Scalability of Wireless Fingerprinting based Indoor Localization Systems

Yingling Mao*, Ke Liu[†], Hao Li[†], Xiaohua Tian[†], Xinbing Wang[†]
* School of Mathematical Sciences
[†]School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University, China
Email:{sandglassmyl, kqft2540253972, iamgeek, xtian, xwang8}@sjtu.edu.cn

Abstract—**Fingerprinting indoor localization systems have been studied in different perspectives in the past decades; however, a vitally important piece in the puzzle is still missing: how does the system scale with the number of users? In this paper, we present a theoretical study of the issue, where the upper and lower bound of the system's localization reliability with respect to the number of users are derived. Results of our theoretical analysis can be verified by experiments thus can provide meaningful guidance for practical system design, which is in contrast to the scaling-law work utilizing asymptotical analysis that is valid only under unverifiable extreme conditions. The theoretical and experimental results of our work reveal two interesting observations, which shed light on the insight into the scalability of the fingerprinting localization system: First, the localization reliability drops dramatically before the number of users increases to a critical point and then decreases smoothly, where the critical point tends to appear when the number of users equals the number of access points (APs) deployed in the service region; second, even if the number of users approaches to infinity, the fingerprinting localization system still retains certain level of reliability.**

I. INTRODUCTION

Extensive efforts have been devoted to wireless fingerprinting based indoor localization in the past two decades [1], [3]–[10], [12], which leverages the existing Wi-Fi access points (APs) thus saves the cost of deploying dedicated infrastructure particularly for localization. The basic idea of fingerprinting localization is to utilize the spatial feature of wireless radio signals from Wi-Fi APs to distinguish one location from another, where the specific feature is termed as the wireless fingerprint. The mostly used fingerprint is the received signal strength (RSS) [3]–[5], [9], [10], [12], which can be handily obtained from the programmable interface of the mobile device. The fingerprinting localization approach is normally divided into two phases, where the radio map associating the location information with RSS fingerprints is constructed in the offline phase, and users' locations are estimated in the online phase by matching the user's reported fingerprints with those stored in the offline phase.

Although recent years have witnessed localization systems achieving centimeter even millimeter level accuracy, which utilize channel state information (CSI) of wireless signals [14], [15], RFID [16] and acoustic signals [17], such systems require special infrastructures such as APs capable of retrieving CSI, high-end RFID readers and acoustic anchor points. Consequently, the wireless fingerprinting based localization system seems more suitable for large-scale deployment [2],

[18]–[20]. While a large body of work has been available to reveal insights into the fingerprinting localization in different perspectives, such as localization accuracy and reliability [8], [12], [21], [24], [25], influence of user mobility [13] and crowdsourcing based radio map construction [19], [20], a vitally important piece in the puzzle is still missing if large-scale deployment is the target: how does the fingerprinting localization system scale with the number of users?

In particular, people are moving around in the indoor space, which continuously changes the radio propagation environment. Intuitively, the more human bodies are within the space, the more serious the consequent shadowing and multipath effects are, which results in more complicated radio propagation environment and more deviation in radio map constructed in the offline phase to the current reality. Although the human body shadowing effect is investigated for the purposes of radio propagation modeling and antenna design [26]–[28], the scalability of fingerprinting localization systems is still unknown.

In this paper, we study the scalability issue in a theoretical perspective, where the upper and lower bound of the system's localization reliability with respect to the number of users are derived. Our contributions are as following:

First, we formulate the scalability issue into the problem of finding how the localization reliability deteriorates with respect to the number of users N (§**Section II**). The fingerprinting localization approach can be modeled as a mapping from the fingerprints sample space to the physical space, where it has been found that a user can be localized within δ neighborhood of the real location only if the user's reported fingerprints fall in certain area E of the sample space [11], [12]. Based on the theory, we infer that once the localization system is deployed, the area E with respect to the specific radio propagation environment is determined. When users appear and move around in the service area, the radio propagation environment change makes characteristics of wireless fingerprints such as the mean and variance vary, but the system still uses the predetermined E to preform location estimation. Then the inconsistency between the predetermined E and the changed fingerprints characteristics incurs deterioration of localization reliability, where more users result in more serious inconsistency and deterioration.

Second, we present mathematical expressions of the deteriorated localization reliability R' with respect to the number

of users N, which is obtained by two-step derivations. The first step is to find the upper and lower bound of R' with respect to the number of impacted APs m; this involves performing integration of a Gaussian probability distribution function (PDF) with only partial information available over a hard-to-profile high-dimensional E in the fingerprints sample space, which is realized through multiple times of coordinate system construction and transformation (§**Section III**). The second step is to find the bounds of m with respect to the number of users N, where the challenge is that it is difficult to determine the interference region associated with the AP and the user. We address the issue by designing simple measure-and-infer procedures, based on which we can obtain the result of N-user case with experimental results of 2-user case (§**Section IV**). Our main theoretical result (§**Section V**) can be verified by experiments thus can provide meaningful guidance for practical system design, which is in contrast to the scaling-law work utilizing asymptotical analysis that is valid only under unverifiable extreme conditions.

We present our experimental and numerical results to verify the theoretical analysis; we reveal two interesting observations, which shed light on the insight into the scalability of the fingerprinting localization system (§**Section VI**): First, the localization reliability drops dramatically before the number of users increases to a critical point and then decreases smoothly, where the critical point tends to appear when the number of users equals the number of access points (APs) deployed in the service region; second, even if the number of users approaches to infinity, the fingerprinting localization system still retains certain level of reliability. Due to limitation of the space, details of mathematical derivations and proofs are put in our technical report [36].

II. PROBLEM FORMULATION

A. Interference Region

We consider an indoor space denoted by S, where there are M APs distributed along the boundaries of S and N users. As is done by many fingerprinting localization systems [1], [2], we divide S into small square grids, where the center of each grid is set as a reference point (RP). In the offline phase, the fingerprints are measured on RPs; in the online phase, the nearest RP to the user's real location is supposed to be the system's estimated location. If a user performs RSS measurement with respect to an AP, then a transmitter-receiver (T-R) pair is formed. Due to the line-of-sight blockage and mutipath effects, when a human body appears in some grids between the T-R pair, the RSS measurement will be impacted. We term those grids as the *interference region* of the corresponding T-R pair. For users, we have no assumption of the specific probability model describing how likely they will appear in which location, but we assume that the same model applies to all the users.

