Fusing Noisy Fingerprints with Distance Bounds for Indoor Localization

Suining He[1] S.-H. Gary Chan[1] Lei Yu[2] Ning Liu[2]
[1]Department of CSE, The Hong Kong University of Science and Technology, Hong Kong, China
[2]School of Software, Sun Yat-sen University, Guangzhou, China
Email: [1]{sheaa, gchan}@cse.ust.hk, [2]{yulei5@mail2, liuning2@mail}.sysu.edu.cn

***Abstract*—Fusing fingerprints with mutual distance information potentially improves indoor localization accuracy. Such distance information may be spatial (e.g., via inter-node measurement) or temporal (e.g., via dead reckoning). Previous approaches on distance fusion often require exact distance measurement, assume the knowledge of distance distribution, or apply narrowly to some specific sensing technology or scenario.**

Due to random signal fluctuation, wireless fingerprints are inherently noisy and distance cannot be exactly measured. We hence propose Wi-Dist, a highly accurate indoor localization framework fusing noisy fingerprints with uncertain mutual distances (given by their bounds). Wi-Dist is a generic framework applicable to a wide range of sensors (peer-assisted, INS, etc.) and wireless fingerprints (Wi-Fi, RFID, CSI, etc.). It achieves low errors by a convex-optimization formulation which *jointly* considers distance bounds and only the first two moments of measured fingerprint signals. We implement Wi-Dist, and conduct extensive simulation and experimental studies based on Wi-Fi in our international airport and university campus. Our results show that Wi-Dist achieves significantly better accuracy than other state-of-the-art schemes (often by more than 40%).

Keywords*—*Indoor localization, convex optimization, fusion, noisy fingerprint, distance bounds, measurement uncertainty.

I. Introduction

Indoor location-based service has attracted much attention in recent years due to its commercial potential. The quality of such service largely depends on the localization accuracy of users. Among all the current indoor localization techniques, fingerprint-based approach emerges as a promising one.

Fingerprint-based indoor localization is usually conducted in two phases. In the offline (survey) phase, a site survey is conducted to collect the vectors of *received signal strength indicator (RSSI)* at reference points (RPs) of known locations. The vectors of these RSSIs form the *fingerprints* of the site and are stored at a database. In the online (query) phase, a user (or a *target*) samples an RSSI vector at his position and reports it to the server. In traditional fingerprinting, the server then compares the received vector with the stored fingerprints using some similarity metric in the signal space (like Euclidean distance in [1]). It then estimates the target position out of the RPs whose fingerprints closely match the target's RSSI (termed the "neighbors").

This work was supported, in part, by Hong Kong Research Grant Council (RGC) General Research Fund (610713), HKUST (FSGRF13EG15), and National Natural Science Foundation of China (61472455).

Error in location estimation is inevitable. This is due to random signal fluctuation in both offline and online measurements. As targets are often considered independently in the above traditional approach, such measurement noise or uncertainty may lead to a disperse set of spatially distant neighbors, which greatly degrades the localization accuracy. It has been observed that localization errors can be very high (more than 10 m [2]) under large open indoor environment such as malls, train stations or airports. This is unsatisfactory for many applications.

In order to reduce the estimation errors, one may incorporate, or *fuse*, wireless fingerprinting with mutual distance information. Embedding such information into the fingerprints can significantly reduce the dispersion, leading to substantial enhancement in localization accuracy.

The distance information can be *spatial*, where the target estimates the distances to some of the nodes or beaconing devices in its neighborhood using, for examples, Bluetooth, Wi-Fi direct, ultrasound, etc. While there have been impressive works on using spatial distance for localization, they often assume accurate distance measurement, resulting in *rigid* constraints over the fingerprints. The rigidity cannot be extended to the more realistic scenarios when distance measurement is often uncertain with given upper and lower bounds.

The distance information can also be *temporal*, where the target estimates its displacement over consecutive time instants (e.g., by step counter or inertial navigation system (INS) provided in one's mobile phone). Previous fusion works in the area are often based on Bayesian approach, assuming some probability distribution in sensor measurement. In reality, such probability distribution is often not known. Furthermore, these works cannot be easily extended to the case of noisy sensor measurement over multiple periods of time.

Due to random signal fluctuations, wireless fingerprint is inherently noisy and distance cannot be exactly measured. In this paper, we propose Wi-Dist, a novel indoor localization approach fusing noisy **wi**reless fingerprints with uncertain mutual **dist**ances given by their bounds. Wi-Dist *jointly* considers distance bounds and noisy fingerprints to reduce indoor localization errors. It requires only the first two moments (mean and variance) of the fingerprint RSSI signals, and optimizes locations based on *Semi-Definite Programming* (SDP). Using SDP relaxation, Wi-Dist solves the localization problem achieving excellent accuracy.

Because Wi-Dist takes as input only the upper and lower

978-1-4799-8381-0/15/$31.00 ©2015 IEEE

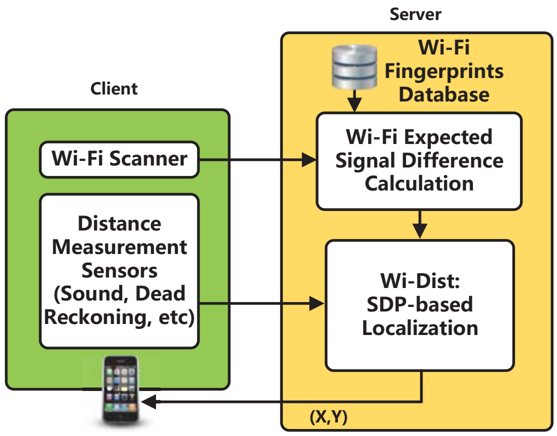

Fig. 1: System framework of Wi-Dist based on Wi-Fi fingerprints.

bounds of distances, it is not based on rigidity constraints or Bayesian approach, and hence does not require accurate distance measurement or knowledge of probability distribution. Using the bounds as constraints, Wi-Dist estimates the location by maximizing the overall similarity with the fingerprint map consisting of random signals. It is applicable to scenarios where distances may be asymmetric, in which case the upper (lower) bounds can be obtained by using the larger (smaller) value of the two directions.

Wi-Dist is a *generic* framework applicable to a wide range of sensing techniques, enabling indoor localization with adaptive spatial and temporal mobile sensing independent of how distance is measured. For example, it may employ peer-to-peer spatial distances in the crowded region. For an unpopulated area, it may then switch to dead reckoning (INS) for distance measurement. Though most of our discussion is in the context of Wi-Fi fingerprints (due to its ease of deployment without extra infrastructure beyond the existing Wi-Fi one), Wi-Dist is general enough to be extended to other wireless fingerprint signals such as RFID [3], [4] or channel state information (CSI) [5].

We show in Figure 1 the overall architecture of Wi-Dist based on Wi-Fi fingerprints. The fingerprint database is initialized by a site survey, storing pairs <*location*, *RSSI vector*> of RPs. In addition to the Wi-Fi RSSI vectors, a target measures the distance bounds and reports them to the localization server. Based on that and a difference metric for random Wi-Fi signals, the server constructs an SDP convex-optimization problem which leads to accurate localization of the target.

