Maxlifd: Joint Maximum Likelihood Localization Fusing Fingerprints and Mutual Distances

Suining He, *Member, IEEE*, S.-H. Gary Chan, *Senior Member, IEEE*, Lei Yu, and Ning Liu

Abstract—Fusing mutual distance information with fingerprints can substantially improve indoor localization accuracy. Such distance information may be spatial (e.g., measurement among users or from installed beaconing devices) or temporal (e.g., via dead-reckoning). Previous approaches on distance-fusion often require deterministic distance measurement, consider fingerprints and distances separately, or are narrowly applicable to some specific sensing technology or scenario. Given the fact that fingerprint and distance measurements are intrinsically random, we propose *Maxlifd*, an accurate indoor localization framework fusing fingerprints and distances of arbitrary distributions via joint maximum likelihood. Maxlifd is a generic statistical/probabilistic framework applicable to a wide range of sensors (peer-assisted, INS, iBeacon, etc.) and fingerprints (Wi-Fi, RFID, etc.). It achieves low localization errors by a novel optimization formulation *jointly* considering mutual distances and fingerprint signals. Using generic probabilistic formulation, we further derive the lower bound on localization error for comprehensive performance analysis. We have implemented Maxlifd, and conducted extensive simulation and experimental trials in an international airport and our university campus. Our results show that Maxlifd achieves significantly lower errors than other state-of-the-art schemes (often by more than 30 percent). We experimentally verify that its performance does not depend sensitively on the exact knowledge of the underlying distributions beyond simple Gaussian.

Index Terms—Indoor localization, fusion, joint maximum likelihood, convex optimization, semi-definite programming, noisy fingerprints, mutual distances, measurement uncertainty

1 Introduction

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

Fingerprint-based indoor localization is usually conducted in two phases [1]. In the offline (survey) phase, a site survey is conducted to collect the vectors of *received signal strength indicator (RSSI)* at known locations, the so-called reference points (RPs). The vectors of these RSSIs form the *fingerprints* of the site and are stored in a database. In the online (query) phase, a *user* (also known as *target*) samples an RSSI vector at his location and reports it to the server, which estimates and returns the user location. (Another possibility is to compute the location locally at mobiles, if the fingerprints have been downloaded.)

Error in location estimation is inevitable. This is due to random signal fluctuation in both offline and online measurements. It has been observed that localization error can be quite high (more than 10 m [1], [2], [3]) under large open indoor environment such as malls, train stations or airports. This is unsatisfactory for many applications. In order to substantially enhance accuracy, one may embed, or *fuse*, fingerprinting with mutual distance information.

Mutual distance information can be *spatial*, where the target estimates the distance to other users or beaconing devices in its neighborhood. This can be done using different sensing techniques: Bluetooth [4], Wi-Fi direct [5], sound [6], etc. While there have been impressive works on using spatial distance for localization, they often assume accurate distance measurements, resulting in *rigid* constraints over the fingerprints. These approaches cannot be easily extended or applied to the more realistic scenarios when distance estimation is random.

Besides spatial, the distance information can also be *temporal*, where the target estimates its displacement over time (e.g., by the step counter or inertial navigation system (INS) provided in one's mobile phone). Previous approaches (like particle filter [3], [7] or Hidden Markov Model [8], [9]) in this setting often treat Wi-Fi fingerprints and pedestrian distance measurements as separate, and hence sequentially consider them. This consideration of independence may be hardly satisfactory for accurate localization, as measured fingerprints and distances are inherently correlated (coupled) and better *jointly optimized*. Given noisy/random sensor measurement, prior works may not fully address the issue of large positioning error.

Due to signal fluctuations, wireless fingerprint is inherently noisy, and estimated distance is random. In this paper, we propose *Maxlifd*, a novel **max**imum **li**kelihood framework fusing noisy wireless **f**ingerprints and random mutual **d**istances to achieve highly accurate indoor localization. Maxlifd *jointly* considers fingerprints and distances, both of arbitrary

- *S. He and S.-H. G. Chan are with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong, China. E-mail: {sheaa, gchan}@cse.ust.hk.*
- *L. Yu and N. Liu are with the School of Software, Sun Yat-sen University, Guangzhou, Guangdong 510275, China. E-mail: yulei5@mail2.sysu.edu.cn, liuning2@mail.sysu.edu.cn.*

Manuscript received 25 Apr. 2016; revised 5 Feb. 2018; accepted 16 May 2018. Date of publication 29 May 2018; date of current version 4 Feb. 2019. (Corresponding author: Ning Liu.)
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TMC.2018.2841842

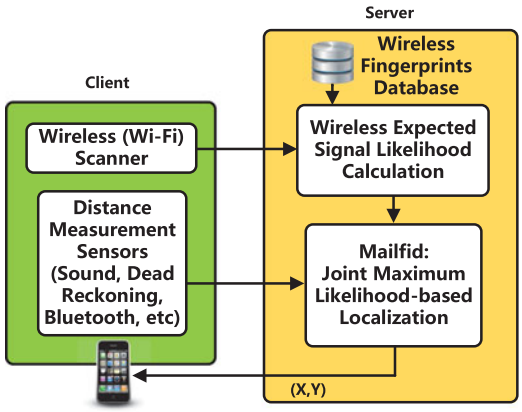


Fig. 1. System architecture of Maxlifd.

distributions. After formulating the location estimation in a *joint likelihood* problem, Maxlifd solves it using novel *semidefinite programming* (SDP), achieving excellent accuracy.

Unlike previous works, Maxlifd is a *generic* statistical or probabilistic framework applicable to a wide range of sensing techniques and scenarios. As a framework, it enables indoor localization combining both spatial and temporal mobile sensing independent of how distance is measured. For example, a target may employ opportunistic peer-to-peer spatial distances in a crowded region. For an unpopulated area, it may switch to dead reckoning (INS) for distance measurement. Meanwhile, the target may also estimate its distances to some fixed beaconing devices (if any).

Based on the *joint maximum likelihood framework* of Maxlifd, we have derived the lower bound on the localization error using *Cramér Rao Lower Bound* (CRLB) [10]. Such derivation enables us to understand the impact of various system parameters over the localization error, and sheds insights on the system design. For example, we show that increasing the number of collaborative users in peer-assisted localization or deploying more iBeacons beyond a certain point would not significantly improve the localization accuracy. We also show effects of other factors like signal noise and distance misestimation over the localization error.

We show in Fig. 1 the overall architecture of Maxlifd based on wireless fingerprints. The fingerprint database is initialized by a site survey, storing the couplets $<$*location*, *RSSI vector*$>$ for each RP. In addition to the RSSI vectors, a target measures the mutual distances (spatial or temporal) and reports them to the localization server. Based on that, the server constructs an SDP-based convex-optimization framework and solves a joint maximum likelihood problem. It then returns the absolute location result to the target.

Maxlifd is applicable to any wireless fingerprint. Though most of our discussion in this paper is on Wi-Fi fingerprints (due to its ease of deployment without extra infrastructure beyond the existing one) for prototype studies, Maxlifd is general enough to be extended to many other existing or emerging fingerprint signals [11], such as RFID and channel state information (CSI), to achieve higher accuracy.

The contributions of our paper are as follows:

- We present a novel and general joint maximum-likelihood-based framework to fuse fingerprints and distances of arbitrary distributions. Beyond the previous fusion systems, Maxlifd generalizes variant distance sensing scenarios, laying a more comprehensive foundation for theoretical and experimental performance analysis.
- We formulate a novel SDP-based optimization problem to fuse fingerprints and distances for indoor location estimation. Using the novel localization framework, we conduct extensive CRLB analysis on the lower bound of localization error;
- We have implemented Maxlifd based on Wi-Fi fingerprints on Android platforms. We have performed large-scale simulation, and conducted testbed experiments in Hong Kong International Airport (HKIA) and our university campus. Using the commonly available indoor sensors of dead reckoning, peer-assisted and beacon-based measurement, we show that Maxlifd achieves much higher accuracy than other state-of-the-art approaches. Furthermore, the performance does not depend sensitively on the exact knowledge of signal or distance distribution, and Gaussian distribution may be applied for simplicity.

The rest of this paper is organized as follows. After reviewing related work in Section 2, we present in Section 2 the framework of Maxlifd based on noisy fingerprint and random distance estimation, followed by an SDP-based localization formulation for the example of Gaussian distribution in Section 4. Given the formulation, we present in Section 5 the CRLB of Maxlifd for Gaussian measurement. Illustrative results based on experimental trials and further simulation studies are discussed in Sections 6 and 7, respectively. We discuss some important implementation issues in Section 8, and conclude in Section 9.

2 Related Work

Fingerprinting techniques [12], pioneered by Radar [1], have been widely studied in recent years. The work by Horus [13] estimates the target location using a probabilistic model, which reflects the signal distribution in the site. Kernel method [14], [15] and compressive sensing [16] 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. Fingerprints provide coarse location information, while mutual relative distances help fine-grain the final results. Furthermore, most recent works only analyze the statistical performance of traditional Wi-Fi fingerprinting localization [17] or distance-based sensor network positioning [18], [19], [20], while fusing noisy fingerprints with variants of statistical sensor distance measurements into a single but generic probabilistic formulation with error bound analysis, has not been comprehensively studied before. Filling such a gap would help the following researchers and practictioners to better understand, interpret and design such fusion systems.

Combining dead reckoning with RF signals has been discussed in some literature. Some of the works, like particle filter, treat multiple temporal Wi-Fi estimations as independent, and have not *jointly* considered their estimation errors for better localization [3]. Some works [21] require extensive motion, trajectory and map information for sophisticated classification model training. Some others treat the outputs from dead reckoning and Wi-Fi fingerprints *sequentially* through particle filter [3], [7] or Markov model [8], [22] rather than jointly, at the cost of lessened useful information. Maxlifd, on the other hand, formulates the localization

problem as a *single joint* convex-optimization problem given distribution of temporal walking distance. This formulation greatly reduces the impact of measurement noise and achieves much higher accuracy.

Our study here is orthogonal to simultaneous localization and mapping (SLAM). In WiFi-SLAM algorithm [23], the state space is modeled by a Gaussian Process (Latent Variable) model. Different from WiFi-SLAM, our Maxlifd considers fusion of fingerprints and distances into one single probabilistic framework, and solves it efficiently through semidefinite programming. Furthermore, SLAM approaches are considered as orthogonal to our work, as Maxlifd considers a novel sensor fusion given a basic floor and signal map. On the other hand, SLAM-based approaches, such as UnLoc [24], can serve as an initial provider of indoor map and fingerprint signals.

There has been some work making use of Wi-Fi direct [5] and high-pitch sound [6], [25] to measure spatial distances between devices and improve localization. Some consider using a rigid graph [5] constructed through rotation and translation [6], while others consider using Bayesian approach to infer the device location [25]. Assuming static devices in office environment, Centaur [25] utilizes the sound to estimate device positions. In contrast, Maxlifd jointly considers distance uncertainty and fingerprint noise, and formulates the joint maximum likelihood problem to estimate the target location. Maxlifd is a much more versatile and generic framework accommodating measurement noises and different application scenarios.

Some beacon-based techniques, pioneered by iBeacon, have recently been proposed for site monitoring and indoor localization [26], [27]. The work in [26] studies the advantages of using iBeacons for indoor localization. The work in [27] investigates the performance of iBeacon fingerprinting. We propose here a new joint optimization framework combining both wireless fingerprints and different emerging beacon techniques (say, iBeacons, AoA [28], [29], ToA [30], Wi-Fi sniffers, sound anchors [31], etc.). Our prototyping with iBeacon has further shown that such formulation achieves much higher accuracy with general applicability.

Based on our deployment practice, we consider spacious and commercial application scenarios (like airport or shopping mall) where acts of adding many new anchor nodes or modifying existing WLAN APs belonging to many different parties of the site are more or less regulated. Thus, we leverage the fingerprint-based technique (where low-cost fingerprints can be collected via crowdsourcing [32]), and beyond this we further embrace any extra information of distances. Specifically, coarse but pervasive fingerprints anchor the rough location of the user, while variant accurate ranging information (e.g., pedometer, phone-to-phone ranging or beacon) is encountered opportunistically to constrain or refine the estimation.