B. Localization Reliability Model

We use a 2-D vector $\vec{r}$ to represent a location in S, then the RSS observed at $\vec{r}$ can be modeled by a random variable

[11], [12]:

$$\mathcal{P}(\vec{r}) = \mu(\vec{r}) + \sigma \mathcal{Y}, \tag{1}$$

where $\mu(\vec{r})$ represents how the mean of RSS readings varies with respect to locations, $\mathcal{Y}$ is the normalized Gaussian random variable with $\mathcal{Y} \sim \mathcal{N}(0, 1)$ and σ is a constant representing the variance of the received signal [12]. Modeling RSS as a Gaussian distributed random variable is adopted in a large number of work in the literature [1], [3], [6], [29]–[31], where the rationale of the modeling is verified with comprehensive experimental results [12].

The location estimation process is essentially a mapping from the fingerprints sample space to the physical space, where it has been found that a user can be localized within δ neighborhood of the real location only if the user's reported fingerprints fall in certain area E of the sample space [11], [12]. We use the simple example as follows to show the modeling approach.
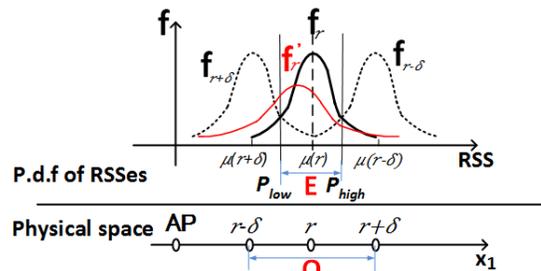

Fig. 1. Reliability model.

Consider an 1-D physical space where the single AP is located at one end, as shown in Fig. 1; the corresponding sample space is shown in the upper part of Fig. 1, where each location is distinguished by the observed mean value of the RSS fingerprints μ, and the PDF of the fingerprints observed at the location following Gaussian distribution based on Eq. (1). According to the principle of maximum likelihood estimation (MLE), the user can be localized in the region Q only if the reported RSS fingerprints fall within the region E in the sample space as shown in Fig. 1. Corresponding reliability R can be obtained by integrating Gaussian PDF over region E in the sample space. The localization in 2-D physical space with high-dimensional sample space is modeled in the similar manner but presents more mathematical challenges. It is proved in [11], [12] that region E in the high-dimensional sample space is a hyper-cylinder with the orthogonal cross-section in the shape of an ellipse.

Our work in this paper leverages the localization reliability model in [11], [12]; however, our focus is on scalability, which is not mentioned in [11], [12]. Moreover, we consider the practical localization scenario, where fingerprints are collected only at RPs [1], [3]–[10]. We define the localization reliability as the probability that a user can be correctly localized in a square region surrounded by the other 8 RPs, and the accuracy is the unit length of the grid. This is in contrast to the model in [11], [12], where it is implicitly but impractically assumed that each point in the indoor space has to be surveyed.

C. Localization Performance Deterioration by Human Body Blockage.

The reliability and accuracy model in [11], [12] as described above implicitly assumes that the PDF representing each location in the offline does not change in the online phase. However, the radio propagation environment in the two phases is factually different in practice, where an important reason is the human body blockage effect. Comprehensive studies show that the mean and the variance of the wireless signals that are observed at a location will change, if the signal's propagation path is changed from the line-of-sight path to the non-line-of-sight one [32], [33]. The deviation of the mean and the variance is especially notable with presence of the human body blockage [26]–[28]. The deviation makes the shape of the PDF represent the same location $\vec{r}$ change in the two phases. Figure 1 shows an example of the blockage effect, where the PDF in the offline phase is $f_{\vec{r}}$ while it becomes $f'_{\vec{r}}$ in the online phase.

D. Strategy of Deriving the Scalability

Consider a fingerprinting localization system, we assume that there is only one user performing the site survey in the offline phase. Then the fingerprints database is constructed without being impacted by other people. If the other $N-1$ users show up, the blockage effect will deteriorate the system performance by reducing the localization reliability. Since the fingerprints database remains unchanged after the user's site survey, then the region E is static. We use R' to denote the system reliability after show-up of the other $N-1$ users, which can be obtained by integrating $f'_{\vec{r}}$ over E.

Although the deformation of $f'_{\vec{r}}$ with respect to $f_{\vec{r}}$ is unpredictable, we intuitively have $R' \leq R$, which indicates the deterioration in localization performance. The physical meaning is that the mismatch of fingerprints' features in the two phases can not result in higher localization reliability; mathematically, due to the PDF nature of $f'_{\vec{r}}$ as shown in Fig. 1, it is impossible for the shape of $f'_{\vec{r}}$ to be both higher and wider than that of $f_{\vec{r}}$, and the peak of $f'_{\vec{r}}$ could deviate from the mid-point of the integration domain, then the result of integration for $f'_{\vec{r}}$ over E is no greater than that for $f_{\vec{r}}$. However, due to the abstraction and high dimension of the sample space and the unknown parameters of the Gaussian PDFs, finding the expression of R' will be challenging. Moreover, the PDF can be changed only if some users are standing in the influence region. The more users are presenting in the indoor space, the more likely there are some users standing in the influence region of some T-R pairs, thus more reliability deterioration can occur.

Consequently, *resolving the scalability issue of the fingerprinting indoor localization system is essentially finding how the localization reliability of the system deteriorates with respect to the number of users N.* Situations such that more than one people are present when the site survey is being performed, or the database is updated periodically are purposely not taken into account in our model, because such detailed scenarios can be easily extended from our model.
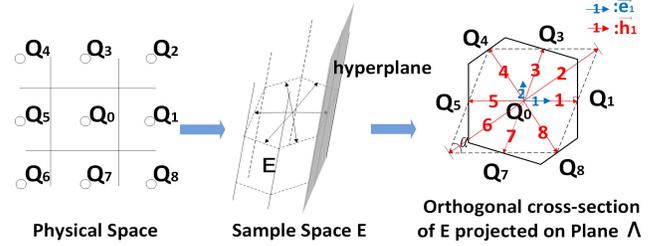

Fig. 2. Revised localization reliability model.