We have implemented Wi-Dist based on Wi-Fi fingerprints on Android platforms and performed large-scale simulation and testbed experiments in Hong Kong International Airport (HKIA) and our university campus. Using commonly used indoor sensors of dead reckoning and peer-assisted measurement, we demonstrate that Wi-Dist achieves much higher accuracy than other approaches.

The rest of this paper is organized as follows. After reviewing related work in Section II, we present the localization problem of Wi-Dist in Section III, and SDP-based localization formulation in Section IV. Illustrative results based on experimental trials and simulation are presented in Sections V and VI, respectively. We conclude in Section VII.

II. Related Work

Wi-Fi fingerprinting techniques, pioneered by Radar [6], have been widely studied in recent years. The work by Horus [7] estimates the target location using a probabilistic model which reflects the signal distribution in the site. Expectation-maximization [8], compressive sensing [9] and signal geometric patterns [10] have been implemented for fingerprint-based indoor localization. The techniques above solely address Wi-Fi fingerprint issues. We study here fusing distance information with fingerprinting to achieve much better accuracy.

Combining dead reckoning with fingerprints has been discussed in [1], [11], [12], [13]. These works assume that the walking path is conditionally independent. Therefore, the distance constraints over multiple temporal Wi-Fi estimations have not been jointly considered. Furthermore, these works treat the outputs from dead reckoning and Wi-Fi fingerprint sequentially [12]. Wi-Dist, on the other hand, formulates the localization problem as a *single joint* convex-optimization problem. This greatly reduces the influence of measurement noise and achieves higher accuracy. The work on Wi-Fi SLAM [13] implements robot odometer to determine the distance accurately. Our work differs by considering step counter whose measurement may be noisy.

There has been much work making use of Wi-Fi direct [14] and high-pitch sound [15], [16] to measure distance between devices. Some consider using a rigid graph constructed through rotation and translation [15], while others consider using Bayesian approach to infer the device location [16]. It is found that higher accuracy can be achieved when each user is 3-connected [17], or the graph satisfies some special requirements such as being pairwise connected [15]. In contrast, Wi-Dist considers jointly distance bounds and fingerprint noise, and formulates an optimization problem to estimate the target location. Therefore, Wi-Dist is a more versatile and realistic framework accommodating measurement noises.

In contrast to all the sensor fusion works above, Wi-Dist is a seamless generic framework applicable to different sensor systems with temporal or spatial distance measurement. It may be extended to different application scenarios with little modification. Wi-Dist is an optimization-based approach for indoor localization, by jointly considering both Wi-Fi signal measurement noise and distance bounds.

III. Problem Formulation and Complexity

In this section, we show how Wi-Dist estimates user locations by means of convex optimization. For concreteness, our discussion is in the context of Wi-Fi fingerprint signal (the extension to other signals is clear and straightforward). We first present the preliminaries of Wi-Dist in Section III-A. Then in Section III-B we present the objective function which is based on a novel difference metric for random Wi-Fi signals, and the problem formulation for location optimization. In Section III-C, we discuss the hardness of the problem. We show in Table I the symbols used in our problem formulation.

TABLE I: Major symbols in the problem formulation.

Notations	Definitions
M	Number of spatial or temporal targets
Q	Number of RPs in fingerprint database
$\widehat{\mathbf{x}}_m$	Estimated 2-D coordinate of target m
$\mathbf{X}$	$M \times 2$ matrix of all target locations
$\mathbf{V}$	Index set of all targets to be located
$\mathbf{Y}$	$M \times M$ matrix for transformation in SDP
$\mathbf{I}_2$	2×2 identity matrix
$\mathbf{r}_q$	2-D coordinate of reference point (RP) q
$\mathbf{R}$	$2 \times Q$ matrix of RPs
ω_{mq}	Weight of RP q to estimate target m
$\mathbf{W}$	$M \times Q$ matrix of weights at RPs
$\mathbf{\Lambda}_m$	Index set of targets to be estimated in $\mathbf{V}$ with distance measurements from m
$\mathbf{\Omega}_m$	Set of distance bounds from m
L	Number of Wi-Fi APs
$\mathbf{\Psi}_q$	RSSI vector received at RP q
$\bar{\psi}_q^l$	Average RSSI of AP l at q (dBm)
σ_q^l	RSSI standard deviation of AP l at q (dB)
$\mathbf{\Phi}_m$	RSSI vector received at m
ϕ_m^l	RSSI of AP l at m (dBm)
$\Delta_l(\mathbf{\Phi}_m, \mathbf{\Psi}_q)$	Expected signal difference of RSSI from AP l between target m and RP q
$\Gamma(\mathbf{\Phi}_m, \mathbf{\Psi}_q)$	Overall expected signal difference between target m and RP q
δ_{mn}	Distance between target m and n (m)
$\underline{\delta}_{mn}$	Lower distance bound between m and n (m)
$\hat{\delta}_{mn}$	Upper distance bound between m and n (m)

A. Preliminaries

In the offline mode, a site survey is conducted with a total of Q reference points (RPs). Let $\mathbf{r}_q$ be the 2-D position of RP q, and $\mathbf{R}$ be a $2 \times Q$ matrix indicating the RP positions, i.e.,

$$\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_Q]. \tag{1}$$

Let $\mathcal{L}$ be the index set of the Wi-Fi access points (APs) that cover the site, i.e., $\mathcal{L} = \{1, \ldots, L\}$.

At each RP, time samples of Wi-Fi RSSI readings are collected. Due to the random nature of radio signal, multiple samples are collected in order to reduce the uncertainty in the signal measurements.

Denote the RSSI at RP q from AP l at time t as $\{\psi_q^l(t), t = 1, \ldots, S, S > 1\}$, with S being the total number of samples collected. Denote the average RSS readings from AP l, $l \in \mathcal{L}$, at RP q as $\bar{\psi}_q^l$, and the unbiased estimate of the variance of the RSS time samples for AP l at RP q as $\left(\sigma_q^l\right)^2$. Then for each RP, the unbiased estimates for the mean RSSI and its corresponding standard deviation at RP q are computed as:

$$\begin{aligned} \bar{\psi}_q^l &= \frac{1}{S}\left(\sum_{t=1}^{S} \psi_q^l(t)\right), \\ \sigma_q^l &= \sqrt{\frac{1}{S-1}\left(\sum_{t=1}^{S}\left(\psi_q^l(t) - \bar{\psi}_q^l\right)^2\right)}. \end{aligned} \tag{2}$$

Then the Wi-Fi RSSI vector at $\mathbf{r}_q$ is

$$\mathbf{\Psi}_q = \left[\bar{\psi}_q^1, \bar{\psi}_q^2, \ldots, \bar{\psi}_q^L\right], q \in \{1, 2, \ldots, Q\}, \tag{3}$$

where, by definition, $\bar{\psi}_q^l = 0$ if AP l is not detected at RP q.