In our joint optimization, we consider the mutual Euclidean distance and have a quadratic objective function. To reduce the computation complexity from the quadratic form, we apply semi-definite relaxation [33] to relax the distances. Semi-definite programing has been applied in wireless communication [34] and sensor networks [33]. In this paper, we implement SDP to fuse the noisy fingerprints and distance distributions for indoor localization. Note that, unlike multidimensional scaling (MDS) and spectral graph drawing (SGD) [35], Maxlifd formulation considers fingerprint matching with distances, and relies less on high sensor connectivity [36], [37]. Hence Maxlifd is more applicable to complex signal measurements in indoor localization.

TABLE 1
Major Symbols in the Maxlifd Formulation

Notations	Definitions
M, $\mathbf{V}$	Number of spatial/temporal targets and their index set
Q	Number of reference points (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
$\mathbf{Y}$	$M \times M$ matrix for transformation in SDP
$\mathbf{r}_q$	2-D coordinate of reference point (RP) q
$\mathbf{R}$	$2 \times Q$ coordinate 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 target m
Ω_m	Set of distances (spatial or temporal) from target m
L	Number of Wi-Fi APs in the site of interest
Ψ_q	RSSI vector received at RP q
$\bar{\psi}_l^q$	Average RSSI of AP l at RP q (dBm)
σ_l^q	RSSI standard deviation of AP l at RP q (dB)
Φ_m	RSSI vector received at target m
ϕ_m^l	RSSI of AP l at target m (dBm)
δ_{mn}	Distance measurement between target m and n (m)
$\mathcal{I}_{mn}$	Indicator on existence of δ_{mn} between m and n

A preliminary version of this work, Wi-Dist, has been reported in [38], which estimates location by fusing fingerprints and distance bounds. In this paper, we advance from Wi-Dist in several major ways: 1) Maxlifd jointly accounts for the measurement noise in both fingerprints and distance estimation, given any arbitrary distribution. As distance bound is a special case of our formulation, Maxlifd is more general; 2) Maxlifd uses joint maximum likelihood to localize the target, which is much more robust to signal noise and hence more accurate; 3) We reformulate Maxlifd through SDP relaxation, and solve it efficiently; 4) We further analyze the lower bound of localization error based on CRLB; (5) We conduct more simulation and experimental studies.

3 Preliminaries & Statistical Framework

In this section, we show the preliminaries and statistical framework on how Maxlifd estimates user locations. For concreteness, our discussion is in the context of Wi-Fi fingerprint signal (extension to other signals is clear and straightforward). We first present the preliminaries of Maxlifd in Section 3.1. Then in Section 3.2, we present the statistical or probabilistic framework of Maxlifd. Table 1 shows some of the important symbols used in our formulation.

3.1 Preliminaries of Fingerprint-Based Localization

In the offline phase, a site survey is conducted with a total of Q 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 $\mathbf{L}$ be the index set of the Wi-Fi access points (APs) that cover the site, i.e., $\mathbf{L} = \{1, \ldots, L\}$.

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 $\{\psi_q^l(t), t = 1, \ldots, S_q^l, S_q^l > 1\}$, with S_q^l being the total number of samples collected. Denote the average RSS readings from AP l, $l \in \mathbf{L}$, 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 $(\sigma_q^l)^2$. Then for each RP, the unbiased estimates for the mean RSSI and its corresponding standard deviation at RP q are

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

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

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

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

In the online phase, 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 locations 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, $l \in \mathbf{L}$. Similar to the RP RSSI vector, we define the target m's sampled RSSI vector as

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

where, by definition, $\phi_m^l = 0$ if AP l is not detected at target m. Given a target m, let Λ_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.,

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

As the measurements of mutual distances often contain noise, we may evaluate the difference (or relationship) between $\|\widehat{\mathbf{x}}_m - \widehat{\mathbf{x}}_n\|^2$ and δ_{mn} according to a certain distribution. Then we obtain the likelihood of δ_{mn} given $\widehat{\mathbf{x}}_m$ and $\widehat{\mathbf{x}}_n$, denoted as $p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n)$. $p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n)$ can be given by a certain distribution formulation. Note that for the temporal pedestrian localization, δ_{mn} exists if m and n are two consecutive targets (i.e., walking distance between two successive Wi-Fi RSSI measurements). For spatial scenarios, we consider δ_{mn} exists if two devices (sensors) are within a certain detection range.

Based on statistical analysis of distance (spatial or temporal), we can obtain the standard deviation of the distance measurement, i.e., σ_{mn} to generalize the inherent uncertainty. Therefore, for each target m, we store the corresponding standard deviation in its distance measurement as $\Omega_m = \{\sigma_{mn}\}, \forall n \in \Lambda_m$.

To summarize, given the M targets (temporal or spatial) to be localized in $\mathbf{V}$, each of them contains the following information in the Maxlifd problem:

$$\Pi_m \triangleq \{m, \widehat{\mathbf{x}}_m, \Phi_m, \Lambda_m, \Omega_m\}, \quad m \in \mathbf{V}. \tag{7}$$

3.2 Statistical Localization Framework

The RP positions $\mathbf{R}$ are used to estimate the locations of the targets. Similar to traditional weighted average estimation in K-NN [1], [2], let ω_{mq} be the weight assigned to RP q to locate target m, so that

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

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

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

Note that in reality the target may not always be located within the convex hull due to the random indoor layout (say, corner at a corridor). This issue can be addressed by augmenting map constraints such that the finally returned locations fall within the accessible area. 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

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

Given the above, we present in the following the general localization framework for Maxlifd problem. Denote the index set of APs received by target m as $\mathbf{L}_m$, i.e., $\phi_m^l \neq 0$ if $l \in \mathbf{L}_m$. Let $p(\phi_m^l|\mathbf{r}_q)$ be the likelihood that signal ϕ_m^l appears at RP q. Then the joint likelihood (considered as independent [13]) that Φ_m appears at RP q is

$$p(\Phi_m|\mathbf{r}_q) = \prod_{l=1}^{l\in\mathbf{L}_m} p(\phi_m^l|\mathbf{r}_q). \tag{12}$$

Note that we only consider the uncorrelated APs in Maxlifd, while the virtual or correlated APs and those tethered by mobiles can be easily removed via existing MAC address filtering [39] or correlation computation [16].

In practice, each target (temporal or spatial) may detect different number of APs. Therefore, for each target m, $p(\Phi_m|\mathbf{r}_q)$ is normalized, i.e.,

$$p(\Phi_m|\mathbf{r}_q)' = \frac{p(\Phi_m|\mathbf{r}_q)}{\sum_{q=1}^{Q} p(\Phi_m|\mathbf{r}_q)}. \tag{13}$$

In summary, the probability of each target estimated in the survey region of Q RPs is defined as the mixture of likelihood (joint probabilistic distribution), i.e.,

$$p(\Phi_m) = \sum_{q=1}^{Q}\omega_{mq}p(\Phi_m|\mathbf{r}_q)'. \tag{14}$$

When estimating target m based on above mixture of likelihood, the RPs with higher potential to be target position, i.e., with larger $p(\Phi_m|\mathbf{r}_q)'$, tend to get larger weight ω_{mq} in

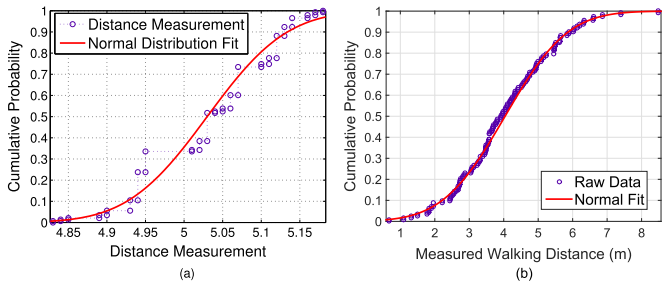


Fig. 2. CDF and normal fit of (a) sound-based distance measurement (5 m between smartphones). (b) Walking distance obtained from pedometer (4 m walking distance).

order to maximize $p(\mathbf{\Phi}_m)$, whose intuition is inspired by the spirits of Gaussian mixture model (GMM) and expectation maximization (EM) [40].

For ease of distance formulation, we denote $\mathcal{I}_{mn} = 1$ if δ_{mn} exists, and $\mathcal{I}_{mn} = 0$ vice versa. Based on Equations (6) and (14), we formulate in Maxlifd a joint maximum likelihood problem for localization. Specifically, we are to find a matrix $\mathbf{W}$ for the M targets (temporal or spatial) such that

$$\begin{aligned} \arg\max_{\mathbf{W}} \quad & \prod_{m=1}^{M} p(\mathbf{\Phi}_m) \prod_{m=1}^{M}\prod_{n=1}^{M} \mathcal{I}_{mn} p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n), \\ \text{subject to:} \quad & \text{Constraints (8), (9), (10) and (11).} \end{aligned} \tag{15}$$

In other words, we are to find the weights of RPs in order to find the target locations which jointly maximize the likelihood of signals and mutual distances. The RPs which satisfy above objective get high weights in final estimation.

4 An Example: Location Estimation with Gaussian Fingerprint & Distance

Given the above statistical framework, we consider here the special case that both signals and mutual distances are Gaussian distributed (Section 4.1). The objective function of Maxlifd has quadratic elements (mutual distances), which lead to high computation overhead. To reduce its hardness, we further implement Semi-Definite Programming relaxation [41] to solve it (Section 4.2). Finally we discuss the application scenarios and complexity analysis (Section 4.3).

4.1 Gaussian Distribution for Fingerprints & Distances

In Section 3.2, the probability distribution, or likelihood, can be calculated through several measures. The selection of distribution forms or likelihood functions may depend on the signal samples and experimental environments. Wireless fingerprint signals have been reported to follow Gaussian distribution [13]. We further conduct extensive data collection in HKUST campus to evaluate the distance measurement randomness and error. 200 mutual distance samples and walking distances are collected. As shown in Figs. 2a and 2b, the pair-wise distance measurements between smartphones and walking distances with step counter may also follow the Gaussian (normal) distribution [4], [6], [42]. Note that other distributions (for example, log-normal or exponential) may be considered in Maxlifd, while here we use Gaussian distribution for illustration and prototyping.

In the derivation of this section, for concreteness and ease of prototyping, we derive the maximum likelihood based on Gaussian distribution. Therefore, in this section, for model derivation and prototyping, we consider Gaussian distribution within the wireless signal noise and distance measurement, i.e., $\phi_q^l \sim \mathcal{N}(\bar{\phi}_q^l, (\sigma_q^l)^2)$ and $\|\mathbf{x}_m - \mathbf{x}_n\|_2 \sim \mathcal{N}(\delta_{mn}, (\sigma_\delta)^2)$. Note that our work is also general enough to be applied in various distributions or statistical measures, including log-normal distribution, histogram-based [13] or kernel-based [14] methods.

Specifically, when comparing the signal vectors at RP q with target signals $\mathbf{\Phi}_m$, we first compute the probability distribution of signal value ϕ_m^l at RP q, i.e.,

$$p(\phi_m^l|\mathbf{r}_q) = \frac{1}{\sqrt{2\pi}\sigma_q^l} \exp\left(-\frac{\left(\phi_m^l - \psi_q^l\right)^2}{2(\sigma_q^l)^2}\right). \tag{16}$$

By definition, if AP $l \in \mathbf{L}_m$ is not detected at RP q, $p(\phi_m^l|\mathbf{r}_q)$ is assigned with a nonzero value $p_{\min}$ (in our evaluation, $p_{\min} = 0.01$). Device heterogeneity in RSSI can be further addressed through existing approaches like offset calibration by crowdsourcing [7] or online RSSI offset learning [15]. For example, we can conduct regression over the RSSIs crowdsourced by many users to find the offset [15], [43].

Then the likelihood that estimated locations $\mathbf{x}_m$ and $\mathbf{x}_n$ match the measured distance δ_{mn} is given by

$$p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n) = \frac{1}{\sqrt{2\pi}\sigma_\delta} \exp\left(-\frac{\left(\delta_{mn} - \|\widehat{\mathbf{x}}_m - \widehat{\mathbf{x}}_n\|_2\right)^2}{2(\sigma_\delta)^2}\right), \tag{17}$$

where σ_δ represents the uncertainty of mutual distance measurement. By definition, if there is no distance measurement between m and n, $p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n) = 0$. The joint probability distribution of the distance matching is then given by

$$p(\boldsymbol{\delta}) = \prod_{m=1}^{M}\prod_{n=1}^{M} \mathcal{I}_{mn} p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n), \tag{18}$$

where indicator $\mathcal{I}_{mn} = 1$ if δ_{mn} exists, and 0 otherwise.

