused by Doppler effect is ignored.

2) *In Rician Fading Channel*: distinction between generalized Rician fading and Rayleigh fading. The ripple-like deep fading is caused by strong interference (constructive or destructive) in multipath environment. When In strong rician fading environment, high K value will weaken the multipath effect and make it difficult to recognize and count the ripples.

However, according to our experiments in typical indoor environment, we found that even in a always-LoS path, high value K rarely happens. Figure 3 plot the K values estimated from two clients during the test using a moment-based estimator according to [1]. The client in the path has always-LoS

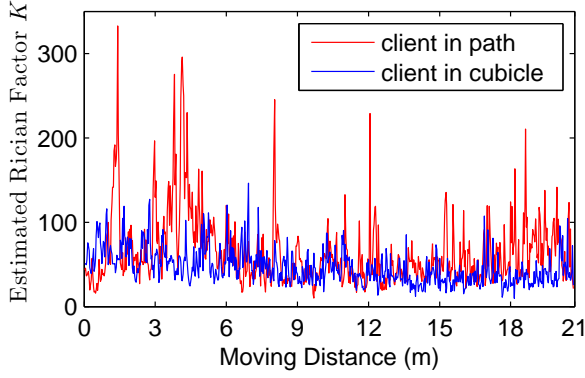


Fig. 3. Estimated Rician Factor K during the test, where the mobile AP is moving towards the end. Higher K denotes stronger LoS components reception. The client in a path has always-LoS connection, while there is no LoS component for the client in the metal-framed cubicle.

connection, while the LoS component is cut-off for the client placed in a metal-framed cubicle.

Out of our expectation, comparing to the client in cubicle, the K value measured by the client in the path doesn't benefit much from the always-LoS connection. Very large K values only happen occasionally at the beginning, while it increases very slowly even when the AP is quite close to the end. The occasional large values K at prophase is simply caused by the indoor structure at these locations, where additional propagation path is introduced. These additional paths dilute and absorb some of the multipath component. Since the large value K rarely happens even in the condition with LoS, the distance measured using the Rician fading could be approximated according to the moving velocity before and after the Rician fading.

C. Speed Estimation Algorithm

As previously described, the speed estimation problem now becomes a specific frequency estimation problem. We design a reasonable and effective processing flow. It includes Data Preprocessing, Noise Cancellation, Fading Enhancement, and Frequency Estimation. the oscillation frequency f_o .

Data Preprocessing: Every frame sent in 802.11n MCS rate at time t has an CSI H_t . It is a complex-number vector with a length of $N_{ss} \times L_{ss}$, where N_{ss} and L_{ss} are the number of MIMO spatial streams and number of measured subcarriers across the Wi-Fi bandwidth. Every complex value $h_t^i \in H_t$ describes the instantaneous amplitude a_t^i and phase θ_t^i of the underlying i -th subcarrier. In order to enable all 802.11n compatible devices to be ready for speed estimation, we only use the first spatial stream (first L_{ss} complex values in H_t) to estimate speed. Moreover, the computation space is greatly reduced. Since in multipath environment the phase θ^i is uniformly distributed between $[0, 2\pi]$ [24], which provides no discriminative information. Thus we drop the phase θ_t^i and only use amplitude $A = |H|$ to estimate speed.

The amplitude matrix $\mathbf{A}_{ori} = \{A_1, \dots, A_n\}^T$ is further defined, where A_i is the i th received column-wise amplitude.

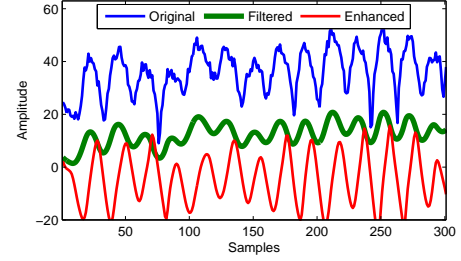


Fig. 4. The effect of each step of processing.

Since the instantaneous reception rate of frames is unstable due to the wireless traffic control, A_{ori} is resampled to a stable reception frequency f_w with the even interval between each slot, and let \mathbf{A}_{re} denote the resampled amplitude matrix.

Noise Cancellation: Convolution based noise cancellation is applied on \mathbf{A}_{re} to filter out the high frequency noise, that $\mathbf{A}_{nc} = \mathbf{A}_{re} * h_b(r)$,

$$h_b(r) = r \cdot \mathbf{1}^s$$

where $\mathbf{1}^s$ is a full-1 square matrix of size s . Currently in our system $s = 6$. This step is of great importance according to the real data evaluation, since the following fading enhancement and frequency estimation is quite sensitive to noise.

Fading Enhancement: An intuitive idea of enhance the fading is first-order derivation of \mathbf{A}_{nc} , however, first-order derivation is quite sensitive to high frequency noise rather than low frequency ripples. Another convolution is used to emphasize the fading that $\mathbf{A}_{en} = \mathbf{A}_{nc} * h_{df}$, where h_{df} is a Sobel-style calculator that

$$h_{df} = \begin{bmatrix} 2 & 5 & 2 \\ 0 & 0 & 0 \\ -2 & -5 & -2 \end{bmatrix} \quad (10)$$

Fig. 4 shows the intermediate results After first 3 processing, and it is now suitable for frequency estimation.

Frequency Estimation: Due to the MIMO configuration or other interference, deep fading(s) in all subcarriers are not guaranteed to appear simultaneously, as shown in Fig. 1, therefore the final decision of f_o are based on the estimations of each underlying subcarriers.

Extracting f_o^i for i -th subcarrier is equivalent to extracting the expected frequency $E(f_c^i)$ in the spectral graph within a frequency interval $f_{min} < f_c < f_{max}$. According to Eq. (14), f_{min} and f_{max} are set according to the speed interval of human walking that

$$f_{min} = 2 \cdot v_{min} / \lambda, \quad f_{max} = 2 \cdot v_{max} / \lambda \quad (11)$$

where the minimum speed v_{min} and the maximum speed v_{max} in our system are set to 0.8m/s and 1.6m/s.

Short-Time Fourier Transformation (STFT) with 50% overlapping window is applied to obtain the Power Spectral Density (PSD) of i -th envelope of A_{re} . It reveals the spectral density of subcarrier i along with time. To reduce the jitter, the

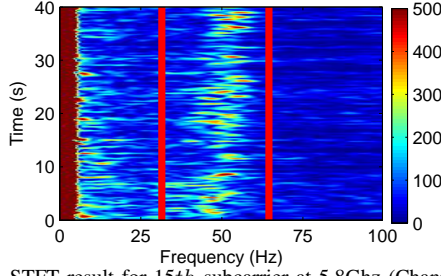


Fig. 5. STFT result for 15th subcarrier at 5.8GHz (Channel 161).

estimated f_o^i is set to the weighed expectation of frequencies between f_{min} and f_{max}

$$f_o^i = \frac{\sum_{f_{min} < s_j < f_{max}} s_j \cdot w_j}{\sum w_j}, \quad (12)$$

where w_j denotes the power of frequency f_j .

Figure 5 shows the STFT result for 15-th subcarrier. We can see very strong power around 50Hz. As $50 \times \frac{\lambda}{2} = 1.29m/s$, and it is quite close to the real walking speed at about $1.3m/s$. Two red bars denote the f_{min} and f_{max} , and they are set to 32 and 64 according the v_{min} and v_{max} settings. The final estimation of f_o is set to the median of all estimated f_o^i

$$f_o = median(f_o^1, f_o^2, \dots, f_o^n), n = N_{ss} \quad (13)$$

Then the moving speed is estimated by

$$v = \frac{\lambda \cdot f_o}{2} \quad (14)$$

It was worth to mention that although the Doppler effect does exist, as discussed in Section II, the Doppler frequency is small comparing to the carrier frequency. Thus in current system design, we did not consider the Doppler effect caused by human walking.