In the online mode, let $\mathbf{V}$ be the set consisting of the indexes of the M target nodes (or simply targets) to be localized in the site, i.e., $\mathbf{V} = \{1, 2, \ldots, M\}$, and each of their 2-D locations to be estimated is denoted as $\widehat{\mathbf{x}}_m, m \in \mathbf{V}$. Note that these targets to be estimated can be either spatial or temporal. Let $\mathbf{X}$ be an $M \times 2$ matrix of all these points, i.e.,

$$\mathbf{X} = [\widehat{\mathbf{x}}_1, \widehat{\mathbf{x}}_2, \ldots, \widehat{\mathbf{x}}_M]^T. \tag{4}$$

For each of the targets (spatial or temporal), let ϕ_m^l be the RSSI value at target location $\widehat{\mathbf{x}}_m$ for Wi-Fi AP l, $l \in \mathcal{L}$. Similar to the RP RSSI vector, we define the target m's sampled RSSI vector as

$$\mathbf{\Phi}_m = \left[\phi_m^1, \phi_m^2, \ldots, \phi_m^L\right], m \in \mathbf{V}. \tag{5}$$

where, by definition, $\phi_m^l = 0$ if AP l is not detected at target m. Given a target m, let $\mathbf{\Lambda}_m$ be the set of its neighbors that have distance measurement with. Let δ_{mn} be the distance (spatial or temporal) between targets $\widehat{\mathbf{x}}_m$ and $\widehat{\mathbf{x}}_n$ for any $m, n \in \mathbf{V}$, i.e.,

$$\|\widehat{\mathbf{x}}_m - \widehat{\mathbf{x}}_n\|^2 = \delta_{mn}^2, \forall n \in \mathbf{\Lambda}_m, n \neq m. \tag{6}$$

Based on statistical analysis of distance (spatial or temporal), we can obtain a distance bound with high confidence level. Given the distance measurement interval, we denote $\underline{\delta}_{mn}$ as the lower bound of the measurement and $\hat{\delta}_{mn}$ as the upper bound. For each Wi-Fi target, $\mathbf{\Omega}_m$ stores the distance bounds, i.e., $\mathbf{\Omega}_m = \{[\underline{\delta}_{mn}, \hat{\delta}_{mn}]\}, \forall n \in \mathbf{\Lambda}_m$.

Given the M targets to be localized in $\mathbf{V}$, each of them contains the following information in the Wi-Dist problem:

$$\mathbf{\Pi}_m \triangleq \{m, \widehat{\mathbf{x}}_m, \mathbf{\Phi}_m, \mathbf{\Lambda}_m, \mathbf{\Omega}_m\}, m \in \mathbf{V}. \tag{7}$$

B. Problem Formulation

The RP positions $\mathbf{R}$ are used to estimate the location of the target. Let ω_{mq} be the weight assigned to RP q to locate m, so that

$$\widehat{\mathbf{x}}_m = \sum_{q=1}^{Q} \omega_{mq} \mathbf{r}_q, m \in \mathbf{V}, \tag{8}$$

where the weights $\omega_{mq}, \forall m$, satisfy

$$\sum_{q=1}^{Q} \omega_{mq} = 1, \omega_{mq} \geq 0, \forall q \in \{1, 2, \ldots, Q\}. \tag{9}$$

Let $\mathbf{W}$ be an $M \times Q$ matrix of ω_{mq}, $\mathbf{r}_q \in \mathbf{R}$, i.e.,

$$\mathbf{W} = \begin{bmatrix} \omega_{11} & \ldots & \omega_{1Q} \\ \vdots & \ddots & \vdots \\ \omega_{M1} & \ldots & \omega_{MQ} \end{bmatrix}. \tag{10}$$

Then the positions of all the targets in $\mathbf{V}$ given $\mathbf{W}$ are given by

$$\mathbf{X} = \mathbf{W}\mathbf{R}^T, \tag{11}$$

The distance between two targets in Equation (6) satisfies bound constraints, i.e., $\underline{\delta}_{mn} \leq \delta_{mn} \leq \hat{\delta}_{mn}$, or equivalently,

$$\underline{\delta}_{mn}^2 \leq \delta_{mn}^2 \leq \hat{\delta}_{mn}^2, [\underline{\delta}_{mn}, \hat{\delta}_{mn}] \in \mathbf{\Omega}_m, \forall m \in \mathbf{V}. \tag{12}$$

Given the above, we present in the following the objective function for Wi-Dist problem. We first introduce a metric to evaluate the difference between the target Wi-Fi samples and the stored fingerprints under measurement noise. (Device heterogeneity in RSSI of online and offline measurement is outside the scope of this paper; interested readers are referred to works like [18] for more information on how to address it.)

We consider fingerprint noise at each RP. Define $\mathbf{J}_{mq}$ as the shared APs between Wi-Fi measurement point m and RP q ($0 < |\mathbf{J}_{mq}| \leq L$). Given a target's Wi-Fi RSSI ϕ_m^l (constant) from AP $l \in \mathbf{J}_{mq}$, the expected signal difference between RP q and the target m's RSSI in AP l is derived as [10]:

$$\begin{aligned}
\Delta_l(\mathbf{\Phi}_m, \mathbf{\Psi}_q) &\triangleq \mathrm{E}\left(\left(\phi_m^l - \psi_q^l\right)^2\right) \\
&= \mathrm{E}\left(\left(\phi_m^l\right)^2 - 2\phi_m^l\psi_q^l + \left(\psi_q^l\right)^2\right) \\
&= \left(\phi_m^l\right)^2 - 2\phi_m^l \mathrm{E}\left(\psi_q^l\right) + \mathrm{E}\left(\left(\psi_q^l\right)^2\right) \\
&= \left(\phi_m^l\right)^2 - 2\phi_m^l \mathrm{E}\left(\psi_q^l\right) + \mathrm{E}^2\left(\psi_q^l\right) + \left(\sigma_q^l\right)^2 \\
&= \left(\phi_m^l - \bar{\psi}_q^l\right)^2 + \left(\sigma_q^l\right)^2. \qquad (13)
\end{aligned}$$

By definition, if either $\phi_m^l = 0$ or $\bar{\psi}_q^l = 0$ (or both), $\Delta_l(\mathbf{\Phi}_m, \mathbf{\Psi}_q) = 0$. Thus the total expected signal difference between the RP q and the target m's RSSI vector is given by

$$\Gamma\left(\mathbf{\Phi}_m, \mathbf{\Psi}_q\right) \triangleq \frac{1}{|\mathbf{J}_{mq}|} \sum_{l=1}^{|\mathbf{J}_{mq}|} \Delta_l(\mathbf{\Phi}_m, \mathbf{\Psi}_q). \qquad (14)$$

If $|\mathbf{J}_{mq}| = 0$ (no shared APs between the target m and RP q), we have by definition $\Gamma\left(\mathbf{\Phi}_m, \mathbf{\Psi}_q\right) = \infty$, i.e., RP q is essentially excluded from the later optimization formulation.

Using Equation (14), we present in the following the objective function for Wi-Dist. To jointly measure the overall difference of all targets with the stored signal map, we find the weights which minimize all the targets' weighted sum of expected signal difference as:

$$\arg\min_{\mathbf{W}} \sum_{m=1}^{M} \sum_{q=1}^{Q} \Gamma(\mathbf{\Phi}_m, \mathbf{\Psi}_q)\omega_{mq}, \qquad (15)$$

which jointly considers the signal difference and the physical distance constraints.