Using Equations (18) and (14), we present in the following the objective function for Maxlifd. Let $\{\mathbf{\Phi}_m\}$ and $\{\delta_{mn}\}$ be the set of RSSI vectors and mutual distances. To measure the matching of all targets (spatial or temporal) with the stored signal map and measured distances, we find the weights in $\mathbf{W}$ which maximize the *joint likelihood* [40], denoted as $\mathcal{L}(\{\mathbf{\Phi}_m\}, \{\delta_{mn}\}|\mathbf{W})$, as the following distribution format [44]:

$$\begin{aligned} \mathcal{L}(\{\mathbf{\Phi}_m\}, \{\delta_{mn}\}|\mathbf{W}) &= \left(\prod_{m=1}^{M} p(\mathbf{\Phi}_m)\right) p(\boldsymbol{\delta}) \\ &= \left(\prod_{m=1}^{M}\left(\sum_{q=1}^{Q} \omega_{mq} p(\mathbf{\Phi}_m|\mathbf{r}_q)'\right)\right) \prod_{m=1}^{M}\prod_{n=1}^{M} \mathcal{I}_{mn} p(\delta_{mn}|\mathbf{x}_m, \mathbf{x}_n). \end{aligned} \tag{19}$$

By taking logarithm on both sides, we get rid of the exponential terms in Equation (17), and transform Equation (19) into a *joint maximum likelihood* formulation, i.e., the log likelihood

$$\begin{aligned} \log \mathcal{L}(\{\mathbf{\Phi}_m\}, \{\delta_{mn}\}|\mathbf{W}) &= \sum_{m=1}^{M} \log\left(\sum_{q=1}^{Q} \omega_{mq} p(\mathbf{\Phi}_m|\mathbf{r}_q)\right) \\ &- \sum_{m=1}^{M}\sum_{n=1}^{M} \frac{\mathcal{I}_{mn}\left(\|\mathbf{x}_m - \mathbf{x}_n\|_2 - \delta_{mn}\right)^2}{2\sigma_\delta^2} - \text{const.} \end{aligned} \tag{20}$$

Note that the logarithm function is convex and therefore our objective function $\log \mathcal{L}(\{\mathbf{\Phi}_m\}, \{\delta_{mn}\}|\mathbf{W})$ is convex, which can be solved through the convex optimization. In our optimization, we can simply leave out the constant term, which is $\text{const} = \sum_{m=1}^{M}\sum_{n=1}^{M} \mathcal{I}_{mn} \log \sqrt{2\pi}\sigma_\delta$.

Our objective function finally becomes

$$\arg\max_{\mathbf{W}} \log \mathcal{L}(\{\mathbf{\Phi}_m\}, \{\delta_{mn}\}|\mathbf{W}), \tag{21}$$

which jointly considers the probability of detected signals and the physical distance constraints.

4.2 Semidefinite Programming (SDP)

To relax the quadratic objective, we describe as follows solving the optimization using SDP. Note that SDP approach in Maxlifd can be applied to other probability distributions by simply replacing the probability distribution in fingerprint RSSI (Equation (16)) and distance (Equation (17)).

Let $\mathbf{e}_{mn}$ be an $M \times 1$ column vector where the mth element is 1 and nth element is -1. The physical distance between node m and n can be therefore represented as

$$\delta_{mn}^2 = \mathbf{e}_{mn}^T \mathbf{X}\mathbf{X}^T \mathbf{e}_{mn}, [\delta_{mn}, \hat{\delta}_{mn}] \in \mathbf{\Omega}_m. \tag{22}$$

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

$$\mathbf{Y} = \mathbf{X}\mathbf{X}^T. \tag{23}$$

Finally, the distance constraint can be rewritten as

$$\mathbf{e}_{mn}^T \mathbf{Y} \mathbf{e}_{mn} = \delta_{mn}^2. \tag{24}$$

Based on the transformation of Equation (24), we can rewrite Formulation (15) into

$$\begin{aligned} &\text{Objective: Equation (21),} \\ &\text{subject to: Constraints (9), (11), (23) and (24).} \end{aligned} \tag{25}$$

Given a symmetric matrix $\mathbf{A}$, let $\mathbf{A} \succeq 0$ represent that $\mathbf{A}$ is a positive semidefinite matrix [41], [45]. 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 (23), with a convex positive semi-definite constraint, i.e.,

$$\mathbf{Y} - \mathbf{X}\mathbf{X}^T \succeq 0. \tag{26}$$

Constraint (26) is a nonlinear constraint, which can be further transformed into linear matrix inequality (LMI) [39]. Through LMI, the quadratic elements within the objective function can be relaxed. Then it can be solved efficiently by a convex optimization solver [41], [45]. The extent of the relaxation can be controlled through the solver. The transformation is through a Schur complement:

Definition 1. *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.,*

$$\mathbf{H} = \begin{bmatrix} \mathbf{B} & \mathbf{E} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \tag{27}$$

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

$$\mathbf{S} = \mathbf{B} - \mathbf{E}\mathbf{D}^{-1}\mathbf{C}. \tag{28}$$

If $\mathbf{S} \succeq 0$, then $\mathbf{H} \succeq 0$ [41]. Let $\mathbf{I}_2$ be a 2×2 identity matrix. Thus, by using Schur complement, we can rewrite Constraint (26) as a matrix form for LMI, i.e.,

$$\begin{bmatrix} \mathbf{Y} & \mathbf{X} \\ \mathbf{X}^T & \mathbf{I}_2 \end{bmatrix} \succeq 0. \tag{29}$$

Formulation (25) is finally transformed into an SDP problem [41]

$$\begin{aligned} &\text{Objective:} \quad \text{Equation (21),} \\ &\text{subject to} \quad \text{Constraints (9), (11), (23), (24) and (29).} \end{aligned} \tag{30}$$

4.3 Localization Applications & Complexity Analysis

The formulation above can be directly applied in peer-assisted localization. For dead-reckoning based localization using inertial navigation system, we may take m as the time stamp (M becomes the number of temporal targets). In fusion with dead reckoning, the device also collects the Wi-Fi RSSI vectors as the user walks. Then using the notation in Equation (7), the index m represents the time stamp during walking. 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 Mth target is returned as the current position. Given $\widetilde{\delta}_{mn}$ at time m from its predecessor, $\mathbf{\Lambda}_m = [m-1]$. Based on Equation (7), a Wi-Fi temporal target $\mathbf{\Pi}_m$ is defined as

$$\mathbf{\Pi}_m \triangleq \{m, \widehat{\mathbf{x}}_m, \mathbf{\Phi}_m, [m-1], \mathbf{\Omega}_m\}. \tag{31}$$

Each temporal target m has distance from its predecessor, $m-1$, and by definition the first target in the sliding window has no distance and predecessor. For beacon-assisted localization, we may let $M = 1$, and a target measures distances from those neighboring beacons with known locations. In practical deployment, at a time only one of above applications is triggered opportunistically given corresponding spatial or temporal distance measurements.

We end by analyzing the computational complexity of the solution. Given Q RPs and L APs, signal likelihood calculation takes $\mathcal{O}(QL)$. Given M temporal or spatial target measurements (usually M is small), the computation of SDP relaxation is bounded by $\mathcal{O}(M^3Q^3)$ [46]. Some commercial SDP solver can solve this problem efficiently [41]. Further computational reduction can be achieved by AP filtering and RP cluster mapping [16]. Specifically, unimportant APs can be first filtered before signal likelihood calculation, and the target can be first mapped to a small area (namely RP cluster) before final Maxlifd calculation [16].

5 CRLB for Localization Error

This section presents the Cramér-Rao lower bound (CRLB) of the Maxlifd problem discussed in Section 4.1. CRLB has been widely used to evaluate the fundamental hardness of an estimation problem [17], [47]. Here we utilize CRLB to evaluate *the localization error bound* of Maxlifd, showing the *optimality* of the formulated problem. Specifically, in Section 5.1, preliminary of CRLB is first introduced. Then, the CRLB of Maxlifd for Gaussian distribution is derived in Section 5.2.

5.1 CRLB Preliminary

CRLB sets a lower limit for the covariance of the unbiased estimates of the unknown parameters [17]. In our formulation,

the location coordinates of the targets (temporal or spatial), denoted as $\boldsymbol{\theta} = [x_1, y_1, x_2, y_2, \ldots, x_M, y_M]^T$, are the corresponding parameters of the estimator. Each target corresponds to a 2-D location, i.e., $\mathbf{x}_m = (x_m, y_m)^T$. Let $\sigma^2_{x_a}$ be the variance of coordinate x_a, and $\sigma_{x_a y_b}$ be the covariance between coordinate x_a and y_b. Therefore, the covariance matrix of all the targets should be a $2M \times 2M$ matrix, i.e.,

$$Cov(\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2_{x_1} & \sigma_{x_1 y_1} & \cdots & \sigma_{x_1 x_M} & \sigma_{x_1 y_M} \\ \sigma_{y_1 x_1} & \sigma^2_{y_1} & \cdots & \sigma_{y_1 x_M} & \sigma_{y_1 y_M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \sigma_{x_M x_1} & \sigma_{x_M y_1} & \cdots & \sigma^2_{x_M} & \sigma_{x_M y_M} \\ \sigma_{y_M x_1} & \sigma_{y_M y_1} & \cdots & \sigma_{y_M x_M} & \sigma^2_{y_M} \end{bmatrix}. \quad (32)$$

Here the diagonal elements in Equation (32) represent the mean squared errors, and the off-diagonal elements are the covariances between the coordinates of the M targets.

The observation in our formulation, denoted as Θ, consists of the received signal strength from APs and the measured pairwise distances between targets, i.e., $\Theta = \left[\phi_{11}, \phi_{12}, \ldots, \phi_{1L}, \ldots, \phi_{ML}, \delta_{12}, \ldots, \delta_{1M}, \ldots, \delta_{M(M-1)}\right]$. Let $f_{\boldsymbol{\theta}}(\Theta)$ be the joint probability density function of the observations Θ conditioned on $\boldsymbol{\theta}$. *Fisher information matrix* (FIM), denoted as $\mathbf{J}(\boldsymbol{\theta})$ (a $2M \times 2M$ matrix), is given by

$$\mathbf{J}(\boldsymbol{\theta}) = -E\left[\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\Theta)}{\partial \boldsymbol{\theta}^2}\right]. \quad (33)$$

CRLB is the inverse of FIM and the covariance of the estimated locations (Equation (32)) is bounded as

$$Cov(\boldsymbol{\theta}) \geq \{J(\boldsymbol{\theta})\}^{-1}. \quad (34)$$

Let $\mathbf{J}_m(\boldsymbol{\theta})$ be the corresponding square block in FIM for target m, $\mathbf{J}_{x_m,x_n}$ be the element at the $(2m-1)$th row and $(2n-1)$th column of FIM $\mathbf{J}(\boldsymbol{\theta})$ (odd row/column), and $\mathbf{J}_{y_m,y_n}$ be that at the $2m$th row and $2n$th column (even row/column). Then the mean squared error of each target m, denoted as $var(\hat{\theta}_m)$, is then bounded by its local geometric relationship $\{\mathbf{J}_m(\boldsymbol{\theta})\}^{-1}$ [10], [18], i.e., $var(\hat{\boldsymbol{\theta}}_m)$ or

$$\begin{aligned} E\Big((\boldsymbol{\theta}_m - \bar{\boldsymbol{\theta}}_m)^T(\boldsymbol{\theta}_m - \bar{\boldsymbol{\theta}}_m)\Big) &\geq \{\mathbf{J}_m(\boldsymbol{\theta})\}^{-1} \\ &= \begin{bmatrix} \mathbf{J}_{x_m,x_m} & \mathbf{J}_{x_m,y_m} \\ \mathbf{J}_{y_m,x_m} & \mathbf{J}_{y_m,y_m} \end{bmatrix}^{-1} \\ &= \frac{1}{|\mathbf{J}_m(\boldsymbol{\theta})|}\begin{bmatrix} \mathbf{J}_{y_m,y_m} & -\mathbf{J}_{x_m,y_m} \\ -\mathbf{J}_{y_m,x_m} & \mathbf{J}_{x_m,x_m} \end{bmatrix}, \end{aligned} \quad (35)$$

where $|\mathbf{J}_m(\boldsymbol{\theta})| = \mathbf{J}_{x_m,x_m}\mathbf{J}_{y_m,y_m} - \mathbf{J}_{x_m,y_m}\mathbf{J}_{y_m,x_m}$.