D. Start/Stop Detection

Fig. 6 (a) presents a CSI sample, where the user starts moving around 400th sample. Observing the degree of disorder before and after the start, we devise a correlation-based start/stop detection method. The basis on the method is to leverage the rapid spatial de-correlation property of CSI. We find that the correlation coefficient ρ between consecutive CSI samples will drop rapidly if the spatial distance d_s between them is larger than $\lambda/2$. Thus, there will be a rapid co-efficiency raising or drop to check the “moving” and “static” status. Fig. 6 (b) shows the samples’ correlation matrix. When device is static, stable and high correlation co-efficiency holds the entire upper-left area, while it disappears immediately when the device starts moving. According to our experimental evaluation, a devices is said to be moving when ρ drops below 0.4, and the final detected time t_d is quite close to the actual time t_a .

E. Minimum Sampling Rate

Similar to sensor-based system, sufficient CSI sampling rate is critical for accuracy. Due to the un-equal time distribution between fading and non-fading, the Nyquist sampling rate of

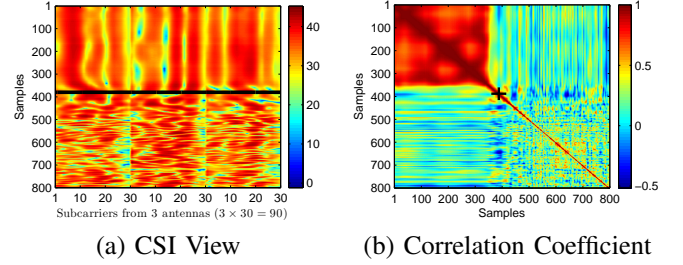


Fig. 6. (a) the CSI image of a sample. The movement starts at about 400 time-slot. (b) the correlation matrix of the CSI images shown in (a), the black cross denotes the detected start.

$f_s = 2 \cdot f_{max}$ is not sufficient. We carried out experiments to find the minimum f_s that can guarantee good accuracy. Evaluations are carried out in a wide range of channel frequencies including 2.4G (channel 1), 5.2GHz (channel 40), 5.5G (channel 100) and the highest 5.8G (Channel 161). During the experiment, testers are walking at the same speed around $v = 1.3m/s$ and the mobile device in their hands are constantly transmitting beacon frames at 500hz. After the experiments, we simulate the sampling rate f_s from 20hz to 500hz by dropping frames uniformly. Fig. 7 presents the results. We can see from the figure that the estimated speed v continuously climbs when f_s is higher than Nyquist rate f_N , and the speed stops raising when f_s is about 4 times of $f_v = 2v/\lambda$. More experiments in other situations have also confirmed the $4 \cdot f_v$ sampling rate. Therefore if we set $f_{max} = 1.6m/s$, the minimum sampling rate is only 100 (or 250) frames/s in 2.4G (or 5.8G) environment, or equivalent to approximately 40KBps or 100KBps traffic.

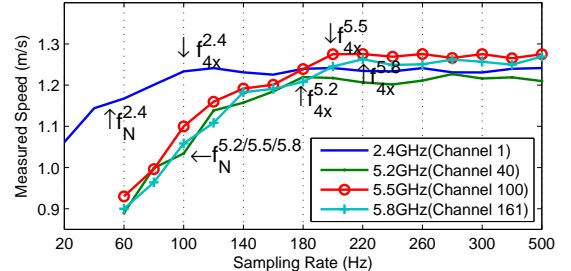


Fig. 7. The measured speeds in different sampling rate. The points denoted by f_N^* denotes the Nyquist sampling rate, while f_{4x}^* denotes the minimum require sampling rate.

The traffic burstiness is another problem. The burstiness, which happened frequently, is obviously against the CSI-based speed estimation. Since the burstiness is usually short-time high-frequency traffic phenomenon, a reasonable assumption could be made to ease this problem: people’s walking speed remain stable during the gap between two burstiness. Fig. 8 presents our solution that during each burstiness the speed is estimated, while in the gap the speed is approximated as the average.

V. RSS FINGERPRINTS EXTRACTION

RSS fingerprint are the most representative RSS point extracted from many samples for a given position, and the

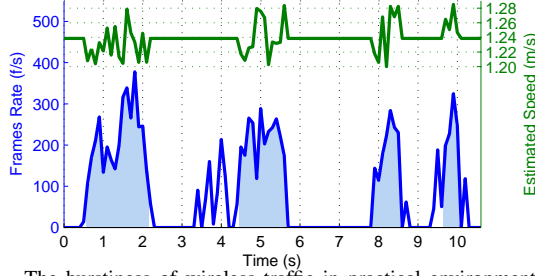


Fig. 8. The burstiness of wireless traffic in practical environment and the estimated speed.

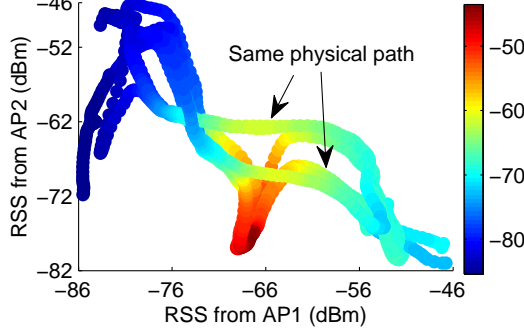


Fig. 9. 4 RSS trajectories measured from 3 APs. The color denotes the value for 3rd AP: deeper color denotes higher RSS for the 3rd AP.

error of the fingerprint will directly affect the mapping and localization accuracy. Besides the strong noise in RSS measurement, there are two main type of errors. The first is the well-known device-based measurement offset, which is mainly caused by the variation of device antenna. We observe another measurement error, called *directional shadowing problem*, which has minor effect on traditional site-survey but has strong interference on crowdsensing based approach.

A. The Directional Shadowing Problem

Human body is a strong electromagnetic energy absorption object, which may cause directional shadowing effect to wireless signal. Therefore, when the same device is placed at different position near human body, *e.g.* shirt pocket or back pocket, the RSS measurements will be obviously deviated.

This effect has been exploited in previous works [25], [33] to achieve Direction-of-Arrival (DoA) detection, however, it may cause severe error to unsupervised fingerprints-to-floor plan mapping, because there may be multiple parallel RSS trajectories corresponding to the same physical path. Fig.9 presents an example of directional shadowing from real measurement data. In this situation, the mapping algorithm may wrongfully think that there are parallel paths between the start and end, and it is very likely to cause mapping failure.