III. LOCALIZATION RELIABILITY DETERIORATION BY BLOCKAGE

Our localization reliability analysis is based on the model as shown in Fig. 2. It is supposed that the user's actual location in the physical space is at Q_0. In order to make the system accurately estimate the user's location, the user's reported RSSs must be the measurement outcome which falls into the region E. The orthogonal cross-section of E projected on a plane is illustrated in the rightmost sub-figure of Fig. 2. With similar derivation process as shown in [11], [12], the resulted reliability R can be obtained. For the convenience of readers, we put the brief derivation process in our technical report [36], so that the methodology used in the following section can be understood. The reliability R factually reflects the localization performance when there is no blockage; however, when users come into the service region, the PDF of the fingerprints observed at $\vec{r}$ changes from $f_{\vec{r}}$ into $f'_{\vec{r}}$, but the event E does not change since it has been determined in the offline phase. We are to derive the localization probability with the existence of the users' blockage effect, and then find the upper and lower bound of the reliability deterioration.

A. Localization Reliability with Blockage

Considering a T-R pair associated with an influence region, if another person appears in the influence region, the RSS value observed by the user is actually profiled by $f'(\vec{P})$. The resulted reliability becomes:

$$R' = \int_E f'(\vec{P})d\vec{P} = \int_E \prod_{i=1}^{M} \frac{1}{\sigma'_i \sqrt{2\pi}} e^{-\frac{(P_i - \mu'_i(\vec{r}))^2}{2\sigma'^2_i}} dP_1...dP_M,$$

where P_i is the user's measurement of the RSS with respect to AP_i, $\mu'_i(\vec{r})$ and σ'_i represent the deviated mean and standard variance resulted from the human body blockage effect, in contrast to the parameters $\mu_i(\vec{r})$ and σ_i without blockage. Note that E does not change since the system still uses the fingerprints database constructed by the single user or survey worker, that is, $E = \{\vec{o} | \vec{h_j}(\vec{o} - \vec{h_j}) \leq 0, j = 1,...,8, \vec{o} = (Y_1,...,Y_M)^T\}$, which is rigorously derived in [36]. For the convenience of derivation, we apply a new coordinates system to the sample space. Let $Y'_i = \frac{P_i - \mu'_i}{\sigma'_i}$, then $Y'_i = \frac{\sigma_i}{\sigma'_i} Y_i + \frac{\mu_i - \mu'_i}{\sigma'_i} \triangleq A_i Y_i + B_i$, where it is straightforward that $A_i > 1$ and $B_i > 0$, then

$$R' = \int_E \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}} e^{-\frac{Y'^2_i}{2}} dY'_1 dY'_2...dY'_M. \tag{2}$$

The coordinates transformation provides a neat expression of $f'(\vec{P})$, but also changes the form of E. We must figure out the expression of E, after which R' can be found.

We denote $\vec{o'} = (Y_1', Y_2', ..., Y_M')^T = (A_1 Y_1 + B_1, ..., A_M Y_M + B_M)^T$, and we need to find the form of restrictions to the event E w.r.t $\vec{o'}$. We define the following vector w.r.t. the coefficient A_i and B_i, $i = 1, 2, ..., M$. For any $j \in \{1, 2, ..., 8\}$, we define $\vec{l} = (\frac{B_1}{A_1}, \frac{B_2}{A_2}, ..., \frac{B_M}{A_M})^T$, $\vec{H_j} = (\frac{\mu_1(\vec{r_j}) - \mu_1(\vec{r})}{2\sigma_1 A_1}, ..., \frac{\mu_M(\vec{r_j}) - \mu_M(\vec{r})}{2\sigma_M A_M})^T$ and $\vec{h_j'} = k_j \vec{H_j}$, where k_j is a constant, $k_j = \frac{|\vec{h_j}|^2 + \vec{l}\vec{h_j}}{|\vec{H_j}|^2}$.

We then have: $\vec{H_1} = (\frac{|\nabla \mu_1(\vec{r})|}{2\sigma_1 A_1} \varepsilon \cos(\phi_1), ..., \frac{|\nabla \mu_M(\vec{r})|}{2\sigma_M A_M} \varepsilon \cos(\phi_M))^T$, $\vec{H_2} = \vec{H_1} + \vec{H_3}$, $\vec{H_3} = (\frac{|\nabla \mu_1(\vec{r})|}{2\sigma_1 A_1} \varepsilon \sin(\phi_1), ..., \frac{|\nabla \mu_M(\vec{r})|}{2\sigma_M A_M} \varepsilon \sin(\phi_M))^T$, $\vec{H_4} = -\vec{H_1} + \vec{H_3}$, $\vec{H_5} = \vec{H_3}$, $\vec{H_6} = -\vec{H_1} - \vec{H_3}$, $\vec{H_7} = -\vec{H_3}$, $\vec{H_8} = \vec{H_1} - \vec{H_3}$ and $\vec{h_j'} = k_j \vec{H_j}$, $j = 1, 2, ..., 8$. Moreover, we have $k_j + k_{j+4} = \frac{2|\vec{h_j}|^2}{|\vec{H_j}|^2}$ for $j = 1, 2, 3, 4$ and $0 < k_j + k_{j+4} \leq 2$, according to the definitions of $\vec{h_j}$ and $\vec{H_j}$.

Based on the definitions and analysis above, it is interesting to find

$$\vec{h_j'}(\vec{o'} - \vec{h_j'}) = k_j(\vec{H_j}\vec{o'} - k_j\vec{H_j}^2)$$
$$= k_j[(\vec{h_j}\vec{o} + \vec{l}\vec{h_j}) - \frac{|\vec{h_j}|^2 + \vec{l}\vec{h_j}}{|\vec{H_j}|^2} \cdot \vec{H_j}^2]$$
$$= k_j(\vec{h_j}\vec{o} - |\vec{h_j}|^2) = k_j\vec{h_j}(\vec{o} - \vec{h_j}), \quad (3)$$

which establishes the connection between E in the two coordinates systems. Then we have the new expression of E: $E = \{\vec{o'} | k_j \vec{h_j'}(\vec{o'} - \vec{h_j'}) \leq 0, j = 1, ..., 8, \vec{o'} = (Y_1', ..., Y_M')^T\}$.