5.2 CRLB for Gaussian Signal Measurement

In the following, we derive the CRLB of Maxlifd under Gaussian distributed signal measurements.

We first derive the joint probability density function $f_{\boldsymbol{\theta}}(\Theta)$. Based on the formulation in Equation (21) of Section 4.1, we can formulate the measured signal strength between target m and AP l as

$$\phi_{ml} = P_l - 10\gamma\prime \log\left(\frac{\delta_{ml}}{d_0}\right) + \mathcal{X}_{ml}, \quad (36)$$

where $\gamma\prime = \frac{\gamma}{\log 10}$ is introduced for later derivation, and $\mathcal{X}_{ml} \sim \mathcal{N}(0, \sigma^2_{ml})$. Similarly, we formulate the probability density function of estimated distance between m and n as normal distribution, i.e.,

$$\delta_{mn} \sim \mathcal{N}\big(\|\mathbf{x}_m - \mathbf{x}_n\|_2, \sigma^2_{mn}\big). \quad (37)$$

Given above, let $\bar{\phi}_{ml} = P_l - 10\gamma\prime \log(\frac{\delta_{ml}}{d_0})$, and $\bar{\delta}_{mn} = \|\mathbf{x}_m - \mathbf{x}_n\|_2$. The objective in Section 4 jointly maximizes the likelihood of wireless signals and mutual distances. Thus for ease of analysis in CRLB, we equivalently transform the objective function in Equation (21) [10], [17], and the joint probability density function $f_{\boldsymbol{\theta}}(\Theta)$ is then given by

$$\begin{aligned} f_{\boldsymbol{\theta}}(\Theta) = &\left(\prod_{m=1}^{M}\prod_{l=1}^{L}\frac{1}{\sqrt{2\pi}\sigma_{ml}}\exp\left(-\frac{(\phi_{ml} - \bar{\phi}_{ml})^2}{2\sigma^2_{ml}}\right)\right) \\ &\times\left(\prod_{m=1}^{M}\prod_{n=1}^{M}\frac{1}{\sqrt{2\pi}\sigma_{mn}}\exp\left(-\frac{(\delta_{mn} - \bar{\delta}_{mn})^2}{2\sigma^2_{mn}}\right)\right), \end{aligned} \quad (38)$$

which is equivalent to the joint maximum likelihood formulation presented in Equation (20).

Therefore, the first derivatives of the objective function in Equation (38), are given by

$$\begin{aligned} \frac{\partial \log f_{\boldsymbol{\theta}}(\Theta)}{\partial x_m} = &\sum_{l=1}^{L}\frac{1}{\sigma^2_{mn}}(\phi_{ml} - \bar{\phi}_{ml})\frac{\partial \bar{\phi}_{ml}}{\partial x_m} \\ &+ \sum_{n=1}^{M-1}\frac{1}{\sigma^2_{mn}}(\delta_{mn} - \bar{\delta}_{mn})\frac{\partial \bar{\delta}_{mn}}{\partial x_m}, \end{aligned} \quad (39)$$

$$\begin{aligned} \frac{\partial \log f_{\boldsymbol{\theta}}(\Theta)}{\partial y_m} = &\sum_{l=1}^{L}\frac{1}{\sigma^2_{mn}}(\phi_{ml} - \bar{\phi}_{ml})\frac{\partial \bar{\phi}_{ml}}{\partial y_m} \\ &+ \sum_{n=1}^{M-1}\frac{1}{\sigma^2_{mn}}(\delta_{mn} - \bar{\delta}_{mn})\frac{\partial \bar{\delta}_{mn}}{\partial y_m}, \end{aligned} \quad (40)$$

where the partial derivatives with respect to coordinate x_m or y_m are given by

$$\begin{aligned} &\frac{\partial \bar{\phi}_{ml}}{\partial x_m} = -\frac{10\gamma\prime d_0}{d^2_{ml}}(x_m - x_l), \quad \frac{\partial \bar{\phi}_{ml}}{\partial y_m} = -\frac{10\gamma\prime d_0}{d^2_{ml}}(y_m - y_l), \\ &\frac{\partial \bar{\delta}_{mn}}{\partial x_m} = \frac{x_m - x_n}{\delta_{mn}}, \qquad \frac{\partial \bar{\delta}_{mn}}{\partial y_m} = \frac{y_m - y_n}{\delta_{mn}}. \end{aligned} \quad (41)$$

Considering $\alpha_{ml} \in [0, 2\pi)$ as the relative angle that target m makes with respect to AP l, and $\beta_{mn} \in [0, 2\pi)$ ($\beta_{nm} = 2\pi - \beta_{mn}$) as that the target m makes w.r.t. target n. Therefore, we have

$$\begin{aligned} &\cos\alpha_{ml} = \frac{x_m - x_l}{d_{ml}}, \quad \sin\alpha_{ml} = \frac{y_m - y_l}{d_{ml}}, \\ &\cos\beta_{mn} = \frac{x_m - x_n}{\delta_{mn}}, \quad \sin\beta_{mn} = \frac{y_m - y_n}{\delta_{mn}}. \end{aligned} \quad (42)$$

Given above elements, we calculate expectation of the second derivatives in Equation (33) [10], [20], and find the entries within the FIM $\mathbf{J}(\boldsymbol{\theta})$, which are given as follows:

$$\begin{aligned}\mathbf{J}_{x_m,x_m} &= -E\left[\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\Theta)}{\partial x_m^2}\right] \\ &= \sum_{l=1}^{L}\frac{1}{\sigma_{ml}^2}\left(\frac{10\gamma\prime d_0}{d_{ml}^2}\right)^2 (x_m - x_l)^2 + \sum_{n=1}^{M-1}\frac{1}{\sigma_{mn}^2}\frac{(x_m - x_n)^2}{\delta_{mn}^2} \\ &= \sum_{l=1}^{L}\frac{(10\gamma\prime d_0)^2}{\sigma_{ml}^2}\frac{(\cos\alpha_{ml})^2}{d_{ml}^2} + \sum_{n=1}^{M-1}\frac{(\cos\beta_{mn})^2}{\sigma_{mn}^2},\end{aligned} \tag{43}$$

$$\begin{aligned}\mathbf{J}_{y_m,y_m} &= -E\left[\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\Theta)}{\partial y_m^2}\right] \\ &= \sum_{l=1}^{L}\frac{1}{\sigma_{ml}^2}\left(\frac{10\gamma\prime d_0}{d_{ml}^2}\right)^2 (y_m - y_l)^2 + \sum_{n=1}^{M-1}\frac{1}{\sigma_{mn}^2}\frac{(y_m - y_n)^2}{\delta_{mn}^2} \\ &= \sum_{l=1}^{L}\frac{(10\gamma\prime d_0)^2}{\sigma_{ml}^2}\frac{(\sin\alpha_{ml})^2}{d_{ml}^2} + \sum_{n=1}^{M-1}\frac{(\sin\beta_{mn})^2}{\sigma_{mn}^2},\end{aligned} \tag{44}$$

$$\begin{aligned}\mathbf{J}_{x_m,y_m} = \mathbf{J}_{y_m,x_m} &= -E\left[\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\Theta)}{\partial x_m \partial y_m}\right] \\ &= \sum_{l=1}^{L}\frac{1}{\sigma_{ml}^2}\left(\frac{10\gamma\prime d_0}{d_{ml}^2}\right)^2 (x_m - x_l)(y_m - y_l) \\ &\quad + \sum_{n=1}^{M-1}\frac{1}{\sigma_{mn}^2}\frac{(x_m - x_n)(y_m - y_n)}{\delta_{mn}^2} \\ &= \sum_{l=1}^{L}\frac{(10\gamma\prime d_0)^2}{\sigma_{ml}^2}\frac{\sin 2\alpha_{ml}}{2d_{ml}^2} + \sum_{n=1}^{M-1}\frac{\sin 2\beta_{mn}}{2\sigma_{mn}^2},\end{aligned} \tag{45}$$

and

$$\mathbf{J}_{x_m,x_n} = \mathbf{J}_{x_n,x_m} = \mathbf{J}_{y_m,y_n} = \mathbf{J}_{y_n,y_m} = \frac{1}{\sigma_{mn}^2}, \tag{46}$$

and

$$\begin{aligned}\mathbf{J}_{x_m,y_n} &= -E\left[\frac{\partial^2 \log f_{\boldsymbol{\theta}}(\Theta)}{\partial x_m \partial y_n}\right] = \frac{1}{\sigma_{mn}^2}\frac{(x_m - x_n)(y_m - y_n)}{\delta_{mn}^2} \\ &= \frac{1}{\sigma_{mn}^2}\frac{\sin 2\beta_{mn}}{2},\end{aligned} \tag{47}$$

while $\mathbf{J}_{x_n,y_m} = \mathbf{J}_{x_m,y_n} = \mathbf{J}_{y_m,x_n} = \mathbf{J}_{y_n,x_m}$. To summarize, with Equation (35) and the above components, we can obtain the corresponding lower bound for estimation variance (mean squared errors) of each involved target m in Maxlifd.

6 Experimental Evaluation

We have developed Maxlifd on Android platforms based on Wi-Fi fingerprints, and conducted extensive experiments. In this section, we first discuss the experimental settings and performance metrics in Section 6.1. Then we present illustrative experimental results for dead reckoning fusion (fusing with inertial navigation system, or INS), ***p***eer-***a***ssisted (PA, i.e., fusing with mutual distances of peers) and ***b***ea***c***on-assisted (BC, i.e., fusing Wi-Fi with sensor beacons) localization in Sections 6.1. Though practically only one distance type is likely operating at a time, our framework is general enough to include all the distance information mentioned.

6.1 Experimental Settings & Performance Metrics

We compare Maxlifd with the following typical and state-of-the-art localization schemes in our experiment:

1) *Fingerprint-based localization (FL)* [1], [2], which evaluates the Euclidean distance of target RSSI vector with each RP fingerprint, and finds interpolation of the top k nearest neighbors for location estimation ($k = 15$).
2) *Sequential Monte Carlo (SMC) localization* [3], [7], the state-of-the-art fusion algorithm based on Sequential Monte Carlo method (particle filter) using INS data and Wi-Fi fingerprinting (FL). Via the propagation along the temporal walking path, the particles move from one location to the next. With map constraints, the spatial distribution of these particles is corrected and resampled. The final estimation is based on the weighted average of particle locations.
3) *Hidden Markov Model (HMM) localization* [8], [9], which fuses INS with Wi-Fi fingerprints (FL), and conducts location positioning based on a traditional probabilistic framework, hidden Markov models.
4) *Localization with Social Interaction (SocialLoc)* [4], which fixes the Wi-Fi fingerprint-based location estimations (FL) via Bluetooth-based interaction of users (encounter/nonencounter information).
5) *Graph-based & fingerprint localization scheme (GB + FL)* [6], which uses graph construction and Wi-Fi fingerprinting (similar to FL) for peer-assisted localization. With the pairwise spatial distances of peer targets, GB+FL constructs the rigid graph consisting of all targets [5], [48]. Then GP+FL searches against the Wi-Fi signal map and finds a set of fingerprints to minimize the objective function $\Sigma_{m=1}^{M}\|\Phi_m - \Psi_q\|^2$ via graph rotation and translation [6].
6) *MMSE* [49]: For indoor beacon-based positioning scenarios, we compare the performance of Maxlifd with minimum mean squared error (MMSE) algorithm [49], which has been applied in many existing iBeacon-based systems. 3 iBeacons with the strongest RSSIs are used for MMSE.

Corresponding state-of-arts are compared in each application scenario. Specifically, in experimental comparison of dead reckoning fusion (INS), we compare Maxlifd with SMC and HMM. In peer-assisted (PA) application, we compare Maxlifd with Social-Loc and GB+FL. In beacon-based localization (BC), we compare Maxlifd with MMSE. Note that FL is compared in all scenarios as the base comparison without sensor fusion. In iBeacon-based localization scenario (BC), we implement FL using Bluetooth fingerprints [26].

Let $\mathbf{x}_m$ be target m's true location and $\widehat{\mathbf{x}}_m$ be the estimated location. The performance metrics used are the mean localization error (unit:m) of the estimated target in set $\mathbf{V}$: $ME = \left(\sum_{m\in\mathbf{V}}\|\mathbf{x}_m - \widehat{\mathbf{x}}_m\|_2\right)/|\mathbf{V}|$.