B. Fingerprints Extraction

Previous approaches usually adopted cluster-based algorithm to extract fingerprints by merging the the nearby RSS samples within a certain threshold. This is a coarse-grained algorithm, and it cannot identify the device offset and directional shadowing. Our solution comes from an intuitive observation: although the RSS samples are highly deviated for the same

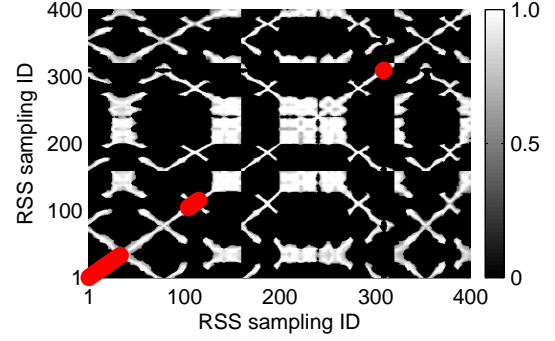


Fig. 10. The correspondence ratio among RSS sampling sequence. Ted dots denote the extracted RSS fingerprints.

physical position if under antenna variation and directional shadowing, the temporal RSS samples transition trends are very similar. If we could correctly identify the matching between these RSS trajectories, the fingerprints could be extracted without two errors mentioned above.

This intuitive idea could be transformed to a n -partite graph matching problem if we see these trajectories as curve-shaped graphs. This can further be done by $n - 1$ times iteratively graph matching between n -th trajectory and previous resulting RSS fingerprints graph. As revisited in Section II, relaxation-based approaches approximate the graph matching elegantly, and the only job is to build the affinity matrix \mathcal{M} .

Given two RSS trajectories, G^P and G^Q , each of them contains n^P and n^Q RSS measurements respectively. A possible assignment a is defined as $a = (i, i')$, where $i \in \mathcal{V}^P$ and $i' \in \mathcal{V}^Q$. Given another possible assignment $b = (j, j')$, the compatibility $m_{a,b} \in \mathcal{M}$ is assigned as follows in current system.

$$\mathcal{M}_{(a,b)} = \begin{cases} e^{-||d_{ij} - d_{i'j'}||}, & \text{if } ||d_{ij} - d_{i'j'}|| > \epsilon \\ 0, & \text{otherwise} \end{cases} \quad (15)$$

where d_{ij} and $d_{i'j'}$ are the *distance* between i and j , and their assignment pairs i' and j' , respectively. Since RSS attenuation along distance is non-linear, in our system the Minkowski distance [15] with value $p = 1.7$ is used to define the distance between RSS samples a and b .

$$d_{ab} = \left(\sum_{i=1}^n |rss_{ai} - rss_{bi}|^p \right)^{\frac{1}{p}} \quad (16)$$

where n represents the numbers of all heard APs, and rss_{ai} for the i -th AP's RSS value of RSS sample a .

In each iteration the fingerprints extraction is a partial graph matching problem, and it is not tight integer constrained. We use spectral matching (SM) algorithm to calculate the optimal column-wise vector \mathcal{X} with length $n^P \times n^Q$. \mathcal{X} is further reshaped to an *association matrix* $\mathbf{A}^{P \times Q}$, where each element \mathbf{A}_{ij} denotes the *matching rate* between $i \in \mathcal{V}^P$ and $j \in \mathcal{V}^Q$. Fig. 10 presents an example of the association matrix between two identical RSS trajectories, where high value denotes highly possibility of matching.

C. Fingerprints Transition Graph

The Fingerprints Transition Graph $\mathcal{G}^F = (\mathcal{V}^F, \mathcal{E}^F)$ records the spatial connectivity of all fingerprints. Since the vertices set \mathcal{V}^F is the fingerprints set \mathcal{C}^F , we only need to determine the edge set \mathcal{E}^F and the weight of edges $W(\mathcal{E}^F)$.

Basically, any two nodes $i, j \in \mathcal{V}^F$ will an edge $e_{ij} \in \mathcal{E}^F$ if they satisfy following two conditions:

- 1) i and j are subsequent RSS fingerprints within the same RSS fingerprints trajectory;
- 2) or when i and j belong to different trajectory, the distance d_{ij} is smaller than ϵ and at least one of both is the *start* or *end* of a trajectory.

The weight of edges will be set to the absolute distance between fingerprints in traditional approaches [7], [23], [31], however, in $\mathbf{C}^2\mathbf{IL}$, there is no absolute distance information due to the absence of physical space measurement.

Fortunately, the walking duration becomes a distance indicator, and we can assign *virtual distance* to the weight of edges. The virtual distance is based on a reasonable assumption: although people may vary their walking speed according to mood or ongoing tasks, they usually maintain a constant speed during a single walking trip. Based on this assumption, we can infer the distance ratio between nearby fingerprints sequences along a trajectory, and eventually obtain a global distance matrix. Since the assignment of virtual distance for a fingerprint sequence is gradually propagated, error accumulation may exist. However, the logical isolation of indoor space is very small, the virtual distance for most of segments can be assigned in less than 3 iteration, the error propagation is therefore ignorable. Our virtual distance assignment algorithm is briefly described as blow.

VI. MAPPING BETWEEN FINGERPRINTS AND FLOOR PLAN

Ordinary floor plan is not friendly for crowdsensing based approaches, especially when client-side doesn't provide direction information. By studying the shortest walking distance (SWD) in indoor space, we realize that there is a highly curly 2D manifold embedded in a 2D floor plan polygon. The unfolded version of this manifold will remove the direction information but preserving the SWD information, which is friendly for crowdsensing-based approaches.

The technical underpinning of automatic establishing the mapping is based on the topological similarity between the unfolded version of floor plan and the fingerprint transition graph, which is a typical graph matching problem. However, challenges exist in many aspects, *e.g.* unacceptable computation cost and low accuracy of large-scale GM, scaling effect, and the unsupervised nature of auto mapping. Here we start from transforming the floor plan, then go through the auto mapping algorithm.

A. Floor Plan in Manifold's Eyes

n -dimensional manifold is a topological space that near each point resembles n -dimensional euclidean space, while globally not euclidean. The indoor floor plan shares the same property. Due to the obstruction of walls, the shortest walking distance

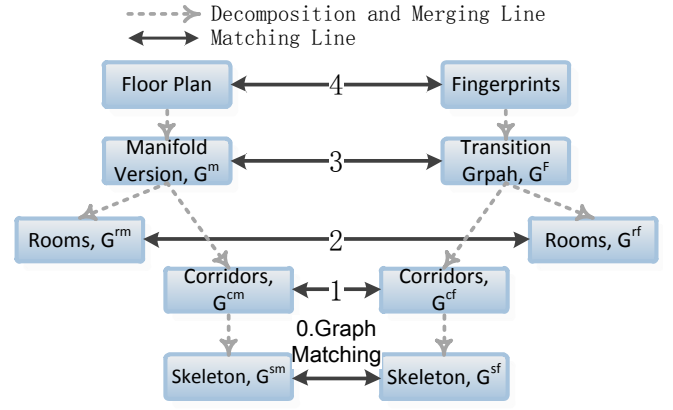


Fig. 11. The Skeleton-based matching algorithm. Skeleton graphs are first extracted from both side. Based on the matching between skeletons, upper levels are gradually matched.

(SWD) between two points i and j in the floor plan \mathbf{P} equals to the euclidean distance $d_{eu}(i, j)$ iff the points i and j are within the same local isolation with direct Line-of-sight distance. When they are not in the same isolation, the SWD would be the geodesic distance $d_{geo}(i, j)$ which detour through various of obstructions. In this way, indoor space could be essentially viewed as a 2D-manifold \mathcal{S} embedded in a 2D-polygon \mathbf{P} .