It can be seen that $f(\vec{P})$ and $f'(\vec{P})$ have the same form although they are in different coordinates systems. And what we need to do is to find the event E in the new coordinates system $\{Y_i'\}$. Since $\vec{h_j'} = k_j \vec{H_j}$, $j = 1, 2, ..., 8$, the coefficient k_j factually reflects how the blockage effect will impact the event E, which results in R'. Figure 3(a), (b), (c) respectively show three possible scenarios how the event E can be impacted by users: a) If there is a single user in the system just like the scenario of site survey, then the fingerprints are not impacted and $k_1 = k_2 = ... = k_8 = 1$, which means that the shape and location of E in sample space are the same as shown in Fig.2; b) if the RSS readings are insignificantly influenced, then $k_j > 0, j = 1, 2, ..., 8$, and the shape and position of E will change but in an insignificant manner, which is due to the relationship among E's boundaries; c) if $k_5 < 0$, both the shape and location of E will change significantly.

Then we rotate the coordinates system $\{Y_i'\}$ to another orthogonal basis $\{\vec{g_i}\}$, $i = 1, 2, ..., n$, where $\vec{g_1}$ is parallel to $\vec{H_1}$ and $\vec{g_2}$ is in the plane Λ which is determined by $\vec{H_1}$ and $\vec{H_3}$. We use β to denote the angle between $\vec{H_1}$ and $\vec{H_3}$, then $\vec{H_1} = |\vec{H_1}| \cdot \vec{g_1}$, $\vec{H_3} = |\vec{H_3}| \cos\beta \cdot \vec{g_1} + |\vec{H_3}| \sin\beta \cdot \vec{g_2}$. Suppose that $\vec{o'} = \sum_i a_i \vec{g_i}$, then we have $-k_5|\vec{H_1}| \leq a_1 \leq k_1|\vec{H_1}|$ and $L' \leq a_2 \leq H'$, where
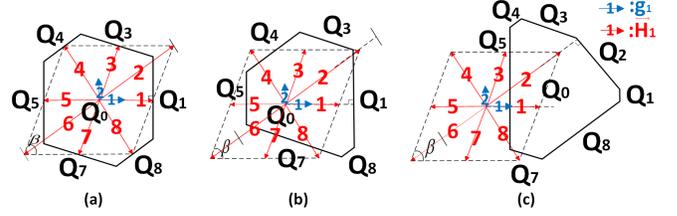

Fig. 3. Influence of blockage effect on E in the sample space.

$$L' = -a_1 \cot\beta + \frac{1}{\sin\beta} \cdot \max\{-k_7|\vec{H_3}|,$$
$$-k_6|\vec{H_3}| - \frac{|\vec{H_1}|}{|\vec{H_3}|}[k_6(|\vec{H_1}| + 2|\vec{H_3}|\cos\beta) + a_1],$$
$$-k_8|\vec{H_3}| - \frac{|\vec{H_1}|}{|\vec{H_3}|}[k_8(|\vec{H_1}| - 2|\vec{H_3}|\cos\beta) - a_1]\}, \quad (4)$$
$$H' = -a_1 \cot\beta + \frac{1}{\sin\beta} \cdot \min\{k_3|\vec{H_3}|,$$
$$k_2|\vec{H_3}| + \frac{|\vec{H_1}|}{|\vec{H_3}|}[k_2(|\vec{H_1}| + 2|\vec{H_3}|\cos\beta) - a_1],$$
$$k_4|\vec{H_3}| + \frac{|\vec{H_1}|}{|\vec{H_3}|}[k_4(|\vec{H_1}| - 2|\vec{H_3}|\cos\beta) + a_1]\}. \quad (5)$$

Since $|\vec{o'}| = \sum_{i=1}^{n} Y_i'^2 = \sum_{i=1}^{n} a_i^2$, which implies the norm of $\vec{o'}$ is unchanged after the rotation of the coordinate system, then the probability that the system correctly estimates the user's location is

$$R' = \int_E \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}} e^{-\frac{Y_i'^2}{2}} dY_1' dY_2'...dY_M'$$
$$= \int_{-k_5|\vec{H_1}|}^{k_1|\vec{H_1}|} \int_{L'}^{H'} \int_{-\infty}^{+\infty} ... \int_{-\infty}^{+\infty} \prod_{i=1}^{M} \frac{1}{\sqrt{2\pi}} e^{-\frac{a_i^2}{2}} da_1...da_M$$
$$= \int_{-k_5|\vec{H_1}|}^{k_1|\vec{H_1}|} \frac{1}{\sqrt{2\pi}} e^{-\frac{a_1^2}{2}} da_1 \cdot \int_{L'}^{H'} \frac{1}{\sqrt{2\pi}} e^{-\frac{a_2^2}{2}} da_2. \quad (6)$$
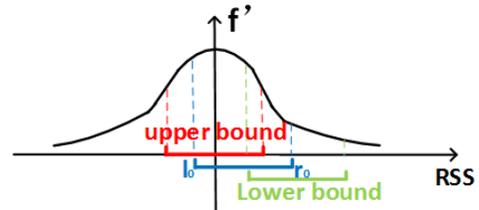
B. Finding Bounds of R'

Fig. 4. Finding bounds of R'.

We can see that Eq. 6 is the expression of localization reliability, where we transfer the deviation of the Gaussian PDF due to the environment change to that of the area E. Obtaining the concrete expression of R' requires to know the exact amount of the deviation incurred by the environment change as illustrated in Fig. 3, which is almost impossible in practice. However, we could manage to find the upper and lower bounds of R', so that the reliability deterioration incurred by the environment change can be quantified.

We here use a simple example in the 1-D sample space to explain our basic idea for deriving the bounds, which is as

shown in Fig. 4. We know that the event E is determined in the offline phase, but the event E represented by the blue line segment in the figure is actually the event in the transformed coordinate systems. The coordinate system transformation operations described in the previous subsection could change both the size and the position of E, as shown in the Fig. 3, which is to have a neat expression of f'. The reliability R' can be obtained by performing integration of f' over E, where it is relatively easy to find the size but difficult to find the displacement of E in mathematical derivations. The difficulty is particularly significant when performing high-dimensional integration, where the specific derivation process to be presented in following subsections can provide a better understanding.