We evaluate Maxlifd 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. Figs. 3a and 3b show the corresponding floor plans of HKIA and HKUST campus. Note that the black triangles in the map correspond to the deployed iBeacons. In the HKIA and the HKUST campus, at each RP we take overall 80 Wi-Fi RSSI vectors using HTC One X+ (total 5 are used).

A quarter of these samples are collected when we are facing north, south, west and east, respectively. Note that the target signals are collected at least one month after the fingerprinting. For all the application scenarios, unless otherwise stated we use the following parameters as baseline: 5 m survey grid size (i.e., width between two neighboring

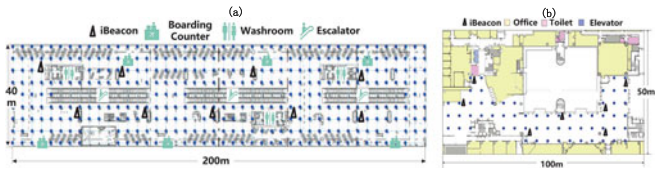


Fig. 3. (a) Boarding area at HKIA. (b) HKUST campus atrium (black triangles: locations of iBeacons).

RPs, which represents the density); 1 Wi-Fi RSSI sample (which takes around 1 second for sampling on our Android platform) is used for each target; no Wi-Fi AP reduction is conducted over the target RSSI vectors. Note that HTC One X takes around 1 second to scan one Wi-Fi RSSI vector as the OS needs to scan all possible channels (2.4 and 5 GHz). Bluetooth fingerprinting is conducted in the same process as the Wi-Fi fingerprinting.

Note that the ranging test is conducted in the noisy environment (including campus and airport). The sound ranging is working at different frequencies from the audible sound, which is overall robust against the environmental dynamics [6], [50].

The signal studies are illustrated as follows. In Fig. 4, we show the CDF of detected AP number at each target and RSSI standard deviation (STD) at RPs (Equation (2)) in HKUST and HKIA. Though in HKIA the number of detected APs at each RP is larger, most of their detected signals are weak (less than -80 dBm). As a larger spectrum of signal variation is observed in HKIA, we expect a larger localization error there. Preprocessing of AP importance, virtual and correlated AP filter has been conducted here, and the implementation details are similar to [14], [37]. Note that we conduct the experiment during working hours with crowds nearby (each collection takes around 6 hours). The sites where we conduct experiments are with marked RSSI noise and low fingerprint differentiation, leading to accuracy degrade of state-of-the-art algorithms.

In fusing with dead reckoning (INS), sliding window size $M = 7$ for Maxlifd, and the influence of different M's is also evaluated in the experiment. 200 particles are used in the particle filter (SMC). A step counter at the smartphone measures user steps and temporal distances. Each step detection is based on the periodic changes in the gravity direction of accelerometer [7]. The walking traces range from 8 to 15 meters. We label the ground truth based on the landmarks nearby (say, the door, hall gate or a pillar). 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 [42]). Step length calibration can be referred to works like [21] or [51] for related approaches.

For some areas visited by many users, peer-assisted localization may be used [6]. Peer ranging can be based on either RSS-distance mapping or sound ranging. Wi-Fi direct is not considered as it may cause extra overhead and interference to ongoing Wi-Fi transmission. Time synchronization between transceivers makes obtaining accurate distance information from Wi-Fi direct difficult. Sound-based distance estimation is easier for time synchronization and implementation. For ease of prototyping, in the experiment we implement and test smartphone-based sound ranging like [6] under quiet and noisy campus environment. The mean peer ranging errors under these two conditions are 0.8 m and 2 m respectively. Since the distance constraint between two peers is asymmetric

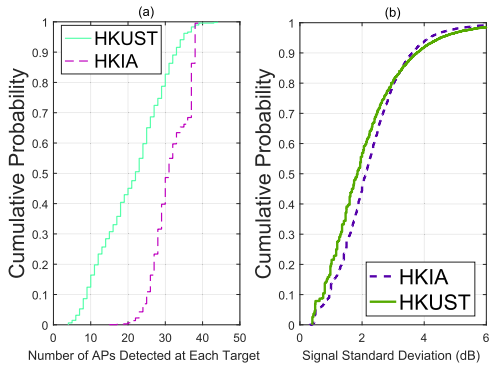


Fig. 4. (a) CDF of detected AP number at each target and (b) CDF of RSSI STD (dB) at RPs in two sites.

due to measurement uncertainty, we use their average value of the measurements as mean in Gaussian likelihood estimation. By default, 5 targets (users) are involved in sound-based distance measurement. We do not exclude the cases when walls may partition some of the peers during localization (with none-line-of-sight or NLoS measurements).

In beacon-assisted localization (BC), we implement Maxlifd upon Bluetooth iBeacons (Fig. 3) with TI CC2540. Besides Wi-Fi fingerprinting, we also measure the distance between the beacon and the target based on the path loss model. The calibration of the iBeacons on RSSI-distance model and heterogeneous devices can be achieved through online learning [15] and crowdsourcing-based methods [7]. To further mitigate the signal noise, we empirically implement a mean filter over a sliding window of three consecutive RSSI samples. After that, a distance estimation result is returned to Maxlifd. Further importance differentiation of spatial distances in both PA and BC applications is outside the scope of Maxlifd and interested readers may refer to established prior arts like [6].

6.2 Illustrative Experimental Results

We evaluate the fusion with dead reckoning (INS) as follows. Fig. 5 shows the cumulative probability of localization errors in the HKIA. We can observe at least 30 percent error reduction using Maxlifd compared with other state-of-the-arts. As the sensor readings are jointly considered in one single formulation, Maxlifd mitigates the influence in walking distance and hence achieves much lower errors than SMC and HMM. Fig. 6 shows the computation time using SMC, HMM and Maxlifd running on a PC with i7 3610QM. We may observe slightly higher computation in Maxlifd than traditional HMM, while Maxlifd achieves better efficiency than the particle filter in SMC. Traditional SMC often suffers from computation under large number of particles in order to locate the target.

Fig. 7 shows the ME of Maxlifd (INS) against M, the number of Wi-Fi temporal target measurements (size of sliding window) in HKIA. We can see that the accuracy improves as we utilize more temporal samples, as joint consideration of more periods further constrains the location estimations. When we further increase the number of measurements, the localization accuracy gradually converges, indicating that given distances already provide sufficient constraints. Thus, to balance between localization accuracy and computational complexity we choose several temporal measurements (like 7 in our experiment) in Maxlifd (INS).

Fig. 8 plots the real-time localization error for Maxlifd (INS), HMM and SMC in HKIA. The estimation error fluctuates as the user walks in the airport. Changes in wall

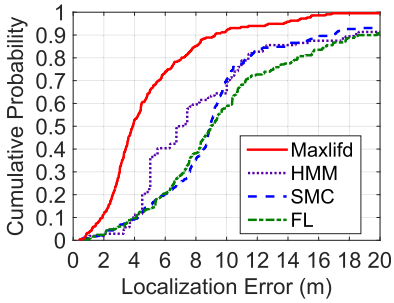


Fig. 5. Cumulative distribution function (CDF) of localization errors (INS) in HKIA.

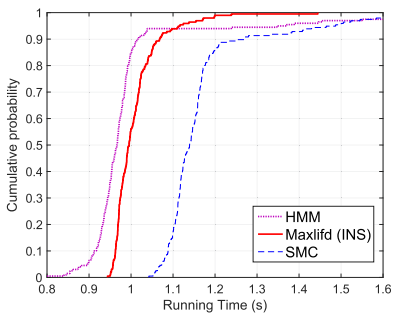


Fig. 6. CDF of running time in fusing with INS (HKIA).

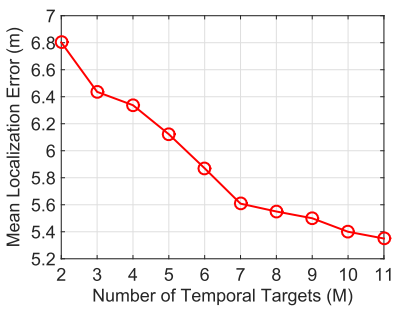


Fig. 7. *ME* versus number of Wi-Fi temporal targets (HKIA).

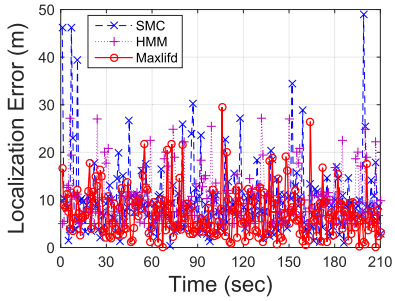


Fig. 8. Real-time position errors versus time in HKIA (INS).

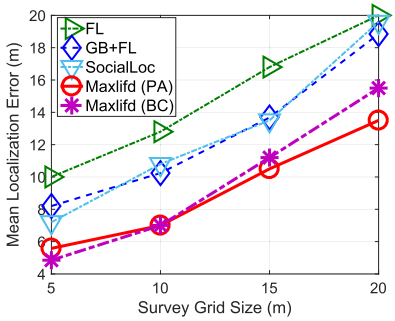


Fig. 9. *ME* versus survey grid size (HKIA).

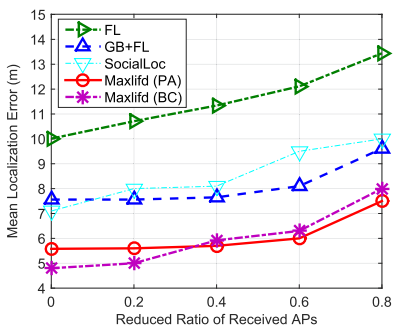


Fig. 10. *ME* versus reduced ratio of detected APs in HKIA.

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, Maxlifd constrains its estimations through the single optimization over the joint maximum likelihood. Therefore, Maxlifd can achieve lower localization errors and smaller estimation fluctuation.

Further evaluation of Maxlifd in peer-assisted and beacon-based localization (BC) is presented as follows. Fig. 9 shows the mean localization errors versus different survey grid size. As grid size increases, the localization accuracy decreases, while a diminishing return is observed we use 5 m as the grid size. Fig. 10 shows the mean localization errors against the proportion of APs removed at targets in HKIA. 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 peer-assisted localization systems marginally rely on the number of received APs. It is because the Wi-Fi samples at multiple users 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. Compared with GB+FL and SocialLoc, Maxlifd uses joint optimization of fingerprints and distances which reduces noise influence and suffers less from AP loss.

We show the CDF of location errors for different algorithms at two sites as follows. Figs. 11 and 12 show the location error CDF of Maxlifd at different scenarios (PA and BC at baseline parameters) in HKIA. Large indoor open space often leads to high uncertainty in Wi-Fi signals [2] and hence the disperse nearest neighbors in signal space. Furthermore, the temporal and spatial distance measurement also contains large signal noise under the crowded scenarios. Therefore, we expect the signals collected (Wi-Fi and mutual distance measurements) contains strong NLoS elements. Compared with other state-of-the-art algorithms, Maxlifd significantly reduces the errors often by around 30 percent in the noisy environment of HKIA. With distance constraints and joint optimization, Maxlifd mitigates the influence of dispersed nearest neighbors in location mapping.

Compared with HKIA, the campus atrium in HKUST is smaller in size with more building partitions, which may influence the peer-distance measurement accuracy. We show the accuracy of Maxlifd (INS, PA and BC) on HKUST campus in Figs. 13, 14 and 15, respectively. Maxlifd achieves much higher localization accuracy (often about 30 percent error

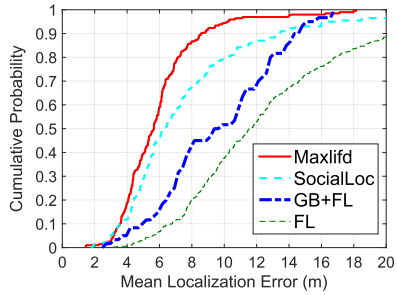


Fig. 11. CDF of localization errors (PA) in HKIA.

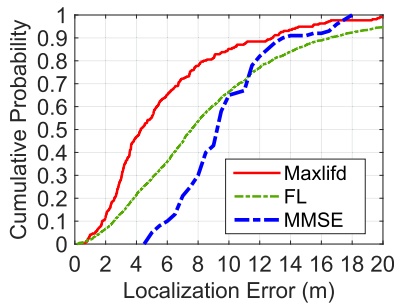


Fig. 12. CDF of localization errors (BC) in HKIA.