To summarize, we are to find a matrix $\mathbf{W}$ so as to satisfy

$$\begin{aligned}
&\text{Objective:} \quad \text{Equation (15)}, \\
&\text{subject to:} \quad \text{Constraints (6), (9), (11) and (12).}
\end{aligned} \qquad (16)$$

C. Problem Hardness

In this section we generalize the problem given by Formulation (16) and study its hardness. (In the next section we will apply relaxation to solve it by semi-definite programming.)

Formulation (16) can be generalized into the following problem:

Definition 1. *Given the targets with RSSI measurements and the distance information between them, is there a set of locations in the fingerprint database such that the total difference between the measured RSSI vectors and the stored signal map is minimized while their relative locations satisfy the measured distance bounds?*

To prove the hardness in Definition 1, we introduce the subset sum problem (SSP), which is stated as follows:

Definition 2. *Given a set* $\mathbf{A}$ *of integer numbers and an integer number* a*, does there exist a subset of* $\mathbf{A}$ *such that the sum of its elements is equal to* a*?*

In reality, solving the problem in Formulation (16) is challenging. Here we are to prove that it is computationally hard by reduction from subset sum problem (SSP).

Theorem 1. *There is no efficient algorithm that solves the problem given by Definition 1 unless* $P = NP$.

Proof: We briefly describe our proof as follows. Suppose we have a polynomial-time algorithm that takes as input the distances between different targets as well as fingerprint measurement points to recover their original positions. Therefore, we minimize the overall sum of target signal difference with the fingerprint map to obtain the candidate locations, and then find those among them which satisfy the corresponding distance bounds.

Then such an algorithm can be used to solve the SSP by applying it to an instance of the problem. After reaching its polynomial time bound, the algorithm will either have returned a solution or not. In the first case, we can check if the solution with pairwise distances returned is consistent with the distance bounds. It is like that in the SSP we check the sum of elements in polynomial time and accept if and only if the check succeeds. In the second case, we can reject the instance. For both cases, we have returned the correct answer for SSP. Since SSP is already NP-hard, our problem is as hard as the SSP. Thus, the problem in Definition 1 is NP-hard. ■

IV. Wi-Dist: SDP-based Location Optimization

As the Wi-Dist problem is NP-hard, we use Semi-Definite Programming (SDP) relaxation [19] to solve it. SDP has been applied in wireless communication [20] and sensor networks [21]. In this work, we implement SDP to fuse the noisy fingerprints and distance bounds for indoor localization.

Given the distance bounds, we apply semi-definite relaxation [21] to relax the distance constraints. Let $\mathbf{e}_{mn}$ be an $M \times 1$ column vector where the m-th element is 1 and n-th element is -1. The physical distance between node m and n can be therefore represented as

$$\delta_{mn}^2 = \mathbf{e}_{mn}^T \mathbf{X}\mathbf{X}^T \mathbf{e}_{mn}, [\underline{\delta}_{mn}, \hat{\delta}_{mn}] \in \mathbf{\Omega}_m. \qquad (17)$$

Denote an $M \times M$ matrix $\mathbf{Y}$ for internal transformation, i.e.,

$$\mathbf{Y} = \mathbf{X}\mathbf{X}^T. \qquad (18)$$

Finally the distance bound can be rewritten as

$$\underline{\delta}_{mn}^2 \leq \mathbf{e}_{mn}^T \mathbf{Y} \mathbf{e}_{mn} \leq \hat{\delta}_{mn}^2, [\underline{\delta}_{mn}, \hat{\delta}_{mn}] \in \mathbf{\Omega}_m. \qquad (19)$$

Based on the distance bound constraints in Equation (19), we can rewrite Formulation (16) into

$$\begin{aligned}
&\text{Objective:} \quad \text{Equation (15)}, \\
&\text{subject to:} \quad \text{Constraints (9), (11), (18) and (19).}
\end{aligned} \qquad (20)$$

Clearly, Constraint (19) is nonconvex due to $\underline{\delta}_{mn}^2 \leq \mathbf{e}_{mn}^T \mathbf{Y} \mathbf{e}_{mn}$. Given a symmetric matrix $\mathbf{A}$, let $\mathbf{A} \succeq 0$ represent that $\mathbf{A}$ is a positive semidefinite matrix [19]. We can then relax this problem into a convex one by replacing the nonconvex equality constraint, $\mathbf{Y} - \mathbf{X}\mathbf{X}^T = 0$ in Constraint (18), with a convex positive semi-definite constraint, i.e.,

$$\mathbf{Y} - \mathbf{X}\mathbf{X}^T \succeq 0. \tag{21}$$

Constraint (21) is a nonlinear constraint, which can be further transformed into a linear matrix inequality [19]. Then it can be solved efficiently by a convex optimization solver. The transformation is through a Schur complement:

Definition 3. *Let* $\mathbf{H}$ *be a matrix partitioned in four blocks, consisting of four matrices* $\mathbf{B}$,$\mathbf{E}$,$\mathbf{C}$ *and* $\mathbf{D}$*, i.e.,*

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \tag{22}$$

where $\mathbf{B}$ *and* $\mathbf{D}$ *are symmetric and nonsingular. The Schur complement of* $\mathbf{D}$ *in* $\mathbf{H}$*, is defined as*

$$\mathbf{S} = \mathbf{B} - \mathbf{E}\mathbf{D}^{-1}\mathbf{C}. \tag{23}$$

If $\mathbf{S} \succeq 0$, then $\mathbf{H} \succeq 0$ [19]. Thus, by using Schur complement, we can rewrite Constraint (21) as a matrix form, i.e.,

$$\begin{bmatrix} \mathbf{Y} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{I}_2 \end{bmatrix} \succeq 0. \tag{24}$$

Then Formulation (20) is finally transformed into an SDP problem [19]:

$$\begin{aligned} &\text{Objective:} && \text{Equation (15),} && (25) \\ &\text{subject to} && \text{Constraints (9), (11), (18), (19) and (24).} \end{aligned}$$

The formulation above can be directly applied in peer-assisted localization. For dead-reckoning based localization, we may take m as the time stamp.

We end by analyzing the computational complexity of the solution. Given Q RPs and L APs, the complexity of signal difference calculation is $\mathcal{O}(QL)$. Given M temporal or spatial target measurements (usually M is small), the computation of SDP relaxation is bounded by $\mathcal{O}(M^3Q^3)$ [22]. Using some commercial SDP solver this problem can be solved efficiently [19]. Further computational reduction can be achieved by AP filtering and RP cluster mapping [9].

V. Experimental Evaluation

We have developed Wi-Dist based on Wi-Fi fingerprints in Android platforms and conducted experiments to study its performance. In this section, we first discuss the experimental settings and performance metrics in Section V-A. Then we present illustrative experimental results for dead reckoning and peer-assisted localization in Sections V-B and V-C, respectively.

A. Experimental Settings and Performance Metrics

We evaluate Wi-Dist in the Hong Kong International Airport (HKIA) boarding area and HKUST campus atrium. In the airport we collect overall $1,400$ RPs in $8,000$ m^2 area. On the campus we collect 394 RPs in $5,000$ m^2 area. Figure 2(a) and Figure 2(b) show their corresponding floor plans.