In a manifold space, the euclidean distance is misleading. The geodesic distance actually reveals the true structure of the manifold. Therefore, we resample the 2D floor plan \mathbf{P} using n points uniformly scatted. A n points graph G^M is then constructed. For every pair of nodes $i, j \in V^M$, there is an edge e_{ij} iff the correspondence points i^P, j^P in the floor plan \mathbf{P} are in their mutual neighborhood with *direct line-of-sight* distance, and the weight for edge is the direct distance that $W(e_{ij}) = d_{eu}(i^P, j^P)$. matrix A^M , which records the pairwise geodesic distance between all sampling points.

B. Unsupervised Accurate Mapping

Since the RSS samples are measured along users' walking trajectories, the RSS fingerprints transition graph G^F also share the same floor plan manifold structure. An intuitive idea of establishing the fingerprint map is to apply graph matching algorithm directly upon G^F and G^M . However, the accuracy and performance of large-scale graph matching (>50 points) is very poor for unsupervised situation. A lightweight relaxation to the problem is to apply graphing matching only on corridor points. Once the corridor points graph are mapped correctly, it is easy to match the rooms points. Unfortunately, the accuracy and performance of graph matching between such corridor points graphs is still not satisfied.

Fortunately, the highly sparse chain structure of corridor points give us a hint, and we devise a method called "Skeleton-Based Matching" to achieve unsupervised accurate mapping between G^F and G^M even for very complex indoor environment. The basic idea is that: due to the high sparsity and chained structure of corridor points graph, we can extract a coarse-grained skeleton graph from it while preserving the identical topological structure. Adopting GM algorithm on

Algorithm 1: Extracting Corridor Points

Data: a sparse graph $G = (V, E)$
Result: a skeleton graph G^S for G
 $G^T \leftarrow G$, $done \leftarrow \text{false}$
while $done = \text{false}$ **do**
 calculate centrality for $v \in V^T$, $done \leftarrow \text{true}$
 forall the $v_i \in V^T$ **do**
 if $C(v_i) \leq \tau$ **then**
 remove v_i from G^T , $done \leftarrow \text{false}$

skeleton graphs will result high accuracy and performance. Once the main structures of G^F and G^M are matched, the rest parts will be matched easily. Fig. 11 illustrates the basic idea of skeleton-based matching.

The algorithm is detailed in following 3 steps, including *skeleton graphs extraction*, *skeleton graphs normalization*, *skeleton graphs matching*, and *find-grained points matching*.

1) *Extract the skeleton graph*: Two sub-steps are required to extract the skeleton graphs. 1). Identifying the corridor points graph $G^{CF} \in G^F$ and $G^{CM} \in G^M$; 2). extracting skeleton graphs G^{SF} and G^{SM} based on corridor points graph $G^{CF} \in G^F$.

In the first sub-step, a customized centrality measure $C(V)$ is designed to identify the core corridor network. For a given point $v \in V$, its centrality $C(v)$ is measured as follows.

$$C(v) = \sum_{s \neq v \neq t \in V} \sigma_{st}(v) \quad (17)$$

where $\sigma_{st}(v)$ is the numbers of shortest path from s to t via v .

Based on this definition, we design a iterative algorithm to remove the non-central points effectively. In each round of iteration, the centrality $C(v)$ is measured for every points. If $C(v)$ is smaller than a low-bound τ , then remove the points from the graph. This procedure repeats until no points is removed. Algorithm.1 shows the pseudo-code of corridor points extraction.

The choose of value τ is critical. Fig.12 shows the relationship between τ and size of graph. Small τ will bring in redundant points, while large τ will remove all points. Since the linear growth of τ will lead to monotonic decrease of number of remaining points, we use binary search to find the optimal τ point.

In the second sub-step, the skeleton graphs V^S is generated by clustering the corridor points graph G^{CF} and G^{CM} . We use spectral clustering [29] (SC) as the clustering algorithm. SC is computationally faster than K-means and it only requires the adjacency matrix which is exactly suitable in our case that both G^{CF} and G^{CM} are represented only in adjacency matrix. By clustering on G^{CF} and G^{CM} , we obtain the vertices set of skeleton graph G^{SF} and G^{SM} . The edge set E^{SF} and E^{SM} follow the underlying points, that if two points i, j , belonging to different clusters c_a and c_b respectively, have an edge, then

there is an edge between c_a and c_b . The weight of edge e_{ab} is defined as the shortest distance between the *central points* of cluster a and b , and the *central point* of a cluster is the point i which has the shortest distances to other points within the cluster.

2) *Graphs Normalization*: Since the graph matching algorithm assigns the matching score according to pairwise distance, G^F and G^M must be normalized to same scale so as to guarantee an accurate matching result.

The scaling effect can be undone by some global normalization of the graph shape. Possible ways includes normalization of bounding box, or normalized Laplace-Beltrami eigenvalues (LBE). The bounding box approaches works only for rigid graph, while LBE is sensitive to deformations. We use the commonly accepted *longest geodesic distance* L_{lg} as scale indicator.

edge weights of G^{SF} and G^{SM} are divided by the scale indicator. In this way, the scaling of both graph is canceled.

3) *Skeletons Matching*: After the extraction and normalization of skeleton graphs, we now find the best correspondence between G^{SF} and G^{SM} . Let positive symmetrical square matrix M^{SR} and M^{SM} represent their adjacency matrices. We build the affinity matrix $M^{SR \times SM}$ for graph matching as follows.

$$M^{SR \times SM} = e^{(\mathbf{1}^{SM} \otimes M^{SR} - \mathbf{1}^{SR} \otimes M^{SM})^2} \quad (18)$$

where \otimes denotes the Kronecker product [11] and $\mathbf{1}^{SR}$ denotes the full-1 matrix with the same size of G^{SR} . The idea behind Eq. 18 is to enumerate all possible matching candidates and store them in a large adjacency matrix $M^{SR \times SM}$. We use RRWM algorithm [8] to perform the graph matching, and Hungarian algorithm is further applied to discretize the \mathcal{X} in order to meet the final integer constraints $\mathcal{X} \in \{0, 1\}^n$.

4) *Corridor Points Matching*: Although G^{SF} and G^{SM} are matching in previous step, however, the corresponding clustering groups in G^{CF} and G^{CM} are not necessarily matched due to the inconsistency of clustering operation.

In the corridor points graphs, we notice that only a few points connect multiple chain structures, and they may serve as the *bridge points*. Since the topological structures of G^{CF} and G^{CM} are identical, if we could identify the correct correspondences of these *bridge points*, the points within the chain structure will be matched easily. In order to identify the bridge node, we introduce a new metric called "bridge centrality", which is equal to the number of shortest paths

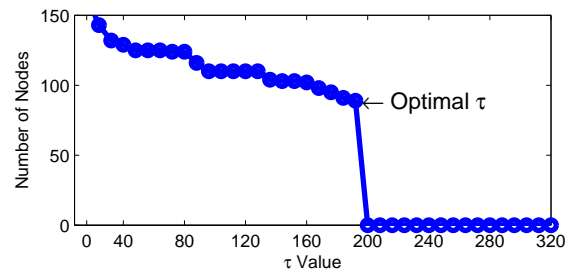


Fig. 12. Figure shows the linear growth of τ , monotonic decrease of graph size, and the optimal τ point.

from all vertices to all others within nearby clusters that pass through that node.

$$C_{bg}(v) = \sum_{\{(s,t)|v \in c^i, s,t \in NN(c^i), s,t \notin c^i\}} \sigma_{st}(v) \quad (19)$$

where $NN(c^i)$ denotes the nearby clusters around c^i . The bridge point will be the point with the highest bridge centrality.

After identifying the of correspondence of bridge points in G^{CF} and G^{CM} , the chain structures are easily matched according to the start and end bridge points.