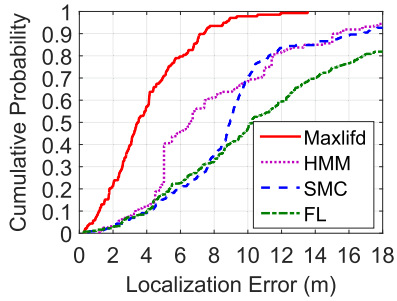


Fig. 13. CDF of localization errors (INS) in HKUST.

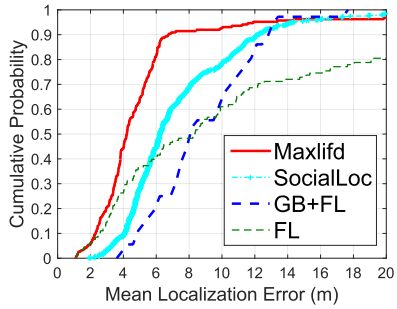


Fig. 14. CDF of localization errors (PA) in HKUST.

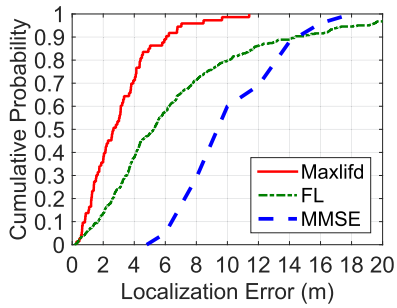


Fig. 15. CDF of localization errors (BC) in HKUST.

reduction) than the other state-of-the-art algorithms. To summarize, for PA scenarios, several users (say, 4 to 5) collect multiple fingerprints and more mutual distances, which is likely to provide more constraints than a single user (INS). Similarly, for BC scenarios, given knowledge of detected beacon locations, the user is also unlikely to be mapped to distant locations. In both cases, the tail of localization error is hence shorter than that in the single-user case. As the results in HKUST are qualitatively similar to those in HKIA, for brevity we do not repeat other experimental results here.

7 Simulation Evaluation

Many environmental parameters are not re-configurable or tunable for a complete experimental studies. To more comprehensively evaluate Maxlifd in large-scale environments under different factors (for example, signal noise, distribution, AP number and survey grid size are limited in the experiment), we simulate a site in the Hong Kong International Airport (HKIA). In this section, we first discuss the simulation setup in Section 7.1. Then we present the illustrative results in dead reckoning fusion (INS), peer-assisted and beacon-assisted localization (BC) in Sections 7.2, 7.3 and 7.4, respectively.

7.1 Simulation Setup

We simulate the Wi-Fi RSSI following the work in [49]. In the signal model, the RSSI ϕ (dBm) from Wi-Fi AP at a distance D can be simulated as

$$\phi = \phi^{TX} - L^0 - 10\gamma \log_{10}\left(\frac{D}{D^0}\right) + \xi, \tag{48}$$

where measurement noise is distributed as $\xi \sim \mathcal{N}(0, \sigma_{db}^2)$. Unless otherwise stated, we use the following settings in our baseline parameters: the transmission power $\phi^{TX} = 25$ dBm; the path loss exponent $\gamma = 4.0$; reference path loss $L^0 = 37.7$ dB; reference distance $D^0 = 1$ m; 215 m $\times$ 40 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 2 seconds during walking.

We also evaluate the CRLB on location errors of Maxlifd given these simulation environment factors. As a result, we evaluate the root mean square error (RMSE) in order to compare with CRLB of Maxlifd, i.e.,

$$RMSE = \sqrt{\frac{1}{|\mathbf{V}|}\sum_{m\in\mathbf{V}} \|\mathbf{x}_m - \widehat{\mathbf{x}}_m\|_2^2}, \tag{49}$$

which is the root of mean square error. We also evaluate the mean localization error (*ME*) in our simulation.

We also compare our Maxlifd with the state-of-the-art algorithms discussed in Section 6.1. We also evaluate the performance difference between Maxlifd and Wi-Dist in some simulations. Detailed settings for each of localization scenarios (INS, PA and BC) are presented as follows.

In INS cases, by default 7 most recent Wi-Fi temporal targets (records) are used. The step count error rate is distributed as $\mathcal{N}(0, \sigma_r^2)$, where $\sigma_r = 25\%$, and the stride length error follows $\mathcal{N}(0, \sigma_l^2)$, where $\sigma_l = 0.2$ m. Additional walking displacement error is assumed to follow $\mathcal{N}(0, \sigma_w^2)$, where $\sigma_w = 3.5$ m. Each user starts walking towards a randomly selected destination point in the site. Based on the random waypoint model, once the user reaches the final

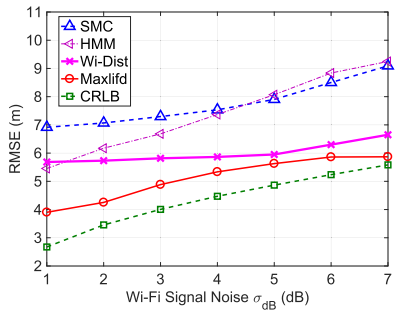


Fig. 16. RMSE versus Wi-Fi signal noise (INS).

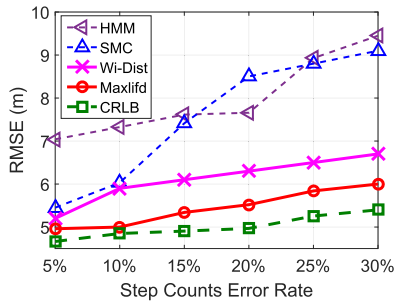


Fig. 17. INS RMSE versus step counts error rate.

location, he or she stops for some random time and starts walking again to another randomly selected destination [4].

In peer-assisted localization, we consider by default noisy mutual distance error $\epsilon_m \sim \mathcal{N}(0, \sigma_m^2)$, σ_m = 3.5 m; 90 users are randomly distributed in the survey site. If two users are within a certain neighborhood detection range (15 m in our simulation), a pairwise link is constructed and a distance constraint is posed upon them. Given Wi-Fi signals collected at each target, by default four peer users together conduct a peer-assisted localization, and further evaluation over peer number is also conducted. k is also set as 15 for FL, SocialLoc and GB+FL.

In the beacon-based localization (BC), we utilize the log-distance path loss model for iBeacon RSSI-distance measurement. The basic approach is to measure the distance based on path loss signal model. In such application, $M = 1$ in Equation (7). For a single target, the corresponding distance δ_n from beacon n, given detected RSSI ϕ_n, can hence be represented as $\delta_n = d_0 10^{(P_l - \phi_n)/\gamma}$. Given a normally distributed ϕ_n, the measured distance δ_n follows a log-normal distribution [49]. Hence in Equations (20) and (38), we replace the normal distribution, with the logarithm of the mutual distances in the final formulation (note that linear signal model [15] may be also applied, and the distance measurement may be modeled as normally distributed). CRLB derivation for the BC case is similar to that using Gaussian distribution, and for brevity we do not repeat here. In our simulation, by default a target can detect an iBeacon if their distance is within 12 m; 80 beacons are deployed.

7.2 Fusing with Dead Reckoning

Fig. 16 plots the RMSE versus the signal variation in Wi-Fi fingerprint (σ_{db} in Equation (48)). We can observe that the accuracy of SMC and HMM degrades when σ_{dB} increases. It is because larger signal noise makes it more difficult to differentiate the fingerprints. For CRLB, the fingerprint signal noise σ_{ml} (coming from σ_{dB} in Equation (48)) decreases the diagonal elements, $\mathbf{J}_{x_m,x_m}$ and $\mathbf{J}_{y_m,y_m}$, hence increasing the inverse of the FIM. CRLB for Maxlifd estimation enlarges. Besides, the difference between Maxlifd and CRLB decreases as the noise increases. It is because under small fingerprint noise, the distance errors become the dominant source of final localization errors. As σ_{dB} further increases, the fingerprint noise becomes dominant in CRLB. Different from SMC and HMM, Maxlifd considers signal uncertainty through the joint maximum likelihood. By jointly maximizing the overall likelihood in signal measurement and mutual distances, Maxlifd reduces the effect of disperse nearest neighbors in signal space and obtains better estimation results.

Fig. 17 shows the RMSE against the step count errors. Clearly, the accuracy of SMC, HMM and Maxlifd degrades with larger step count error. SMC locates the user based on the particle filter and HMM considers conditional independence. They have not jointly combined fingerprints and step count information into one formulation. From CRLB, we can observe that the measurement errors σ_{mn} in Equations (43) and (44) decrease as distance noise increases, leading to increase of $\{\mathbf{J}(\boldsymbol{\theta})\}^{-1}$. Therefore, the CRLB increases with the step count errors. When step count accuracy degrades, the displacement error increases and the particles become spatially sparse, making it difficult for SMC to converge to the correct location. Similarly, HMM considers conditional independence without joint consideration of measurements, and information loss in sensor fusion happens. Therefore, we observe higher errors in their performance.

To the contrary, Wi-Dist and Maxlifd both localize the target more accurately because the joint consideration of Wi-Fi fingerprints and distances of multiple periods (temporal targets) reduces the influence of measurement uncertainty. Compared with Wi-Dist, Maxlifd constructs a joint maximum likelihood formulation, which is more general and robust to signal noise than using bounds. Under large distance errors, such a flexible formulation reduces the misestimation if the bounds do not properly cover the correct location.

Fig. 18 shows the RMSE against the walking displacement errors. We can see that as the displacement error increases, the overall location accuracy decreases. Similarly, from CRLB we can observe that the measurement errors σ_{mn} within temporal distances in Equations (43) and (44) decrease as distance noise increases, leading to the increase of $\{\mathbf{J}(\boldsymbol{\theta})\}^{-1}$. Therefore, the CRLB increases as the step count error becomes larger. Localization error in SMC increases because the particles converge slowly given large distance errors and noisy Wi-Fi fingerprints. Unlike SMC, Maxlifd achieves more accurate results as it utilizes the joint maximum likelihood instead of actual distance measurement. By maximizing the joint likelihood in fingerprints and mutual distances, Maxlifd is more robust to distance uncertainty. As Maxlifd also outperforms Wi-Dist for fusion estimation accuracy and robustness, in the following we focus on comparing Maxlifd with other state-of-the-art algorithms.

7.3 Peer-Assisted Fusion

Fig. 19 shows the RMSE versus the number of users. Clearly, more peer assistance provides more distance constraints over the involved users and improves the localization accuracy. For CRLB, more peer assistance increases the diagonal elements in FIM $\mathbf{J}(\boldsymbol{\theta})$. As more users participate in the localization, we can observe that CRLB decreases, and better localization accuracy is expected. Compared

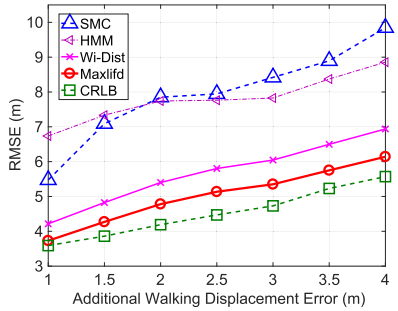


Fig. 18. INS RMSE versus additional walking displacement errors.

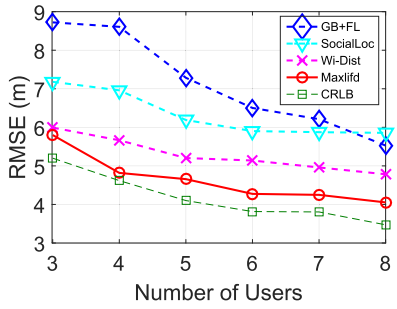


Fig. 19. RMSE versus peer user number (PA).

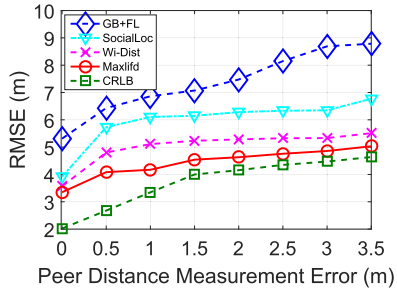


Fig. 20. RMSE versus peer distance errors (PA).