In the airport and the campus, we take overall 80 Wi-Fi samples at each RP using HTC One X+. A quarter of these samples are collected when we are facing north, south, west and east respectively. For all the application scenarios, we use the following parameters as baseline: 5 m survey grid size; 1 Wi-Fi RSSI sample is used for each target; no Wi-Fi AP reduction is conducted over target RSSI vectors.

We compare Wi-Dist with the following typical localization schemes in our experiments:

- *Fingerprint-based localization (FL)*, the classical algorithm such as [6], [2], which evaluates the Euclidean distance of target RSSI vector with each RP fingerprint and finds the interpolation of top k nearest neighbors for location estimation ($k = 15$ in our experiment).
- *Sequential Monte Carlo (SMC) localization*, a typical fusion algorithm [12] based on Sequential Monte Carlo method (particle filter) which fuses INS data and Wi-Fi fingerprinting (FL). Through the propagation along the temporal walking path, the particles translate from one location to the next. With map constraints, the spatial distribution of these particles gets corrected and resampled [11]. The final estimation is based on the weighted average of particle locations.
- *Graph-based and fingerprint localization scheme (GB + FL)*, which uses graph construction and Wi-Fi fingerprinting for peer-assisted localization. With the pairwise spatial distances of peer targets, the server constructs the rigid graph consisting of all targets [14], [15], [23]. Then the system searches against the Wi-Fi signal map and finds a set of fingerprints to minimize the objective function $\Sigma_{m=1}^{M} \|\mathbf{\Phi}_m - \mathbf{\Psi}_q\|^2$ through rotation and translation [15].

Let $\mathbf{x}_m$ be target m's true location and $\widehat{\mathbf{x}}_m$ be the estimated location. The performance metric in our experiment is the mean error (unit:m) of the estimated target in set $\mathbf{V}$:

$$\mu_e = \frac{1}{|\mathbf{V}|} \sum_{m \in \mathbf{V}} \|\mathbf{x}_m - \widehat{\mathbf{x}}_m\|. \tag{26}$$

B. Dead Reckoning

In this section we study how Wi-Dist performs for a mobile user with INS (step counter) in his smartphone.

Temporal walking distance of the target is estimated from the INS sensor which counts the steps of a walking target. Each step detection is based on the periodic changes in the vertical direction of the accelerometer readings [11]. Based on the number of steps, the distance travelled, or motion offset, can be estimated by multiplying the average stride length of the target (which is related to walking frequency as in [24]).

As the target walks, the device also collects the Wi-Fi RSSI vectors. Using the notation in Equation (7) the index m

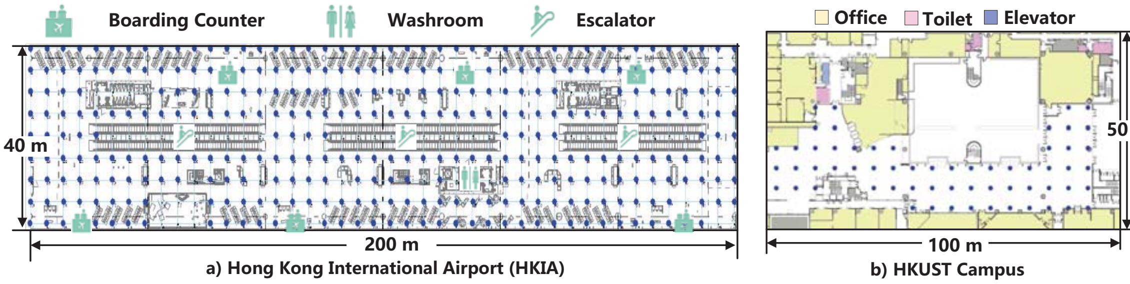

Fig. 2: (a) Map of the HKIA boarding area. (b) Map of the HKUST campus atrium. Blue points are the RPs (5 m grid size).

here represents the time stamp. Each target location $\widehat{\mathbf{x}}_m$ now corresponds to a temporal measurement of a single target. The most recent M temporal targets and $M-1$ distances between them form a sliding window in time domain, and the estimation of the M-th target is returned as the current position.

With the fast Wi-Fi scanning on smartphones [1], the small curvature between two consecutive Wi-Fi samples can be approximated as distance. Let β be the range of confidence interval for estimating the displacement and σ_{mn} be the statistical standard deviation based on experimental results. Given a distance measurement $\widetilde{\delta}_{mn}$ at time m from the last location with RSSI measurement ($\mathbf{\Lambda}_m = [m-1]$), each of the distance bounds in $\mathbf{\Omega}_m$ is defined as

$$\widetilde{\delta}_{mn} - \beta\sigma_{mn} \leq \delta_{mn} \leq \widetilde{\delta}_{mn} + \beta\sigma_{mn}, n = m-1, m > 1. \quad (27)$$

Based on Equation (7), a Wi-Fi temporal target $\mathbf{\Pi}_m$ at time m is defined as

$$\mathbf{\Pi}_m \triangleq \{m, \widehat{\mathbf{x}}_m, \mathbf{\Phi}_m, [m-1], \mathbf{\Omega}_m\}. \quad (28)$$

The distance bound for initial or the first target in the sliding window is defined to be null. Based on the empirical test, β is set to 2 in Equation (27). The size of sliding window M is 7. 200 particles are used in the particle filter of SMC algorithm.

Figure 3 plots the localization accuracy with respect to time for Wi-Dist and SMC. The estimation error fluctuates as the user walks in the airport. Changes in wall partitions, crowded people, user walking direction and smartphone holding gesture introduce measurement noise in Wi-Fi and INS signals. SMC sequentially considers the fingerprints and INS measurements. It does not jointly consider the Wi-Fi fingerprints and the distances from the multiple time periods. Therefore, large error in location estimation happens. In contrast, Wi-Dist constrains its estimations through the distance bounds in a joint optimization formulation. Therefore, Wi-Dist can achieve lower localization errors and smaller estimation fluctuation.

Figure 4 shows the mean displacement measurement and corresponding standard deviation at each true walking distance. In the empirical studies, the major error of displacement comes from misestimation in step counts and step length. Meanwhile the device initialization and walking curvature also leads to additional displacement errors [1], [24]. Based on such empirical analysis, we obtain the displacement variance for each measured distance, which constitutes the distance bounds (Equation (27)) in Wi-Dist temporal measurement.

Figure 5 shows the mean localization errors against the number of Wi-Fi temporal target measurements. We can see that the accuracy improves as we utilize more temporal samples. It is because joint consideration of more periods further constrains the location estimations. When we further increase the number of measurements, the accuracy gradually converges, indicating that distance bounds already provide sufficient constraints. Thus, to balance between localization accuracy and computational complexity we choose several temporal measurements (like 7 in our experiment) in Wi-Dist.

C. Peer-Assisted Localization

For some areas visited by many users, peer-assisted (PA) localization may be used [15]. Peer ranging can be based on either RSS-distance mapping or sound ranging. In the experiment, we implement and test sound ranging under quiet and noisy campus environment. The mean peer ranging errors under these two conditions are 0.8 m and 2 m respectively. Since distance constraint between two peers is asymmetric due to measurement uncertainty, we use the larger value in the distance measurements as the upper bound $\hat{\delta}_{mn}$ and the smaller one as the lower bound $\underline{\delta}_{mn}$. In the peer-assisted localization, 5 targets are involved in sound-based distance measurement. We do not exclude the cases when walls may partition peers during localization.