5) *Rooms Points Matching*: If we remove the corridor points graph G^{CM} from G^M , the rooms points will naturally forms several clusters. Each room points cluster C_R connects to the corridor by a door point, and the rooms matching is also easy by matching the door points in both G^F and G^M .

But if there is more than one rooms connect to a single door point in the corridor, e.g. two rooms in opposite side along the corridor, there may be mis-matching. This kind of mis-matching can be canceled using coarse-grained propagation model. Along with the corridor points matching, the coarse-grained position estimation for APs can be done. With the rough location of APs, the rooms mis-matching can be easily eliminated by checking the RSS values. The smaller RSS difference means higher probability of being in the candidate rooms.

P, the extracted RSS FTM also reveals the true structure the manifold, which means the RSS FTM graph G^T and \mathfrak{S} 's unfolded version G^M are naturally matching, and we are here to find the matching.b

$i \in V^P$ and $i' \in V^Q$. We next define the *affinity matrix* M that $M(a, b) = f(i, i', j, j') = f(a_e^P, a_e^Q)$, where $a_e^P = A^P(e_{ij})$ and $a_e^Q = A^Q(e_{i'j'})$, and $f(\cdot)$ measures the compatibility of assignments a and b . Higher $M(a, b)$ denotes higher compatibility and vice versa. In such setting, the graph matching problem transformed to finding the optimal set C of assignments that will maximize the inter-cluster score $S = \sum_{a,b \in C} M(a, b)$. Cluster C can be further represented by an indicator vector x , and we can rewrite the inter-cluster score as: $x^T M x$ inter-cluster score S is the principal eigenvector of M , and there are various methods to find the principal eigenvector x^* . M . In our setting, M is a $n^P \times n^Q$ symmetric non-negative matrix, where each element $M(a, b)$ is defined as

the points distance. manifold is not globally equal to the euclidean distance that $\forall p_i, p_j \in \mathfrak{S}: d_{eu}(p_i, p_j) \leq d_{geo}(p_i, p_j)$, smaller than the euclidean distance. distance between RSS fingerprints along users' walking trajectories. Apparently, the original indoor space manifold \mathfrak{S} should be transformed to an unfolded version so as to help establish the mapping between the RSS space and indoor space. to transform the \mathfrak{S} into a Manifold Unfolded Space (MUS). In MUS the euclidean distance between any two points is equal to the geodesic distance on original manifold \mathfrak{S} , $\forall p_i, p_j \in S^p: d_{eu}^{S^p}(p_i, p_j) = d_{geo}^{\mathfrak{S}}(p_i, p_j)$. geodesic distance matrix D_{geo} to the classical multi-dimensional scaling (MDS) [16], and matrix D_{geo} is obtained by calculating the shortest walking distance

between all pairs of points on the manifold. in Section.V, the correspondence matrix X is still obtained through the process of building compatibility matrix M and calculating the principal eigenvector x^* . What is different from the trajectories matching is the distance function for adjacency matrix of two graphs, here the distance is the euclidean distance in MUS. structure is of highly symmetry, the correspondence shown in X may contain multiple candidate mapping. methods. to obtain the coarse-grained RSS distribution in the indoor space. The reversion of predicted RSS distribution at some sampling position will correspond to the predicted RSS measurement of the APs. the RSS distribution obeys the estimated trends. Therefore, the true correspondence leaves behind.

VII. LOCALIZATION AND TRACKING

C²IL provides a unified localization and tracking service by treating the direction localization request as a tracking request without historical data. Here we mainly focus on tracking technique in **C²IL**. Unlike the stateless K-NN based method which is widely adopted in previous approaches, in our solution, the users' trajectories are globally determined from the very beginning by transforming the tracking problem to a graph matching problem between the measured RSS samples transition graph G^S and the fingerprint transition graph G^T . After the graph matching, the accuracy is further improved by bringing in the CSI-based speed estimation through a particle-filter based fusion. Here we start introducing these two steps.

A. Graph Matching Based Tracking

Graph matching based tracking is to find the best correspondence between the sequence of RSS samples of tracking request and extracted fingerprints. This is exactly the same matching process undertaken in Section V, except for the differences that the tracking is to find the matched points, while in Section V the un-matched RSS samples are added to the fingerprints database.

Let $\mathcal{X}^{n^S \times n^F}$ represent the association matrix obtained through spectral matching where n^S and n^F are the numbers of RSS samples and candidate fingerprints respectively. Due to the error in RSS measurement and fingerprints map construction, a single RSS samples p_i may correspond to multiple fingerprints in \mathcal{X} , e.g., a RSS sample may correspond to two fingerprints, one is in corridor, and another is in a room. Fortunately the temporal correlation can help eliminate those false correspondence by checking the spatial continuity between current and subsequent candidate fingerprints. After eliminating the false correspondence, the globally estimated coordinates sequence will be given by $T_{GS} = f_{GT}(G^S)$, where $f: fingerprint \rightarrow G^M$ represents the mapping from fingerprints to the floor plan manifold G^M , and T_{GS} are the resulting coordinates sequence in G^M .

However, full graph matching between all fingerprints set and RSS samples is time-consuming, therefore a matching candidates pruning process is introduced to meet practical demand of real-time tracking of multiple clients.

Matching Candidates Pruning: The main idea of the pruning is to find the probable walking area for the tracking request using coarse-grained nearest neighbour (NN) method, which will significantly reduce the search space.

Let $R^T = \{r_1^T, \dots, r_n^T\}$ represent the RSS sequence of the tracking request at time t , where r_i is a n -dimensional RSS vector. The length n is smaller than a positive integer \mathcal{L} , such that the tracking algorithm provides limited backtracking. Let $R^F = \{r_1^F, \dots, r_n^F\}$ denote the RSS fingerprints set and graph G_P^F denote the positions of fingerprints in the floor plan. The pruned matching candidate set S_u is defined as follows.

$$S_u = \{r^F | r^F \in R^F, r^F \in \epsilon\text{-NN}_{R^F}(\text{NN}_{R^F}(R^T))\} \quad (20)$$

where $\text{NN}_{R^F}(R^T)$ denotes the nearest neighbours of R^T within the fingerprints set R^F , and $\epsilon\text{-NN}_A(B)$ denotes the B 's neighbourhood within distance ϵ in the set A . Here the ϵ is set to 3m.

B. Fine-Grained User Tracking

The movement estimation from trajectories matching (TM) and CSI-based speed estimation (CBSE) are naturally complementary. TM is slow-responsive yet accurate in long time, while CBSE may drift with time cumulation yet be accurate in short time. Obviously, a fusion scheme for TM and CBSE will significantly improve the accuracy. Due to the linearity constraints and difficulties in correctly estimating the error covariance matrices, we don't use the conventional Kalman filter. The more robust particle filter (PF) is adopted as the fusion algorithm.

The state space of the tracking is a two dimensional vector $X_t = [M_i, v_t]$, where M_i represents the i -th node in the floor plan manifold M , and v for the walking speed. The main challenge here is that the speed v has no direction information, therefore every round the PF evolves and there are multiple candidate predictions. For instance, when walking in a corridor without optional paths, there are two candidate directions, forward and backward, and therefore two state candidates $X_{for} = [M_j, v_t]$ or $X_{rev} = [M_h, v_t]$, where M_j and M_h are two nearby nodes in different directions.