Although it is difficult to find the displacement of E with respect to the origin of the coordinate system as shown in Fig. 4, it is straightforward to see that the upper bound of R' can be obtained if we move E left to the position of the line segment in red. For the lower bound, it is obvious that 0 is an option, which however provides negligible information. We want to find the greatest computable lower bound so that the reliability can be tightly bounded. In the process of mathematical derivation, we keep moving the event E towards the right until the bound is found. This is the basic idea how we derive the bound of R'. For our problem, obtaining the upper and lower bounds is factually to carry out the process above with a 2-D Gaussian PDF, which is very complicated and long thus put in the technical report [36] due to limitation of the space. The resulted bounds are as following:

Theorem 1. *For a localization system with the fingerprinting approach, if there are m APs influenced by the human body blockage effect, then*

$$e^{-\Theta(g(m))}(1 - e^{-\Theta(g(m))}) \le R'(E) \le 1 - e^{-\Theta(g(m))1}, \quad (7)$$

where $g(m) = \frac{M^2}{(1-A^2)\cdot m + A^2 \cdot M}$; m is the number of APs that are blocked; A is the greatest ratio of σ_i and σ'_i with $A < 1$.

In the numerical and experimental results to be presented later, we adopt the result in [32], [33] and let $A = \frac{3}{4}$.

IV. NUMBER OF IMPACTED ACCESS POINTS W.R.T. NUMBER OF USERS

At first, we consider to use the specific shape of the interference region to derive the relationship between the number of APs m and the number of users N. We note that existing work implicitly models the interference region in the shape of the ellipse [34], [35]. However, our experiments which can be seen in the technical report [36] show the ellipse model is too theoretical and has a high requirement for the environment. In practice, the interference region can not always be modeled as ellipse.

---

[1]We abuse $\Theta(\cdot)$ in this paper and modify the definition as following: $f(n) = \Theta(g(n))$ means $f(\cdot)$ is upper and lower bounded by $g(\cdot)$, that is, $\exists k_1^*, k_2^* > 0, \forall n > 0 : k_1^* \cdot g(n) \le f(n) \le k_2^* \cdot g(n)$. Note that the bounds are valid for any given value of n. They are not asymptotical bounds, and there is no need for $n \to \infty$, which is in contrast to the traditional definition of Θ.

Another method to find the number of impacted APs by different numbers of users is to conduct experiments; however, experimenting all scenarios with different combinations of the number of users and the positions of users can be labor-intensive. In this section, we present our method to find the number of impacted APs by a number of users, where there is no need for knowledge of the specific shape of the interference region or labor-intensive experiments.

A. Bounding the Number of Impacted APs

Due to the difficulty for accurate modeling the shape of the influence region, we switch to another method to resolve the issue. It is interesting to find that the number of impacted APs in the N-user case can be derived from the simple 2-user case, the result of which can be easily obtained with simple experiments.

Assume that the location of a normal user is denoted by $\vec{r}$, and the interferer's location is $\vec{r^*}$, which causes $\hat{m}$ APs to be impacted. This means that the interferer enters into the interference regions of the T-R links between the normal user and those $\hat{m}$ APs. As the APs' locations are fixed, the influence regions of the T-R links are determined by $\vec{r}$; therefore, $\hat{m}$ is determined by $\vec{r}$ and $\vec{r^*}$, which can be denoted by

$$\hat{m} = \hat{m}_{\vec{r}}(\vec{r^*}), \forall \vec{r} \in S, \quad (8)$$

where $\hat{m}_{\vec{r}}(\cdot)$ is determined by the shape of the influence region. It is straightforward that $\hat{m}$ is a random variable with cumulative distribution function (CDF):

$$F_{\vec{r}}(x) = P(\hat{m} \le x) = P(\vec{r^*} \in \{\vec{r^*} | \hat{m}_{\vec{r}}(\vec{r^*}) \le x\}). \quad (9)$$

Theorem 2. *For a localization system with N users within the region S and M APs distributed along the region's boundaries, if m out of M APs are impacted by human body blockage effect, then*

$$(a+1)(1 - b^{N-1}) \le m \le \min\{c(N-1), M\}, \quad (10)$$

where a, b, c are determined by the practical radio propagation environment, and can be obtained through the procedures to be presented in Section IV-B.

Proof: Given a normal user in the system, the rest of the $N - 1$ users can be regarded as interferers denoted by U_i, $i = 1, 2, ..., N - 1$. We use m_i to denote the number of APs which have been impacted by U_i, and m is the total number of APs that have been impacted, then it is straightforward that

$$\begin{cases} m_1 + m_2 + ... + m_{N-1} \ge m, \\ \max\{m_1, m_2, ..., m_{N-1}\} \le m. \end{cases} \quad (11)$$

Since $m_1, m_2, ..., m_{N_1}$ are all random variables, we can transform the inequality set (11) into following:

$$\begin{cases} E(m_1 + m_2 + ... + m_{N-1}) & \ge m, \\ E(\max\{m_1, m_2, ..., m_{N-1}\}) & \le m. \end{cases} \quad (12)$$

• Upper bound of m.

According to the inequality set (12), we have

$$m \leq E(m_1 + m_2 + ... + m_{N-1}) = \sum_{i=1}^{N-1} E(m_i) \qquad (13)$$

Since all the users share the show-up probability in a given position, $E(m_1) = E(m_2) = ... = E(m_{N-1}) = c$ and c is independent of N and m, then

$$m \leq \sum_{i=1}^{N-1} E(m_i) = (N-1)E(m_1) = c(N-1), \forall N. \qquad (14)$$

• Lower bound of m.