Figure 6 shows the mean localization errors against the proportion of APs removed at targets. We randomly remove some received APs of each target to evaluate the influence of AP reduction due to wall partitioning or crowds of people. We can see that Wi-Dist and GB+FL marginally rely on the number of received APs. It is because the multiple users' Wi-Fi samples reduce the effect of sparse AP deployment. To the contrary, FL relies on the APs to differentiate the RPs and therefore its estimation error increases as more APs are pruned.

Figure 7 shows that the location estimation errors against the number of Wi-Fi samples at each target in PA localization. All the algorithms improve with more Wi-Fi samples. It is because as the number of Wi-Fi samples increases, noise from the random sampling can be reduced [7]. Compared with GB+FL and FL, Wi-Dist achieves higher localization accuracy because it further jointly considers the measurement noise in the optimization formulation and reduces the uncertainty. However, increasing the number of Wi-Fi samples means that we need to wait for more samples before final estimation. A balance has to be made between accuracy and latency depending on application scenarios.

Figure 8 shows the location errors against the survey grid size. As the minimum grid size is five meters, lines or rows of RPs are removed to form grid size with multiples of

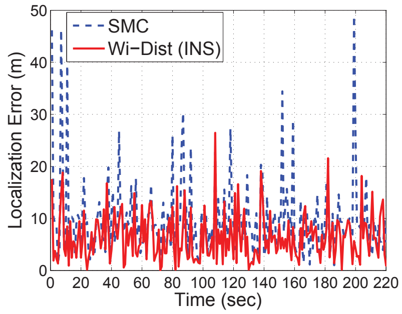

Fig. 3: Localization errors over time.

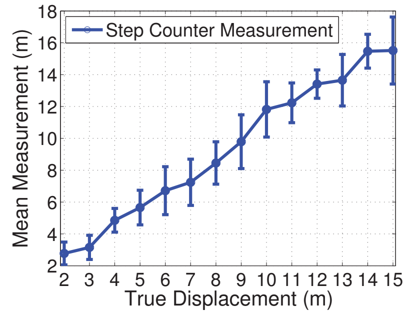

Fig. 4: Walking distance measurements using step counter.

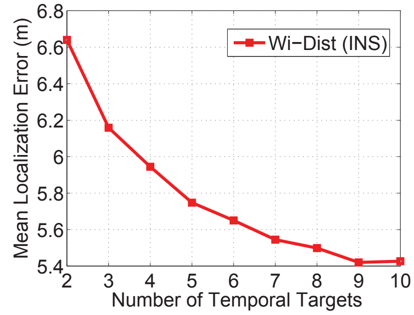

Fig. 5: Localization errors vs. number of Wi-Fi temporal targets.

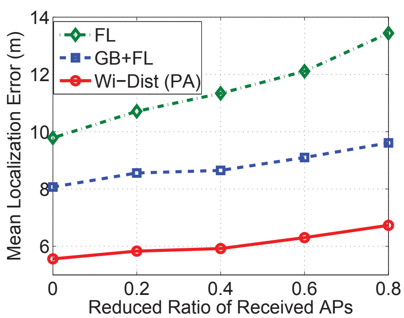

Fig. 6: Localization errors vs. removed proportion of received APs.

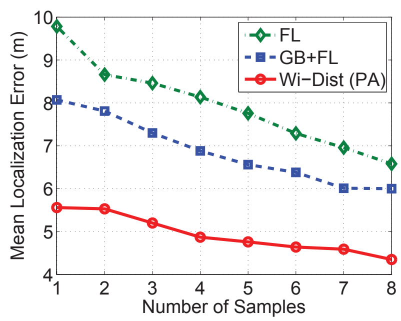

Fig. 7: Localization errors vs. number of Wi-Fi samples.

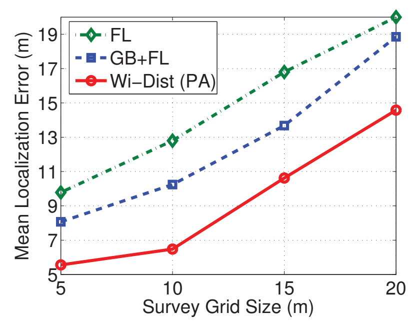

Fig. 8: Localization errors vs. survey grid size.

five. Clearly, as the grid size increases, the accuracy of the three algorithms decreases. Though less labor-intensive, larger survey grid size may more easily lead to dispersed nearest neighbors under large signal noise. Therefore, traditional algorithms like FL may not accurately differentiate these RPs. Under different survey density, Wi-Dist and GB+FL achieve more accurate location estimations with the constraints of peer-to-peer distances. However, the rigid graphs of targets in GB+FL still suffers from pairwise distance measurement noise. By fusing signal uncertainty and distance bounds, Wi-Dist achieves higher estimation accuracy under different grid sizes.

Figure 9 and Figure 10 show the overall performance of Wi-Dist at different scenarios (INS and PA at baseline parameters) in HKIA. Large indoor open space often leads to high uncertainty in Wi-Fi signals [2] and disperse nearest neighbors in signal space. Furthermore, the temporal and spatial distance measurement also contains large noise under the crowded scenarios. Compared with other state-of-the-art algorithms, Wi-Dist significantly reduces the estimation errors in HKIA. With distance constraints and joint optimization, Wi-Dist mitigates the effect of disperse nearest neighbors.

Compared with the airport, the campus atrium is smaller with more building partitions, which may influence the peer-distance measurement accuracy. We show the performance of Wi-Dist (INS and PA) on HKUST campus in Figure 11 and Figure 12 respectively. Wi-Dist achieves higher localization accuracy than the other state-of-the-art algorithms. As the results in HKUST are qualitatively similar to those in HKIA, for brevity we do not repeat other experimental results here.

VI. ILLUSTRATIVE SIMULATION RESULTS

To evaluate more comprehensively the performance of Wi-Dist in large-scale indoor environment with many users, we have simulated Wi-Dist for the scenarios as mentioned in Section V. In this section, we first discuss the simulation setup (Section VI-A), followed by the results for dead reckoning and peer-assisted localization (Sections VI-B and VI-C).

A. Simulation Setup

We simulate the Wi-Fi signal strength following the work in [25]. In the signal model, the RSSI Φ (dBm) from Wi-Fi AP at a distance D can be simulated as

$$\Phi = \Phi^{TX} - L^0 - 10\alpha \log_{10}\left(\frac{D}{D^0}\right) + \epsilon, \tag{29}$$

where measurement noise is distributed as $\epsilon \sim \mathcal{N}(0, \sigma_{db}^2)$. Unless otherwise stated, we use the following as our baseline parameters: the transmission power $\Phi^{TX} = 25$ dBm, the path loss exponent $\alpha = 4.0$, reference path loss $L^0 = 37.7$ dB, reference distance $D^0 = 1$ m, 200 m $\times$ 80 m survey site with 5 m grid size; Wi-Fi signal noise $\sigma_{db} = 6$ dB; 10 APs are uniformly distributed in the survey area; a target takes a Wi-Fi sample every 3 seconds.