Fortunately, the PF can handle this problem elegantly by sprinkling different amounts of particles in different directions. In the prediction phase of each filter iteration, the particles will be re-sampled and enumerate all possible candidate directions.

Fig. 13 illustrates the particles distribution along a path, in each round the particles enumerate all possible candidate directions. When there are optional paths, the particles will enumerate all possible options. After the fusion, the M_i is transformed to floor plan coordinates.

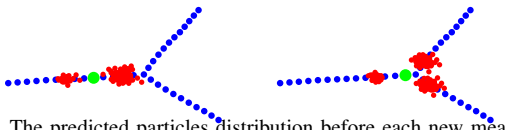


Fig. 13. The predicted particles distribution before each new measurement: red dots represent the particles, and green dots for the determined position in last round.

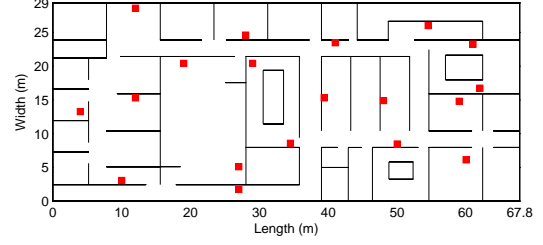


Fig. 14. The floor plan of our test area, and red dots denote the deployed customized APs.

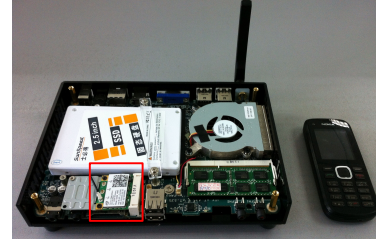


Fig. 15. Prototype system for C^2IL with Intel Atom-based Mini PC and 5300NIC as AP.

VIII. SYSTEM EVALUATION

A. Prototype System with Customized APs

We designed and developed a prototype hardware system for C^2IL . The prototype system is deployed in a large $2000m^2$ office environment with circular corridor network as shown in Fig. 14. To maximize the participation rate, C^2IL is deployed at AP-end. In total, 18 customized APs are sparsely deployed across the office that provide both wireless networking service and localization service. They will forward the measured RSS and CSI values to a central localization server. A developed simple client App is installed on testers' Android smartphones. The localization server will calculate the clients' position based on the networking traffic, and clients only need to read the location coordinate from server.

Although CSI is a standard PHY layer information, currently only Intel 5300 NIC can export it to user level. Our customized AP is simply an Intel ATOM-based mini PC with 5300 NIC. Figure 15 shows our customized AP. It equips with single-core 1.6Ghz ATOM CPU and 5300 NIC. The total cost is about \$90. The OS is Ubuntu 12.04 64-bit, and AP function is hosted by hostapd. Besides the ordinary AP interface, a monitor virtual interface is also added to overhear the wireless traffic. Both measured CSI and RSS are uploaded to localization server in real-time.

B. CSI-based Speed Estimation

we made extensive experiments to evaluate the universality of CSI-based speed estimation. Three groups of experiments mainly cover all wireless channel combinations, propagation scenarios, and a comprehensive evaluation in a typical office environment.

1) *Performance in different channel combinations:* In this section, we mainly focus on the performance of speed estimation in different channels and transmission modulations

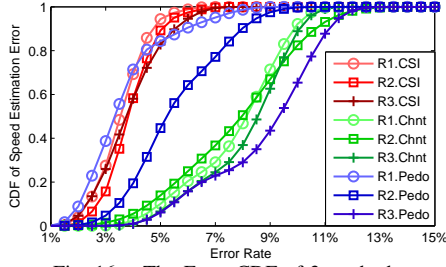


Fig. 16. The Error CDF of 3 methods.

(MCS). The experiments covered all supported channels (channel 1-13 in 2.4Ghz, channel 38-163 in 5Ghz) and MCS values (MCS 0-7). During each $\langle \text{channel}, \text{MCS} \rangle$ combination, a testing laptop moves along a 42m long corridor, and it transmits random messages at 800Hz in the given $\langle \text{channel}, \text{MCS} \rangle$ combination. Several APs along the path will record the CSI value during the movement.

2) *Performance in different propagation scenarios:* Performance in a typical strong multipath environment has been evaluated in previous section. To evaluate the performance of speed estimation in a pure non-multipath to moderate multipath environments, 3 groups of experiments are carried out in a fully-open square, a rooftop platform with several cooling infrastructures, and a large gymnasium. In each environment, 7 tests are carried out to cover the most commonly used $\langle \text{channel}, \text{MCS} \rangle$ combination, $\langle 6, 0 \rangle$, $\langle 40, 0 \rangle$, $\langle 60, 0 \rangle$, $\langle 100, 0 \rangle$, $\langle 120, 0 \rangle$, $\langle 140, 0 \rangle$, and $\langle 161, 0 \rangle$, covering the channels from 2.4Ghz to 5.8Ghz.

3) *Performance in practical office environment:* In this experiment, 10 students are asked to walk 3 times around the 129m long circular corridor as shown in Fig. 14. They are asked to maintain constant walking speed in first round, and the speed may change slightly and remarkably in second and third round. The measured speed will be integrated to walking distance D_w . We mainly considered the error rate $e_w = |D_{\text{path}} - D_w| / D_{\text{path}}$.

Two other approaches, pedometer-based and channel coherent-time based [22], are also developed as comparison. For pedometer we use NASC [23] method to detect steps, and the step length is predefined according to the training data. For coherent-time based approaches, we use a constant $\xi = 0.396$, which is manually optimized specific for this experimental field. The laptop held by students constantly transmits beacon-frames at 500Hz, and accelerometer data used for HD active protection is recorded for step detection.

Fig. 16 presents the error CDF of three methods in three walking manners: with constant speed (R1), slightly varying speed (R2), and with marked change (R3). The experiment shows that, comparing to pedometer-based approach, C^2IL with CSI can achieve better accuracy without requiring predefined constants or training data, meanwhile, C^2IL significantly outperforms pedometer-based (Pedo) approaches when speed is varying or pre-defined step length is out of effectiveness. Coherent-time based approach (Chnt) has the similar speed-invariance feature, however, the accuracy is considerably poorer than C^2IL and pedometer in constant speed.

The accuracy of CSI-based distance estimation is then evaluated. To evaluate the influence of richness of multipath components to the distance estimation error, we carried out experiments in 3 typical environments, a compact corridor, a large office environment, and a very large hall. In each environment we walked along a 50m straight line for 10 times. Two APs were simultaneously used to estimate walking distance. One was placed at the end of the path with strong LoS component, and the other was placed in a cubicle to cut-off the LoS components to simulate strong Rayleigh fading. Figure 17 (a) plots the Rician K Factor [1] along the walking path in different environment, which estimate the degree of LoS components. Very low and stable K appears in corridor environment which means there are rich multipath components, while in office and hall the multipath components is significantly reduced due to the weak reflection in large wide-open space. Figure 17 (b) and (c) plot the CDF of estimated walking distance by the AP in path and cubicle respectively. We see in the best situation that in a corridor with strong multipath component, there is only 3% error. In Figure 17 (c), we also see small error happened in the path, and in the worst situation, in a large hall with very weak multipath components, the averaged estimated error is less than 10%.

C. Mapping Accuracy Test

During the mapping system test, to fasten the mapping system convergence, 4 students are asked to walk through the environment arbitrarily for 10 minutes, so as to ensure all rooms and path are covered. During the walking, the android smart phones in their pockets are constantly communicating with AP to create enough wireless traffic.