For convenience of presentation, we use $F(x)$ to present $F_{\vec{r}}(x)$. By the inequality set (12), we have

$$m \geq E(\max\{m_1, m_2, ..., m_{N-1}\}).$$

If $\widetilde{m} = \max\{m_1, m_2, ..., m_{N-1}\}$, then $m \geq E(\widetilde{m})$. We know that

$$P(\widetilde{m} \leq x) = P(m_1 \leq x, m_2 \leq x, ..., m_{N_1} \leq x)$$
$$= \prod_{i=1}^{N-1} P(m_i \leq x) = F^{N-1}(x), \qquad (15)$$

where $x \in \{0, 1, ..., M\}$.

In particular, $F(M) = P(m_i \leq M) = 1$, thus $F^{N-1}(M) = 1$, and $\forall N$

$$E(\widetilde{m}) = \sum_{k=0}^{M} k \cdot P(\hat{m} = k)$$
$$= \sum_{k=0}^{M} k \cdot [P(\hat{m} \leq k) - P(\hat{m} \leq k-1)]$$
$$= \sum_{k=0}^{M} k \cdot [F^{N-1}(k) - F^{N-1}(k-1)]$$
$$= M \cdot F^{N-1}(M) - F^{N-1}(M-1) - ... - F^{N-1}(0)$$
$$= M - F^{N-1}(M-1) - ... - F^{N-1}(0), \qquad (16)$$

Besides the basic properties of common CDFs, we need to pay particular attention to the following properties of $F(x)$: First, $F(x)$ is monotonically increasing: $\forall x \in \{1, 2, ..., M\}$, we have $F(x+1) \geq F(x)$; second, $F(M) = 1$; third, $F(0) < 1$, since $F(0) = 1$ means that the interferer will definitely not impact any AP which is not true in practice. Then we can find a constant $a \in \{0, 1, 2, ..., M-1\}$, such that $F(a+1) = 1$ and $F(a) < 1$; therefore,

$$m \geq E(\widetilde{m}) = M - F^{N-1}(M-1) - ... - F^{N-1}(0)$$
$$\geq M - 1 \cdot (M - a - 1) - (a+1) \cdot F^{N-1}(a)$$
$$= (a+1)(1 - F^{N-1}(a))$$
$$\triangleq (a+1)(1 - b^{N-1}), \qquad (17)$$

where $a = \sup_{F(x)<1} x$, $b = F(a)$ are both determined by the aggregated radio propagation environment and independent of N and m. ∎

Note that the concrete form of the bounds for the total number of impacted APs is highly dependent on the aggregated radio propagation environment, which is reflected by the parameters a, b and c; however, the values of the three parameters are just dependent on $\hat{m}_{\vec{r}}(\vec{r^*})$, which can be

obtained by merely investigating how one interferer impacts the system, instead of examining all possible scenarios with $N-1$ interferers.

B. Determine the Environment Dependent Parameters

We here present procedures to obtain values of the three parameters so that our theory can be verified in practice. The basic idea is to first construct $\hat{m}_{\vec{r}}(\vec{r^*})$ as a discrete function through practical measurements, and then derive values of the three parameters.

There are many possible values of $\vec{r}$, but we only select some special locations for measurement in practice. This is based on an important observation: the longer distance between the T-R pair, the larger interference region of the T-R pair will be. This is intuitive since the closer the transmitter is to the receiver, the less likely the AP will get interfered by others, and a larger interference region means that the corresponding AP is more likely to be impacted; therefore, we only need to study the locations that are the nearest to and farthest from all APs. The resulted values of $\hat{m}_{\vec{r}}(\vec{r^*})$ for such locations will lead to extreme values of m, which can be used for deriving the scalability bounds of the fingerprinting localization system.

We place k mobile devices at those special locations $\vec{r_1}, \vec{r_2}, ..., \vec{r_k}$, and let an interferer stand in different reference points. Then the mobile devices can measure the observed RSS with respect to different APs as the location of the interferer $\vec{r^*}$ changes. Given a value of $\vec{r^*}$, each mobile device can compare the observed RSS of each AP with that stored in the training phase, based on which the number of impacted APs can be recorded. In this manner, the $\hat{m}_{\vec{r}}(\vec{r^*})$ in the form of a discrete function can be obtained. Then we have

$$a = \max_{\vec{r^*}} \hat{m}_{\vec{r_j}}(\vec{r^*}) - 1, \quad b = \frac{||\{\vec{r^*}|\hat{m}_{\vec{r_j}}(\vec{r^*})=a, \vec{r^*} \in S\}||}{||\{\vec{r^*}|\hat{m}_{\vec{r_j}}(\vec{r^*} \in S\}||} \quad \text{and}$$

$c = E(\hat{m}_{\vec{r_i}}(\vec{r^*}))$, according to the analysis above.

V. MAIN RESULTS

Theorem 3. *For the indoor localization system with M APs distributed along boundaries of the region S, which is designed to support N users, the localization reliability of the system R' satisfies that*

$$e^{-\Theta(\phi(N))}(1 - e^{-\Theta(\phi(N))}) \leq R' \leq 1 - e^{-\Theta(\varphi(N))}, \quad (18)$$

where

$$\begin{cases} \phi(N) = \frac{M^2}{(1-A^2)\min\{c(N-1), M\} + A^2 M}, \\ \varphi(N) = \frac{M^2}{(1-A^2)(a+1)(1-b^{N-1}) + A^2 M}, \end{cases} \quad (19)$$

a, b, c can be obtained by the procedure described above and A is the ratio of variances mentioned in Theorem 1.

Proof: We can obtain Theorem 3 by combining Theorem 1 and Theorem 2. With Theorem 1, we have

$$e^{-\Theta(g(m))}(1 - e^{-\Theta(g(m))}) \leq R'(E) \leq 1 - e^{-\Theta(g(m))}, \quad (20)$$

where $g(m) = \frac{M^2}{(1-A^2)\cdot m + A^2 \cdot M}$ and $A < 1$;

Note that the RHS of the inequality above is a monotonically decreasing function with respect to m, then we must choose m's lower bound expression about N to obtain the final upper bound of $R'(E)$ with respect to N. A careful

examination of the LHS of the inequality reveals that the LHS can be either monotonically decreasing or first increasing and then decreasing with respect to m, if we do not consider the physical meaning of the inequality. However, since the reliability $R'(E)$ will definitely decrease with the number of impacted APs m increases, the LHS must be monotonically decreasing with m; therefore, we have to choose m's upper bound expression about N to obtain the final lower bound of $R'(E)$ with respect to N. According to Theorem 2,we have: $(a+1)(1-b^{N-1}) \leq m \leq \min\{c(N-1), M\}$.