B. Dead Reckoning

For dead reckoning, we use a random way-point mobility model with resting [26], and 7 most recent Wi-Fi records are used for INS fusion. The step count error rate is distributed as $\mathcal{N}(0, \sigma_r^2)$, where $\sigma_r = 20\%$, and the stride length error follows $\mathcal{N}\left(0, \sigma_l^2\right)$, where $\sigma_l = 0.2$ m. Additional walking displacement error is assumed to follow $\mathcal{N}(0, \sigma_w^2)$, where $\sigma_w = 2$ m.

Figure 13 shows the localization accuracy against the step count errors. Clearly, the performance of SMC and Wi-Dist degrades with larger step count error. SMC locates the user based on the particle filter, which sequentially considers the Wi-Fi and INS measurements. Therefore, when step count accuracy degrades, the displacement error increases and the particles become spatially sparse, making it difficult for SMC to converge to correct locations. To the contrary, Wi-Dist localizes the target more accurately because the joint consideration of Wi-Fi fingerprints and distance bounds of multiple periods reduces the influence of measurement uncertainty.

Figure 14 shows the localization accuracy against the walking displacement errors. We can see that as the displacement error increases, the overall location accuracy decreases. Localization error in SMC increases because the particles converge slowly given large distance errors and noisy Wi-Fi measurement. Different from SMC, Wi-Dist achieves more accurate results because it utilizes the distance bounds instead of actual distance measurement. By constraining the target estimation within the intersection of these bounds, Wi-Dist is more robust to distance uncertainty.

Figure 15 plots the localization errors versus the signal noise in Wi-Fi measurement (Equation (29)). We can observe that the performance of both SMC and Wi-Dist degrades when the random signal noise increases. It is because larger signal noise makes it more difficult to differentiate the fingerprints. Different from SMC, Wi-Dist considers signal uncertainty through the expected signal difference. By minimizing the signal difference within constraints of distance bounds, Wi-Dist reduces the effect of disperse nearest neighbors and obtains better estimation results.

C. Peer-Assisted Localization

We assume noisy peer-assisted distance error $\epsilon_m \sim \mathcal{N}(0, \sigma_m^2)$, σ_m = 3.5 m; neighborhood detection range is 15 m; four peers together initiate a peer-assisted localization; 90 users are randomly distributed in the survey site.

Figure 16 shows the localization errors versus the number of users. Clearly, more peer assistance provides more distance constraints over the involved users and improves the localization accuracy. Different from GB+FL, Wi-Dist shows less dependency on user connectivity. It is because Wi-Dist considers the measurement uncertainty in the optimization and jointly constrains all the users. Therefore, it does not have to involve many users to achieve high localization accuracy.

Figure 17 shows the localization accuracy against the peer distance errors. We assume a Gaussian noise is added to the inter-device distance measurement. When the peer distance error is small, both algorithms achieve high accuracy given only Wi-Fi measurement noise in fingerprints. As distance error further increases, both algorithms degrade in localization accuracy. GB+FL constructs a rigid graph to constrain relative positions of different users. However, the graph shape deforms under large distance measurement errors. Wi-Dist, in contrast, shows more robustness by using joint optimization based on fingerprints and distance bounds. Without assuming a rigid graph, Wi-Dist can achieve more robust localization estimation.

VII. Conclusion

In this paper, we have proposed Wi-Dist, a novel and convex-optimization framework fusing wireless fingerprints with mutual distance information for indoor localization. The mutual distance can be temporal or spatial between different target measurements (as obtained from dead reckoning or peer-assisted manner). Due to random signal fluctuation, fingerprints are noisy in nature and distance cannot be measured exactly. Wi-Dist formulates a single semi-definite programming (SDP) problem which fuses noisy fingerprints with uncertain distance measurement, where the fingerprint noise is considered through only its first two moments while the distance needs only upper and lower bounds. Wi-Dist is generic, and hence is applicable to a wide range of sensing devices and wireless fingerprint signals.

We have conducted extensive simulation and experimental trials based on Wi-Fi fingerprints in our Hong Kong International Airport and university campus. We implement Wi-Dist using INS (temporal distance) and peer-assisted distance measurement (spatial distance). Our results show that Wi-Dist can significantly improve Wi-Fi localization accuracy, often achieving substantial improvement as compared with other state-of-the-art algorithms (40%).

References

[1] W. Sun, J. Liu, C. Wu, Z. Yang, X. Zhang, and Y. Liu, "MoLoc: On distinguishing fingerprint twins," in *Proc. IEEE ICDCS*, Jul. 2013.

[2] D. Han, S. Jung, M. Lee, and G. Yoon, "Building a practical Wi-Fi-based indoor navigation system," *IEEE Pervasive Computing*, vol. 13, no. 2, pp. 72–79, 2014.

[3] W. Zhuo, B. Zhang, S. Chan, and E. Chang, "Error modeling and estimation fusion for indoor localization," in *Proc. IEEE ICME*, July 2012, pp. 741–746.

[4] X. Guo, D. Zhang, K. Wu, and L. Ni, "MODLoc: Localizing multiple objects in dynamic indoor environment," *IEEE Trans. Parallel and Distributed Systems*, vol. 25, no. 11, pp. 2969–2980, Nov 2014.

[5] K. Wu, J. Xiao, Y. Yi, D. Chen, X. Luo, and L. Ni, "CSI-based indoor localization," *IEEE Trans. Parallel and Distributed Systems*, vol. 24, no. 7, pp. 1300–1309, July 2013.

[6] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, 2000.

[7] M. Youssef and A. Agrawala, "The Horus location determination system," *Wireless Networks*, vol. 14, no. 3, pp. 357–374, Jun. 2008.

[8] A. Goswami, L. E. Ortiz, and S. R. Das, "WiGEM: A learning-based approach for indoor localization," in *Proc. ACM CoNEXT*, 2011, pp. 3:1–3:12.

[9] C. Feng, W. Au, S. Valaee, and Z. Tan, "Received-signal-strength-based indoor positioning using compressive sensing," *IEEE Trans. Mobile Computing*, vol. 11, no. 12, pp. 1983–1993, Dec 2012.

[10] S. He and S.-H. Chan, "Sectjunction: Wi-Fi indoor localization based on junction of signal sectors," in *Proc. IEEE ICC*, June 2014, pp. 2605–2610.

[11] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proc. ACM MobiCom*, 2012, pp. 293–304.

[12] Y. Gao, Q. Yang, G. Li, E. Y. Chang, D. Wang, C. Wang, H. Qu, P. Dong, and F. Zhang, "XINS: The anatomy of an indoor positioning and navigation architecture," in *Proc. MLBS (ACM UbiComp Workshop)*, 2011, pp. 41–50.

[13] B. Ferris, D. Fox, and N. D. Lawrence, "WiFi-SLAM using gaussian process latent variable models." in *IJCAI*, vol. 7, 2007, pp. 2480–2485.

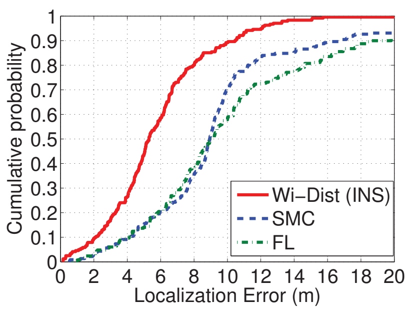

Fig. 9: Performance of Wi-Dist (INS) in HKIA.