To have a view of extracted fingerprints, we use Multi-Dimensional Scaling (MDS) algorithm to visualize the fingerprints transition graph G^F in Fig. 18 (a) and other graph structure, e.g. Fig. 18 (a), (b), (c), and (g). Fig. 18 (b) presents the extracted corridor points graph G^{CF} from fingerprints set. To remove all point except corridor points, the optimal threshold τ^* is 127 after 6 round of attempts. G^{CF} is clustered to extract skeleton graphs, and it is shown in Fig. 18 (c).

On the other hand, the floor plan is sampled by totally 381 point using 2×2 grid, and the extract corridor points are shown in Fig. 18 (d). Despite quite similar to Fig. 18 (c), they have very different clustering, and the extracted skeleton is presented in Fig. 18 (e). For both graphs G^{CF} and G^{CM} we detect the “bridge point” according to the clustering, the these bridge points are shown as magenta points in Fig. 18 (b) and Fig. 18 (d).

After graph normalization, Fig. 18 (f) presents the skeleton matching results. We can see both skeletons are matched even if there are mild mis-match between Fig. 18 (c) and Fig. 18 (e). The corridor points are completely matching in Fig. 18 (g) according to the skeleton matching and the bridge points. The above points are corridor points from fingerprints, while the bottom points are from floor plan. Fig. 18 (h) presents the final mapping error. We see more than 80% of corridor points are matched within an error of 1m, and the largest error

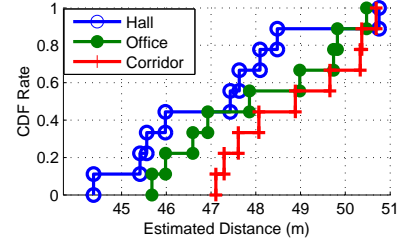
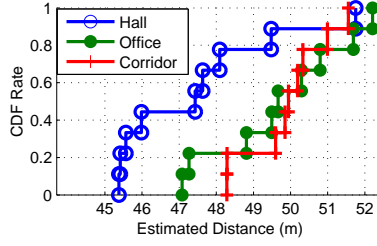
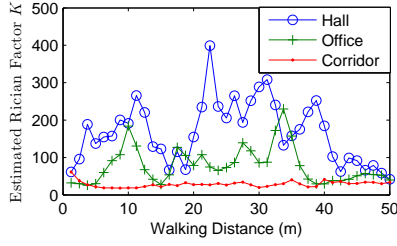
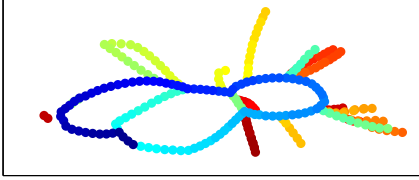
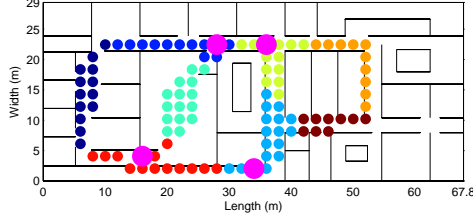


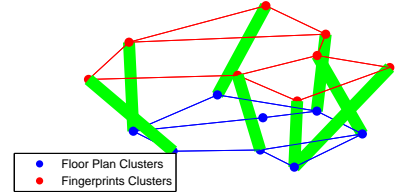
Fig. 17. (a) the Rician factor K along the path in different environment. (b) the CDF of estimated distance without LoS components (c) the CDF with strong LoS components.



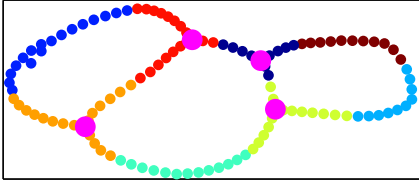
(a) MDS View of Fingerprints, G^F



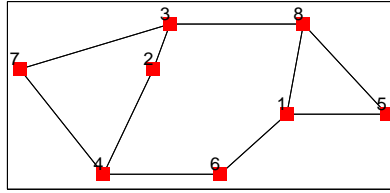
(d) Corridor Points of Floor Plan, G^{CM}



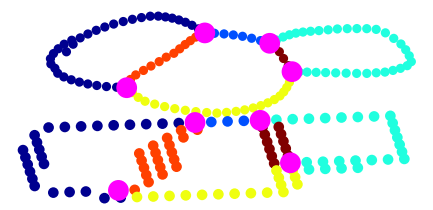
(f) Skeletons Matching



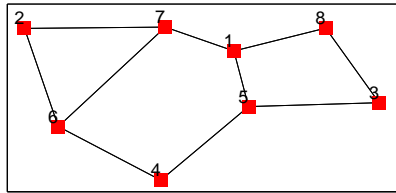
(b) Corridor Points of Fingerprints, G^{CF}



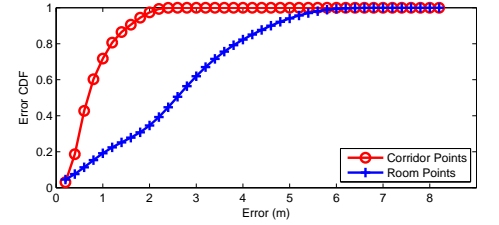
(e) Extracted Skeleton Graph, G^{SM}



(g) Corridor Points Matching



(c) Extracted Skeleton Graph, G^{SF}



(h) Matching Error CDF

Fig. 18. Figure (a), (b), and (c) shows the MDS view of fingerprints graph G^F , extracted corridor points graph G^{CF} , and skeleton graph G^{SF} respectively. Figure (d), and (e) shows the MDS view of extracted corridor points graph of floor plan G^{CM} , and skeleton graph G^{SM} . Figure (f) and (g) presents the matching results of skeletons and corridor points. Figure (h) shows the matching error CDF of both corridor points and room points.

is merely around 2m. The room mapping error is larger. The error for more than 80% of room points is within 3m, however it is acceptable, since there is no reference points to align the room points with the room layout.

rooms for 20 minutes to cover all accessible area. The RSS and CSI data is collected, and Fig. 18 (b) and (b) shows the visualization of G^T and G^{CT} respectively. Please note that our mapping algorithm requires only the adjacency matrix of graphs, and the MDS-based position is only for visualization. The graph matching result is shown in figure (d). For each graph G^{SF} or G^{SM} there are 8 points, the graph matching under such small scale could achieve stably and accurate mapping. Figure (e) shows the matching of corridor points under the guide of skeleton matching and the "bridge points". Figure (f) shows the error of both corridor points mapping and room points mapping. We can see clear that the corridor points are highly matched. The corridor points maximum error

is still under 2m, while the maximum room matching error is under 6m.

D. Localization and Tracking

In this test, we site-survey the testing environment with approximately 3×3 m grid and implement RADAR system as comparison to C^2IL . A student walks along a pre-defined testing trace for 5 rounds, and we randomly select ℓ successive RSS samples for testing.