Consequently,

$$e^{-\Theta(\phi(N))}(1 - e^{-\Theta(\phi(N))}) \leq R' \leq 1 - e^{-\Theta(\varphi(N))}, \quad (21)$$

where

$$\begin{cases} \phi(N) = \frac{M^2}{(1-A^2)\min\{c(N-1),M\}+A^2M}, \\ \varphi(N) = \frac{M^2}{(1-A^2)(a+1)(1-b^{N-1})+A^2M}, \end{cases} \quad (22)$$

and a, b, c can be obtained by the procedures described in the previous section and A is the ratio of variances mentioned in Theorem 1.

VI. EVALUATIONS OF MAIN RESULTS

A. Numerical Results

We here verify the main results by examining the trends of both the upper and the lower bound of R' as N increases. To obtain the numerical results of the bounds, we first need to obtain the environment dependent parameters a, b and c following the procedures presented in Section IV-B.

We set up the testbed in a square $100m^2$ gymnastics room, which is gridded into $1m \times 1m$ cells. We uniformly deploy 12 mobile Wi-Fi APs along the edges of the room. This is to create a comparatively ideal environment to eliminate unexpected interference from unnecessary details. We choose two special locations to derive the discrete function $\hat{m}_{\vec{r_i}}(\vec{r^*})$: the geometric center and a corner of the room denoted by $\vec{r_1}$ and $\vec{r_2}$ respectively; because it is easy to mathematically prove the center is the nearest point to all APs' and the corner is the farthest. Figure Fig. 5(a) and Fig. 5(b) show the results when placing the receiving mobile devices in the corner and at the center of the room respectively, where the different colors mean the total numbers of impacted APs observed when the interferer is at different RPs in the room. Based on such results, we have $E(\hat{m}_{\vec{r_1}}(\vec{r^*})) = 5.76$ and $E(\hat{m}_{\vec{r_2}}(\vec{r^*})) = 6.92$, thus $a = 8, b = 0.96, c = 6.92$.

With the environment dependent parameters, we could profile the trend of upper and lower bound. Since the expression with respect to $\Theta(\cdot)$ just characterizes the shape of the bounds, we must determine values of k_1^* and k_2^* of $\Theta(\cdot)$ before obtaining concrete bounding curves. We take the equations from both sides of the inequalities of Theorem 3. For the upper bound, we let $R' = 1 - e^{-k_1^* \varphi(N)}$ equal the highest reliability observed when $N = 1$ in the offline phase, then we could find the value of k_1^*; for the lower bound, we let $R' = -k_1^* \phi(N)(1 - e^{-k_2^* \phi(N)})$ equal the lowest reliability when $N = 1, 2$ in experiments of obtaining a, b and c, then we could find the values of k_1^* and k_2^* as $\phi(N)$ can be obtained once the value of N is fixed. Figure 5(c) and Fig. 5(d) show the numerical results of the upper and lower bound

of R'. Figure 5(e) shows the total number of impacted APs as N increases. We can see that as the number of impacted APs increases, the localization reliability decreases, which is realistic.

B. Experimental Results

We now present the experimental results for localization, which is compared with the derived theoretical bounds. We first traverse each cell of the area and record the corresponding RSSs of each AP to construct the radio map. We next let one normal user and one interferer enter into the area walking around, then we let them stop and record the estimated and real locations of the user in the presence of the interferer. We repeat the experiments 100 times under 8 kinds of user distribution, and then we can find the reliability of location estimation in the 1-interferer scenario. In this way, we increase the number of interferers one by one to 10, and then we can find the reliability of location estimation error in different scenarios. To calculate the reliability, we need to specify the tolerance error δ; we consider one-time localization is successful if the estimated location is in the δ neighborhood of the real location. In our experiments, we calculate the reliability when $\delta = 2m$ and $\delta = 3m$.

Figure 5(f) shows the experimental results of localization reliability as N increases under different accuracy standards. Figure 5(g) and Figure 5(h) show the comparative positions of the theoretical bounds curves and the experiments curves. We can see that the experimental results are bounded by our theoretical bounds, which validates the main results.

We can see the first-increase-then-decrease part of the reliability curve in those figures. This is primarily caused by the experimental error. The theoretical reliability is in fact the expectation of the corresponding localization reliability of all possible user-distribution scenarios; however, our experiments only record several scenarios. The curve will be more smooth if more experiments can be conducted. We also note that the resulted reliability from the experiments almost achieves the lower bound when $N = 11$, because all the 12 APs have been impacted in this case according to Fig. 5(e). We can expect that reliability curve will be flat if more interferers were added.

C. Important Observations and Analysis

Observation 1: *Localization reliability drops dramatically before the number of users increases to a critical point and then decreases smoothly, where the critical point tends to appear when the number of users equals that of APs deployed in the region.* With Fig. 5(c) and Fig. 5(d) showing the trend of upper and lower bounds of R', we can see that R' decreases as the number of users N increases, because more people in S means there are users in the interference region between the AP and the normal user with a higher probability; however, it is notable that the overall trends of R' first drop dramatically as N increases to a critical point, and then decrease smoothly. This observation can be explained as following. When the number of users in the system is small, adding a new-coming user can dramatically impact the radio propagation environment; however, if there have been a number of users in the system, it is likely that most of the