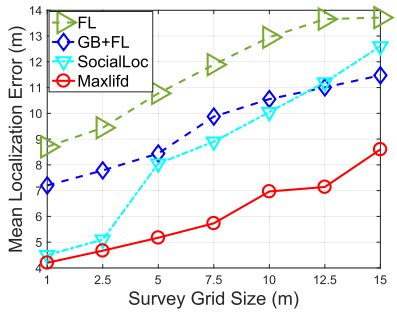


Fig. 21. *ME* versus site survey grid size (PA).

with GB+FL, Maxlifd shows less dependency on the user connectivity. It is because Maxlifd 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. SocialLoc considers dynamic encounters and nonencounters among users, and hence is not sensitive to user number. Maxlifd jointly optimizes the target estimations and achieves better results than SocialLoc. Accuracy gain from number of involved users for Maxlifd becomes smaller as more users participate in location estimation. It is mainly because interaction between distant users contributes for all users in each location estimation, which is also reflected from a stabler CRLB.

Fig. 20 shows the RMSE against peer distance errors. We assume a zero-mean Gaussian noise is added to the mutual distance measurement. When the distance error is small, all algorithms achieve high accuracy given only Wi-Fi fingerprint noise. As distance error further increases, all algorithms suffer from lower accuracy. GB+FL constructs a rigid graph to constrain relative positions of different users. The graph shape deforms under large distance measurement errors and lead to misestimation. SocialLoc considers independently over the pairwise distance constraints between users who encounter. Without joint consideration, their accuracy still degrades with large signal noise. Maxlifd, in contrast, leverages a more flexible probabilistic localization based on fingerprints and distance likelihood. Without assuming a rigid graph, Maxlifd can achieve more robust localization estimation. For CRLB, as more distance noises, σ_{mn}, are introduced to the Fisher information matrix, the diagonal elements in the FIM become smaller, leading to higher CRLB. Therefore, we can observe the increase within the lower bound of estimation errors.

Fig. 21 shows *ME* versus site survey grid size. As the grid size increases, all algorithms degrade in accuracy. FL degrades due to ambiguity of fingerprints under large grid size. Without joint consideration of distances and fingerprints, GB+FL and Social-Loc cannot achieve high accuracy with low site survey density. Different from above schemes, Maxlifd jointly considers both the distance and fingerprint uncertainty and relies less on fingerprint density. Hence Maxlifd can achieve better localization accuracy. Fig. 21 also shows saturation in accuracy improvement under small grid size due to noise and measurement uncertainty.

7.4 Beacon-Assisted Fusion

In the simulation with iBeacon, we assume that when the target is within the range of the beacons, the target can detect and measure the distances from them. Fig. 22 shows the RMSE versus the iBeacon detection range. It shows the localization error decreases when the iBeacon detection range increases. It is mainly because larger iBeacon coverage increases the connectivity of each target, leading to more constraints over the final location estimation. Larger range introduces more distance intersections and more constraints. Therefore, the diagonal elements in $\mathbf{J}(\boldsymbol{\theta})$ increase and CRLB correspondingly decreases. Then the trend of decrease slows down because each target point has obtained sufficient constraints. Similarly in CRLB, as more distance constraints are included, the influence over final target estimation decreases.

Fig. 23 shows the RMSE versus the iBeacon deployment density. It shows that the RMSE of both MMSE and Maxlifd decreases proportionally with the iBeacon deployment density. Then the trend of decrease diminishes after a certain number of beacons. However, dense deployment may introduce large survey efforts, and thus in real application we deploy beacons at area with sparser Wi-Fi coverage. When the density of iBeacon increases, the connectivity for each target increases. The diagonal elements of $\mathbf{J}(\boldsymbol{\theta})$ also increase, leading to a lower CRLB in location estimation.

Fig. 24 shows the *ME* under different probability distributions in fingerprints and distances for Maxlifd (BC). We consider a uniform signal noise (i.e., $\xi \sim \mathcal{U}(-\sigma_{db}, \sigma_{db})$ in Equation (48)), and the log-normal signal noise (i.e., $\Phi \sim \log \mathcal{N}(\Phi^{TX} - L^0 - 10\gamma \log_{10}(\frac{D}{D^0}), \sigma_{db}^2)$) exists within the Wi-Fi RSSIs, respectively (similar forms are applied in the mutual distance measurements). We compare Maxlifd

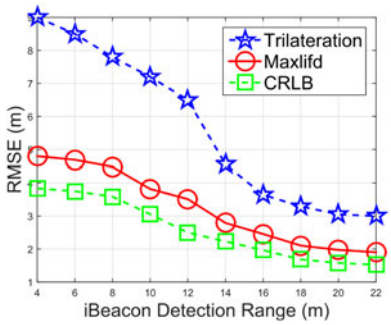


Fig. 22. RMSE versus iBeacon detection range (BC).

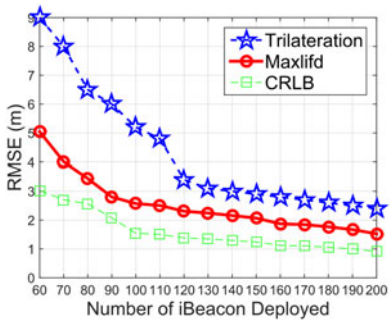


Fig. 23. RMSE versus iBeacon deployment density (BC).

under the corresponding distribution formulations and that considering Gaussian distribution (see Section 4). From the marginal sacrifice of accuracy and estimation fluctuation, we can see Maxlifd does not depend sensitively on knowing exactly the underlying signal or distance distribution, where the simple Gaussian representation suffices to offer reasonable accuracy. In practice, the calculation in probability distribution is also similar (the overall computation is bounded by the optimization process), and the small difference in positioning error indicates that the influence over the optimality is also marginal.

8 Discussion

We further discuss some important deployment issues related to Maxlifd as follows.

- *Reducing fingerprinting cost*: How to learn and reduce the fingerprints has been an important topic for indoor localization. Various mechanisms have been proposed recently, including crowdsourcing/bootstrapping [32] and signal map reconstruction [52], in order to get rid of detailed, costly and labor-intensive survey. Given context of ubiquitous crowdsourcing and appropriate preprocessing [53], fingerprints can be collected or updated at much lower expenses [54]. Despite our focus here in locating target given sensor fusion, these studies can be integrated with Maxlifd to further reduce deployment costs, achieving more ubiquitous localization.
- *Adapting to power adjustment and environmental change*: Signal dynamics due to power adjustment [52] and environmental alternation [55] have attracted research attention recently. Maxlifd is evaluated in practical environments of strong signal dynamics (including power alteration and multi-path effects), and shown to outperform other schemes due to its joint optimization. To make Maxlifd more robust, one may further integrate altered AP detection/filtering [51], and adapts the fingerprint map to the spatial-temporal dynamics [32], [55]. Interested readers may refer to these works, and for brevity we do not explore deeper here.
- *Enhancing ranging performance*: None-line-of-sight (NLoS) ranging may affect the mutual distance constraints. Take the sound-ranging as an example. Inside narrow and NLoS environment (like small rooms), the accuracy of PA may be compromised due to imperfection of their ranging models. We may adopt some effective schemes (see [56]) to detect strong multi-path effect within the sound reflection, thus foreseeing potential degradation in sound ranging accuracy. As a general framework, Maxlifd may also switch to other more applicable sensors (say, inertial pedometer) to conduct fusion-based localization.

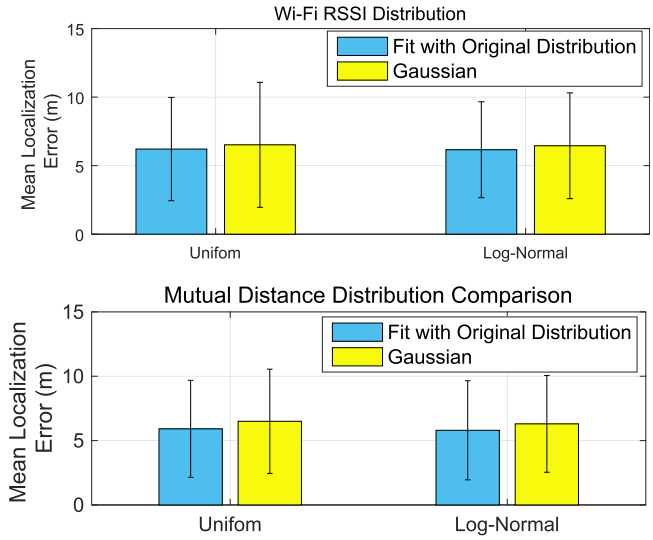


Fig. 24. *ME* and error standard deviations of Maxlifd (BC) under different distributions: (1) Wi-Fi RSSI distribution and (2) mutual distance distribution.

9 Conclusion

We have proposed Maxlifd, a novel joint maximum likelihood framework fusing wireless fingerprints with mutual distances for indoor localization. The mutual distances can be temporal or spatial (as obtained from dead reckoning, beacon-based, peer-assisted manner, etc.). Due to random signal fluctuation, both fingerprints and distance measurements are noisy in nature. Given the probability distributions in distances and fingerprints, Maxlifd formulates a single semi-definite programming problem, fusing the noisy fingerprints with uncertain distance measurement to localize the target based on the joint likelihood maximization. Maxlifd is generic, and hence is applicable to a wide range of sensing devices and wireless fingerprint signals.

We have implemented Maxlifd using INS (temporal), peer-assisted and beacon-assisted distance measurement (spatial), validating its accuracy and robustness. We have also evaluated its CRLB of localization errors, which provides theoretical insights for deployment.

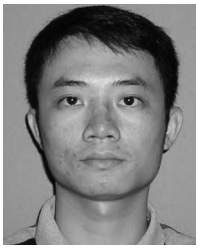
ACKNOWLEDGMENTS

This work was supported, in part, by the National Natural Science Foundation of China under Grant (61472455), and the Guangzhou Science Technology and Innovation Commission (GZSTI16EG14/201704030079).