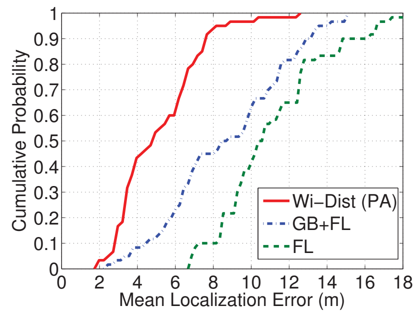

Fig. 10: Performance of Wi-Dist (PA) in HKIA.

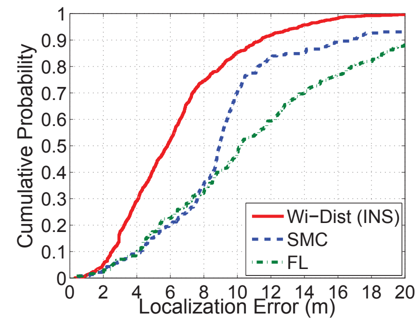

Fig. 11: Performance of Wi-Dist (INS) on HKUST campus.

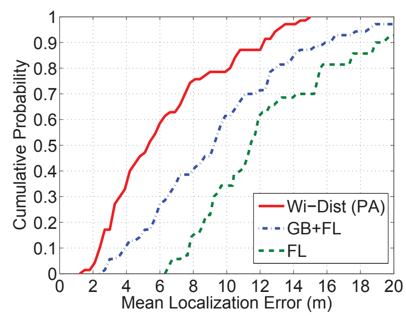

Fig. 12: Performance of Wi-Dist (PA) on HKUST campus.

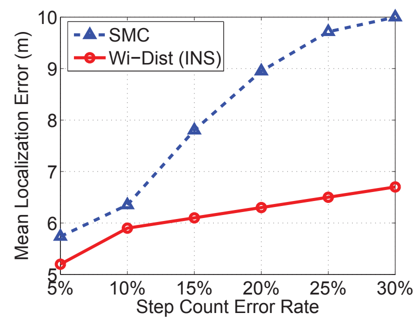

Fig. 13: Localization errors vs. INS step counts error rate.

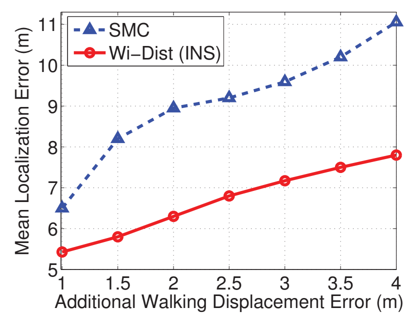

Fig. 14: Localization errors vs. additional walking displacement errors.

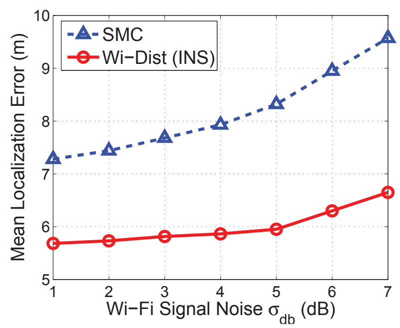

Fig. 15: Localization errors vs. Wi-Fi signal noise.

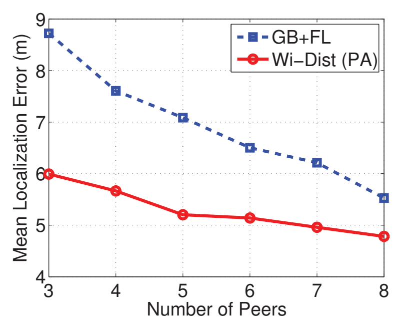

Fig. 16: Localization errors vs. peer user number.

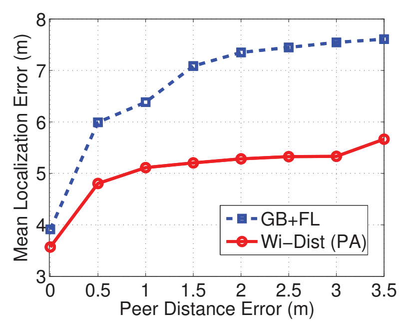

Fig. 17: Localization errors vs. peer distance errors.

[14] N. Banerjee, S. Agarwal, P. Bahl, R. Chandra, A. Wolman, and M. Corner, "Virtual compass: Relative positioning to sense mobile social interactions," in *Proc. Pervasive*. Springer-Verlag, Mar. 2010.

[15] H. Liu, J. Yang, S. Sidhom, Y. Wang, Y. Chen, and F. Ye, "Accurate WiFi based localization for smartphones using peer assistance," *IEEE Trans. Mobile Computing*, vol. 13, no. 10, pp. 2199–2214, Oct 2014.

[16] R. Nandakumar, K. K. Chintalapudi, and V. N. Padmanabhan, "Centaur: Locating devices in an office environment," in *Proc. ACM MobiCom*, 2012, pp. 281–292.

[17] J. Aspnes, T. Eren, D. K. Goldenberg, A. S. Morse, W. Whiteley, Y. R. Yang, B. D. Anderson, and P. N. Belhumeur, "A theory of network localization," *IEEE Trans. Mobile Computing*, vol. 5, no. 12, pp. 1663–1678, 2006.

[18] A. Mahtab Hossain, Y. Jin, W.-S. Soh, and H. N. Van, "SSD: A robust RF location fingerprint addressing mobile devices' heterogeneity," *IEEE Trans. Mobile Computing*, vol. 12, 2013.

[19] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[20] X. Fan, J. Song, D. P. Palomar, and O. C. Au, "Universal binary semidefinite relaxation for ML signal detection," *IEEE Trans. Communications*, vol. 61, no. 11, pp. 4565–4576, November 2013.

[21] S. Ji, K.-F. Sze, Z. Zhou, A. M.-C. So, and Y. Ye, "Beyond convex relaxation: A polynomial-time non-convex optimization approach to network localization," in *Proc. IEEE INFOCOM*, 2013, pp. 2499–2507.

[22] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," *SIAM Journal on Optimization*, vol. 10, no. 2, pp. 443–461, 2000.

[23] F. Dabek, R. Cox, F. Kaashoek, and R. Morris, "Vivaldi: A decentralized network coordinate system," in *Proc. ACM SIGCOMM*, Portland, OR, USA, Aug. 2004, pp. 15–26.

[24] F. Li, C. Zhao, G. Ding, J. Gong, C. Liu, and F. Zhao, "A reliable and accurate indoor localization method using phone inertial sensors," in *Proc. ACM UbiComp*, 2012, pp. 421–430.

[25] N. Alsindi, R. Raulefs, and C. Teolis, *Geolocation Techniques: Principles and Applications*. Springer, 2012.

[26] J. Jun, Y. Gu, L. Cheng, B. Lu, J. Sun, T. Zhu, and J. Niu, "Social-Loc: Improving indoor localization with social sensing," in *Proc. ACM SenSys*, 2013, pp. 14:1–14:14.