Fig. 19 shows the error CDF of RADAR and C^2IL with different length of historical data. Based on the commonly accepted 3×3 m grid site-survey, RADAR has less than 3.5m error in 80% situations, while C^2IL is only 1m error larger than RADAR. When in tracking scenarios, every 3m of increased historical data brings approximately 0.7m tracking accuracy improvement. When historical data is 10m, C^2IL tracking accuracy is very close to RADAR-based tracking. In

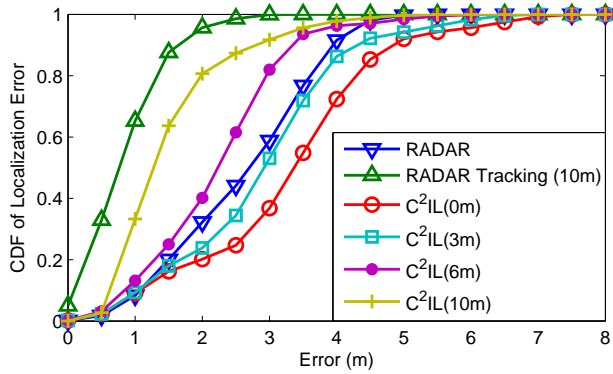


Fig. 19. The localization & tracking error CDF with different amount ℓ of historical RSS fingerprints data. The RADAR scheme is as comparison. The denotes the direct single point localization, while denotes the tracking service with a historical data of 10m.

practical tracking scenarios, 10m of historical data is a very low requirement.

IX. CONCLUSION

In this paper, we proposed an indoor localization and tracking scheme, C^2IL . Our scheme does not require using additional sensors, except the availability of 802.11n wireless connection. We believe that C^2IL is the first scheme that the really benefits from the multipath effect in complex environment. An innovative method is proposed to accurately estimate the moving speed and distance purely based on 802.11n CSI, which should find a wide range of applications alone. Based on this accurate distance estimation, we built the mapping between RSS fingerprints and location using unsupervised learning, and unified the localization and tracking. Our extensive evaluation results indicate that our scheme C^2IL successfully handles very complex indoor structure and simultaneously provides the best performance in contribution rate, localization cost, and localization/tracking accuracy.

REFERENCES

- [1] ABDI, A., TEPEDELENLIOGLU, C., KAVEH, M., AND GIANNAKIS, G. On the estimation of the k parameter for the rice fading distribution. *Communications Letters, IEEE* (2001).
- [2] AZEMI, G., SENADJI, B., AND BOASHASH, B. Mobile unit velocity estimation based on the instantaneous frequency of the received signal. *Vehicular Technology, IEEE Transactions on* (2004).
- [3] AZIZYAN, M., CONSTANDACHE, I., AND ROY CHOUDHURY, R. Surroundsense: mobile phone localization via ambience fingerprinting. In *ACM Mobicom'2009*.
- [4] BAH, P., AND PADMANABHAN, V. Radar: An in-building rf-based user location and tracking system. In *IEEE INFOCOM'2000*.
- [5] BRAUN, W., AND DERSCH, U. A physical mobile radio channel model. *Vehicular Technology, IEEE Transactions on* (1991).
- [6] CHEN, S., AND JENG, S. An sbr/image approach for radio wave propagation in indoor environments with metallic furniture. *Antennas and Propagation, IEEE Transactions on* (1997).
- [7] CHINTALAPUDI, K. K., IYER, A. P., AND PADMANABHAN, V. N. Indoor localization without the pain. In *ACM Mobicom'2010*.
- [8] CHO, M., LEE, J., AND LEE, K. Reweighted random walks for graph matching. *Computer Vision-ECCV* (2010).
- [9] CLARKE, R. H. A statistical theory of mobile-radio reception.
- [10] CONSTANDACHE, I., CHOUDHURY, R., AND RHEE, I. Towards mobile phone localization without war-driving. In *IEEE INFOCOM'2010*.
- [11] GARSIA, A. M., AND REMMEL, J. Shuffles of permutations and the Kronecker product. *Graphs and Combinatorics* (1985).

- [12] GUHA, S., PLARRE, K., LISSNER, D., MITRA, S., KRISHNA, B., DUTTA, P., AND KUMAR, S. Autowitness: locating and tracking stolen property while tolerating gps and radio outages. In *ACM Sensys'2006*.
- [13] HALPERIN, D., HU, W., SHETH, A., AND WETHERALL, D. Predictable 802.11 Packet Delivery from Wireless Channel Measurements. In *ACM SIGCOMM'2010*.
- [14] HILL, D. *Electromagnetic fields in cavities: deterministic and statistical theories*. Wiley-IEEE Press, 2009.
- [15] ICHINO, M., AND YAGUCHI, H. Generalized Minkowski metrics for mixed feature-type data analysis. *IEEE Transactions on Systems, Man, and Cybernetics* (1994).
- [16] KRUSKAL, J. B. Nonmetric multidimensional scaling: A numerical method.
- [17] LEORDEANU, M., AND HEBERT, M. A spectral technique for correspondence problems using pairwise constraints. In *IEEE ICCV'2005*.
- [18] LIU, H., GAN, Y., YANG, J., SIDHOM, S., WANG, Y., CHEN, Y., AND YE, F. Push the limit of wifi based localization for smartphones. In *ACM Mobicom'2012*.
- [19] LIU, J., PRIYANTHA, B., HART, T., RAMOS, H., LOUREIRO, A., AND WANG, Q. Energy efficient gps sensing with cloud offloading.
- [20] MOHANTY, S. Vepsd: a novel velocity estimation algorithm for next-generation wireless systems. *Wireless Communications, IEEE Transactions on* (2005).
- [21] NANDAKUMAR, R., CHINTALAPUDI, K. K., AND PADMANABHAN, V. N. Centaur: locating devices in an office environment. In *ACM Mobicom'2012*.
- [22] PRICOPE, B., AND HAAS, H. Experimental Validation of a New Pedestrian Speed Estimator for OFDM Systems in Indoor Environments. In *IEEE GLOBECOM'2011*.
- [23] RAI, A., CHINTALAPUDI, K. K., PADMANABHAN, V. N., AND SEN, R. Zee: zero-effort crowdsourcing for indoor localization. In *ACM Mobicom'2012*.
- [24] RAPPAPORT, T. *Wireless communications: principles and practice*. IEEE press, 1996.
- [25] SEN, S., CHOUDHURY, R. R., AND NELAKUDITI, S. Spinloc: spin once to know your location. In *ACM HotMobile'2012*.
- [26] SEN, S., RADUNOVIC, B., CHOUDHURY, R., AND MINKA, T. You are facing the mona lisa: spot localization using phy layer information. In *ACM Mobisys'2012*.
- [27] SEN, S., RADUNOVIC, B., CHOUDHURY, R. R., AND MINKA, T. You are facing the mona lisa: spot localization using phy layer information. In *ACM Mobisys'2012*.
- [28] SKLAR, B. Rayleigh fading channels in mobile digital communication systems. i. characterization. *Communications Magazine, IEEE* (1997).
- [29] VON LUXBURG, U. A tutorial on spectral clustering. *Statistics and computing, Springer* (2007).
- [30] WU, K., XIAO, J., YI, Y., GAO, M., AND NI, L. M. FILA: Fine-grained indoor localization. In *IEEE INFOCOM'2012*.
- [31] YANG, Z., WU, C., AND LIU, Y. Locating in fingerprint space: wireless indoor localization with little human intervention. In *ACM Mobicom'2012*.
- [32] YOUSSEF, M., AND AGRAWALA, A. The horus location determination system. *Wireless Networks* (2008).
- [33] ZHANG, Z., ZHOU, X., ZHANG, W., ZHANG, Y., WANG, G., ZHAO, B., AND ZHENG, H. I am the antenna: Accurate outdoor ap location using smartphones. In *ACM Mobicom'2011*.
- [34] ZONOZI, M., AND DASSANAYAKE, P. Shadow fading in mobile radio channel. In *IEEE PIMRC'1996*.