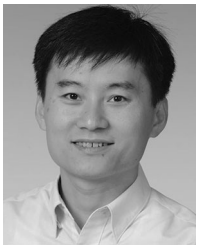
REFERENCES

[1] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. 19th Annu. Joint Conf. IEEE Comput. Commun. Societies*, 2000, pp. 775–784.
[2] D. Han, S. Jung, et al., "Building a practical Wi-Fi-based indoor navigation system," *IEEE Pervasive Comput.*, vol. 13, no. 2, pp. 72–79, Apr.–Jun. 2014.
[3] Y. Gao, Q. Yang, et al., "XINS: The anatomy of an indoor positioning and navigation architecture," in *Proc. 1st Int. Workshop Mobile Location-Based Serv.*, 2011, pp. 41–50.
[4] J. Jun, Y. Gu, et al., "Social-Loc: Improving indoor localization with social sensing," in *Proc. 11th ACM Conf. Embedded Netw. Sensor Syst.*, 2013, pp. 14:1–14:14.
[5] N. Banerjee, S. Agarwal, et al., "Virtual Compass: Relative positioning to sense mobile social interactions," in *Proc. Int. Conf. Pervasive*, Mar. 2010, pp. 1–21.
[6] H. Liu, J. Yang, et al., "Accurate WiFi based localization for smartphones using peer assistance," *IEEE Trans. Mobile Comput.*, vol. 13, no. 10, pp. 2199–2214, Oct. 2014.
[7] A. Rai, K. K. Chintalapudi, et al., "Zee: Zero-effort crowdsourcing for indoor localization," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, 2012, pp. 293–304.
[8] J. Seitz, J. Jahn, et al., "A hidden Markov model for urban navigation based on fingerprinting and pedestrian dead reckoning," in *Proc. 13th Conf. Inf. Fusion*, Jul. 2010, pp. 1–8.
[9] M. K. Hoang, J. Schmalenstroeer, et al., "A hidden Markov model for indoor user tracking based on WiFi fingerprinting and step detection," in *Proc. 21st Eur. Signal Process. Conf.*, Sep. 2013, pp. 1–5.
[10] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.
[11] S. He and K. G. Shin, "Geomagnetism for smartphone-based indoor localization: Challenges, advances, and comparisons," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 97:1–97:37, Dec. 2017.
[12] S. He and S. H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," *IEEE Commun. Surveys Tutorials*, vol. 18, no. 1, pp. 466–490, 2016.
[13] M. Youssef and A. Agrawala, "The Horus location determination system," *Wireless Netw.*, vol. 14, no. 3, pp. 357–374, Jun. 2008.
[14] A. Kushki, K. Plataniotis, and A. Venetsanopoulos, "Kernel-based positioning in wireless local area networks," *IEEE Trans. Mobile Comput.*, vol. 6, no. 6, pp. 689–705, Jun. 2007.
[15] L. Li, G. Shen, et al., "Experiencing and handling the diversity in data density and environmental locality in an indoor positioning service," in *Proc. 20th Annu. Int. Conf. Mobile Comput. Netw.*, 2014, pp. 459–470.
[16] C. Feng, W. Au, et al., "Received-signal-strength-based indoor positioning using compressive sensing," *IEEE Trans. Mobile Comput.*, vol. 11, no. 12, pp. 1983–1993, Dec. 2012.
[17] A. Hossain and W.-S. Soh, "Cramer-Rao bound analysis of localization using signal strength difference as location fingerprint," in *Proc. IEEE INFOCOM*, Mar. 2010, pp. 1–9.
[18] Y. Shen, H. Wymeersch, and M. Win, "Fundamental limits of wideband localization–Part II: Cooperative networks," *IEEE Trans. Inf. Theory*, vol. 56, no. 10, pp. 4981–5000, Oct. 2010.
[19] L. Doherty, K. S. J. Pister, and L. E. Ghaoui, "Convex position estimation in wireless sensor networks," in *Proc. 20th Annu. Joint Conf. IEEE Comput. Commun. Societies*, 2001, pp. 1655–1663.
[20] C. Chang and A. Sahai, "Estimation bounds for localization," in *Proc. 1st Annu. IEEE Commun. Soc. Conf. Sensor Ad Hoc Commun. Netw.*, Oct. 2004, pp. 415–424.
[21] Z. Xiao, H. Wen, et al., "Indoor tracking using undirected graphical models," *IEEE Trans. Mobile Comput.*, vol. 14, no. 11, pp. 2286–2301, Nov. 2015.
[22] W. Sun, J. Liu, et al., "MoLoc: On distinguishing fingerprint twins," in *Proc. IEEE 33rd Int. Conf. Distrib. Comput. Syst.*, Jul. 2013, pp. 226–235.
[23] B. Ferris, D. Fox, and N. D. Lawrence, "WiFi-SLAM using Gaussian process latent variable models," in *Proc. 20th Int. Joint Conf. Artif. Intell.*, 2007, pp. 2480–2485.
[24] H. Wang, S. Sen, et al., "No need to war-drive: Unsupervised indoor localization," in *Proc. 10th Int. Conf. Mobile Syst. Appl. Serv.*, 2012, pp. 197–210.
[25] R. Nandakumar, K. K. Chintalapudi, and V. N. Padmanabhan, "Centaur: Locating devices in an office environment," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, 2012, pp. 281–292.
[26] X. Zhao, Z. Xiao, et al., "Does BTLE measure up against WiFi? A comparison of indoor location performance," in *Proc. 20th Eur. Wireless Conf.*, May 2014, pp. 1–6.
[27] R. Faragher and R. Harle, "Location fingerprinting with Bluetooth low energy beacons," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2418–2428, Nov. 2015.
[28] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proc. 13th Usenix Conf. Netw. Syst. Des. Implementation*, 2016, pp. 165–178.
[29] M. Kotaru, K. Joshi, et al., "SpotFi: Decimeter Level Localization Using WiFi," in *Proc. ACM Conf. Special Interest Group Data Commun.*, 2015, pp. 269–282.
[30] J. Xiong, K. Sundaresan, and K. Jamieson, "ToneTrack: Leveraging frequency-agile radios for time-based indoor wireless localization," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 537–549.
[31] K. Liu, X. Liu, and X. Li, "Guoguo: Enabling fine-grained smartphone localization via acoustic anchors," *IEEE Trans. Mobile Comput.*, vol. 15, no. 5, pp. 1144–1156, May 2016.
[32] C. Wu, Z. Yang, and Y. Liu, "Smartphones based crowdsourcing for indoor localization," *IEEE Trans. Mobile Comput.*, vol. 14, no. 2, pp. 444–457, Feb. 2015.
[33] S. Ji, K.-F. Sze, et al., "Beyond convex relaxation: A polynomial-time non-convex optimization approach to network localization," in *Proc. IEEE INFOCOM*, 2013, pp. 2499–2507.
[34] X. Fan, J. Song, et al., "Universal binary semidefinite relaxation for ML signal detection," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4565–4576, Nov. 2013.
[35] A. Symington and N. Trigoni, "Encounter based sensor tracking," in *Proc. 13th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2012, pp. 15–24.
[36] Y. Shang, W. Ruml, et al., "Localization from mere connectivity," in *Proc. 4th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2003, pp. 201–212.
[37] F. Xiao, C. Sha, et al., "Noise-tolerant localization from incomplete range measurements for wireless sensor networks," in *Proc. IEEE Conf. Comput. Commun.*, 2015, pp. 2794–2802.
[38] S. He, S.-H. G. Chan, et al., "Fusing noisy fingerprints with distance bounds for indoor localization," in *Proc. IEEE Conf. Comput. Commun.*, Apr. 2015, pp. 2506–2514.
[39] E. Martin, O. Vinyals, et al., "Precise indoor localization using smart phones," in *Proc. 18th ACM Int. Conf. Multimedia*, 2010, pp. 787–790.
[40] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[41] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge University Press, 2004.
[42] F. Li, C. Zhao, et al., "A reliable and accurate indoor localization method using phone inertial sensors," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 421–430.
[43] S. He, T. Hu, and S. H. G. Chan, "Toward practical deployment of fingerprint-based indoor localization," *IEEE Pervasive Comput.*, vol. 16, no. 2, pp. 76–83, Apr. 2017.
[44] L. Devroye, L. Györfi, and G. Lugosi, *A Probabilistic Theory of Pattern Recognition*. Berlin, Germany: Springer, 1996.
[45] Z.-Q. Luo, W.-K. Ma, et al., "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.
[46] S. J. Benson, Y. Ye, and X. Zhang, "Solving large-scale sparse semidefinite programs for combinatorial optimization," *SIAM J. Optimization*, vol. 10, no. 2, pp. 443–461, 2000.
[47] K. Lui, W.-K. Ma, et al., "Semi-definite programming algorithms for sensor network node localization with uncertainties in anchor positions and/or propagation speed," *IEEE Trans. Signal Process*, vol. 57, no. 2, pp. 752–763, Feb. 2009.
[48] F. Dabek, R. Cox, et al., "Vivaldi: A decentralized network coordinate system," in *Proc. Conf. Appl. Technol. Archit. Protocols Comput. Commun.*, Aug. 2004, pp. 15–26.
[49] N. Alsindi, R. Raulefs, and C. Teolis, *Geolocation Techniques: Principles and Applications*. Berlin, Germany: Springer, 2012.
[50] H. Liu, Y. Gan, et al., "Push the limit of WiFi based localization for smartphones," in *Proc. 18th Annu. Int. Conf. Mobile Comput. Netw.*, Sep. 2012, pp. 305–316.

[51] L. Zhang, K. Liu, et al., "Montage: Combine frames with movement continuity for realtime multi-user tracking," in *Proc. IEEE INFOCOM*, Apr. 2014, pp. 799–807.
[52] S. He, W. Lin, and S. H. G. Chan, "Indoor localization and automatic fingerprint update with altered AP signals," *IEEE Trans. Mobile Comput.*, vol. 16, no. 7, pp. 1897–1910, Jul. 2017.
[53] X. Zhang, Z. Yang, et al., "Free market of crowdsourcing: Incentive mechanism design for mobile sensing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 25, no. 12, pp. 3190–3200, Dec. 2014.
[54] L. Chang, J. Xiong, et al., "iUpdater: Low cost RSS fingerprints updating for device-free localization," in *Proc. IEEE 37th Int. Conf. Distrib. Comput. Syst.*, Jun. 2017, pp. 900–910.
[55] Y.-C. Chen, J.-R. Chiang, et al., "Sensor-assisted Wi-Fi indoor location system for adapting to environmental dynamics," in *Proc. 8th ACM Int. Symp. Model. Anal. Simul. Wireless Mobile Syst.*, 2005, pp. 118–125.
[56] Y. C. Tung and K. G. Shin, "Use of phone sensors to enhance distracted pedestrians' safety," *IEEE Trans. Mobile Comput.*, vol. 17, no. 6, pp. 1469–1482, Jun. 2018.

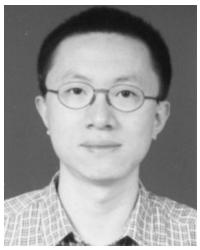


Suining He received the PhD degree from the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology (HKUST), in 2016. He was selected as a Google PhD fellow in Mobile Computing (China and East Asia), in 2015. He is currently working as a postdoctoral research fellow with the Real-Time Computing Laboratory, The University of Michigan, Ann Arbor. His research interest includes indoor localization, smartphone sensing, and mobile computing. He is a member of the IEEE.



S.-H. Gary Chan (S'89-M'98-SM'03) received the BSE degree (highest honor), in electrical engineering from Princeton University, Princeton, NJ, in 1993, with certificates in applied and computational mathematics, engineering physics, and engineering and management systems, and the MSE and PhD degrees in electrical engineering from Stanford University, Stanford, CA, in 1994 and 1999, respectively, with a minor in business administration. He is currently professor with the Department of Computer Science and Engineering, The Hong Kong University of Science and Technology (HKUST), Hong Kong. He is also the director of the Entrepreneurship Center, and chair of the Committee on Entrepreneurship Education Program, Center for Education Innovation, HKUST. His research interest includes multimedia networking, mobile computing, data/user analytics, and IT entrepreneurship. He has been an associate editor of the *IEEE Transactions on Multimedia* (2006-2011), and a vice-chair of the Peer-to-Peer Networking and Communications Technical Sub-Committee of the IEEE Comsoc Emerging Technologies Committee (2006-2013). He is and has been guest editor of *Elsevier Computer Networks* (2017), the *ACM Transactions on Multimedia Computing, Communications, and Applications* (2016), the *IEEE Transactions on Multimedia* (2011), the *IEEE Signal Processing Magazine* (2011), the *IEEE Communication Magazine* (2007), and *Springer Multimedia Tools and Applications* (2007). He was the TPC chair of the IEEE Consumer Communications and Networking Conference (IEEE CCNC) 2010, Multimedia Symposium of IEEE Globecom (2007 and 2006), IEEE ICC (2007 and 2005), and the Workshop on Advances in Peer-to-Peer Multimedia Streaming in ACM Multimedia Conference (2005). He has co-founded several startups deploying his research results. Due to their innovations and commercial impacts, his projects have received local and international awards (2012-2018). He is the recipient of the Google Mobile 2014 Award (2010 and 2011) and the Silver Award of Boeing Research and Technology (2009). He has been a visiting professor and researcher in Microsoft Research (2000-2011), Princeton University (2009), Stanford University (2008-2009), and the University of California, Davis (1998-1999). He was undergraduate programs coordinator with the Department of Computer Science and Engineering (2013-2015), director of the Sino Software Research Institute (2012-2015), co-director of the Risk Management and Business Intelligence Program (2011-2013), and director of the Computer Engineering Program (2006-2008) with HKUST. He was a William and Leila fellow with Stanford University (1993-94), and the recipient of the Charles Ira Young Memorial Tablet and Medal, and the POEM Newport Award of Excellence at Princeton (1993). He is a chartered fellow of The Chartered Institute of Logistics and Transport (FCILT), and a member of honor societies Tau Beta Pi, Sigma Xi, and Phi Beta Kappa. He is a senior member of the IEEE.

Lei Yu received the bachelor's degree in software engineering from Sun Yat-sen University, and the double master's degree in electrical and computer engineering from Carnegie Mellon University and Sun Yat-sen University. He received many prizes during his undergraduate period, such as the National Scholarship of China, Google Excellence Scholarship, etc.



Ning Liu received the BS degree in computational mathematics from Northwest University, Xian, China, and the PhD degree in computer science from Sun Yat-sen University (SYSU), Guangzhou, China, in 1996 and 2004, respectively. He is currently an associate professor with Sun Yat-sen University. He has served as a reviewer for several important conferences and journals. His current research interests include computer vision, peer-to-peer networks, and machine learning algorithms.