(a) Num. impacted APs when the interferer in different positions

(b) Num. impacted APs when the interferer in different positions

(c) Numerical results with initial reliability at 90%

(d) Numerical results with initial reliability at 50%

(e) Num. impacted APs

(f) R' w.r.t N

(g) Experimental and numerical results, $\delta = 3$

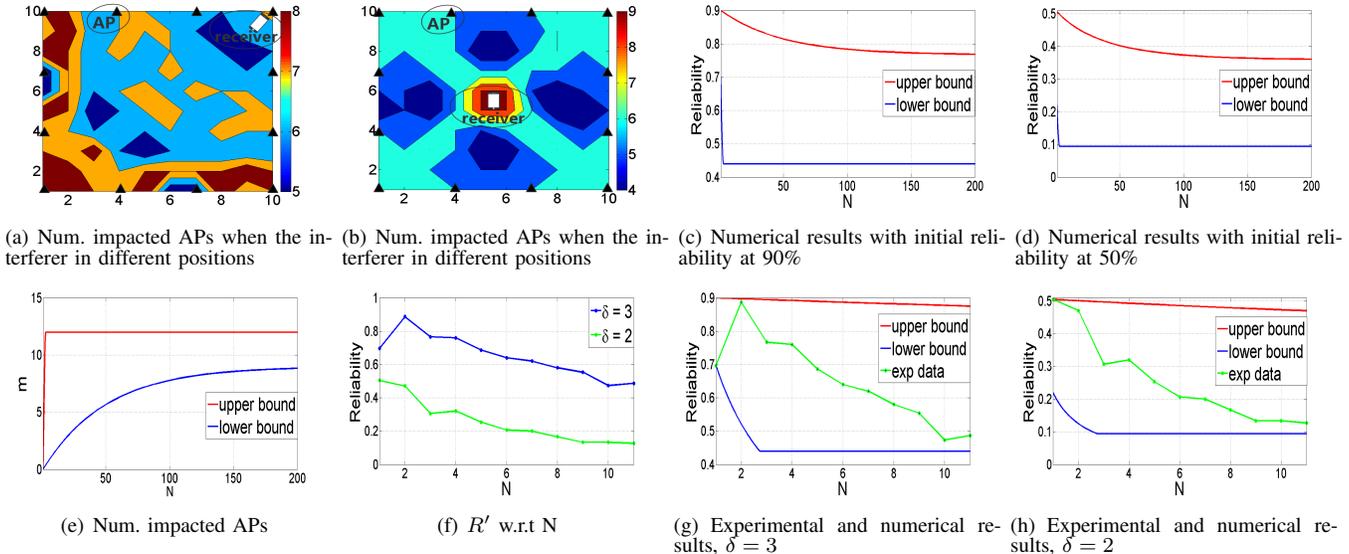(h) Experimental and numerical results, $\delta = 2$

Fig. 5. Experimental and numerical results

APs have been impacted, thus adding a new-coming user will make the change of the radio propagation environment not that dramatic.

However, it is difficult to find the exact value of the critical point. Observing the trend shown in Fig. 5(c) and Fig. 5(d), the values of the critical point for the upper and lower bound are different. This is reasonable since the upper bound indicates the best case, where more interferers can be tolerated, but the lower bound indicates the worse case, where it is intolerant to interferers. This is why the critical point for the lower bound appears much earlier than that for the upper bound. The experimental results as shown in Fig. 5(g) and Fig. 5(h) provide some hint for finding the critical point: it seems that the value of the critical point is around the value of M, the number of APs deployed in the region. This also makes sense to some extent because all the APs could be impacted if $N = M$, with each interferer impacting an AP.

Observation 2: *Even if the number of users $N \to \infty$, the fingerprinting localization system still retains certain level of reliability.* Our experimental results corroborate our theoretical results; however, what will be the resulted reliability if we keep increasing the number of users N? It is difficult to experiment the scenario of $N \to \infty$, but Theorem 3 has shed some light to the answer.

With Eq. (18)-(19), we have $R' \geq e^{-k_1^* \phi(N)}(1 - e^{-k_2^* \phi(N)})$, where $\phi(N) = \frac{M^2}{(1-A^2)\min\{c(N-1),M\}+A^2 M}$ and the parameters c, k_1^*, k_2^* are all independent of the number of users N in the system, but could be related to the number of all APs M. When $N \to \infty$, $\phi(N) = \frac{M^2}{(1-A^2)M+A^2 M} = M$, thus $R' \geq e^{-pM}(1 - e^{-qM}) > 0$, meaning the localization system will still retain certain level of reliability even if there are infinite number of users.

We now consider the extreme-case lower bound in Theorem 3, which is $e^{-k_1^* M}(1 - e^{-k_2^* M})$ occurring when $N \to \infty$. Mathematically, the trend of the worst-case lower bound can be monotonically increasing or decreasing-first-then-increasing; however, the only possible trend in practice should

be monotonically increasing. This can be obtained by using the similar approach of proving Theorem 3. Consequently, an interesting byproduct of the observation is that: deploying more APs can improve the extreme-case lower bound.

VII. RELATED WORK

Theoretical work has been done to evaluate the performance of fingerprinting based indoor localization systems; however, the main focus is on the accuracy and reliability. Cramér-Rao Bound (CRB) analysis can provide a lower bound on the variance achievable by any unbiased estimator, which is utilized to evaluate accuracy of the fingerprinting localization [22]. The derived performance bound is compared with experimental results; however, the comparison indicates that the derived CRB is inaccurate [23], which is due to the inaccurate modeling of radio propagation.

Wen *et al.* propose a general probabilistic model to evaluate the accuracy and reliability of the fingerprinting based localization systems [11], where the probability of correct location estimation can be obtained given the requirement of accuracy. The radio propagation model in [11] is a generalized log-normal path loss (LNPL) model containing the effects of major radio propagation characteristics, based on which the location estimation process is modeled as a mapping from the fingerprints sample space to the physical space. In contrast to [11] assuming the observable fingerprints are static, our work in this paper takes the deviation of observable fingerprints in the online phase into account, which is introduced by presence of users.

Tian *et al.* evaluate how the imperfect fingerprints database will impact the reliability and accuracy of the fingerprinting based localization systems [12], where the imperfect fingerprints database results from the adoption of unreliable crowdsourcing paradigm in the offline phase. The work in [12] presents a very preliminary analysis on how the deviation of the mean will impact the localization performance only for the 1-D localization case. In contrast to [12], our work in this

paper is based on a revised reliability model, and we present complete and rigorous analysis on the high-dimensional scenario in both sample and physical space. Our work in this paper initiates the study of scalability on the fingerprinting based localization.

VIII. Conclusions

In this paper, we have studied the scalability issue of the fingerprinting localization system in a theoretical perspective, where the upper and lower bound of the system's localization reliability with respect to the number of users are derived. The theoretical and experimental results of this work have revealed two interesting observations that shed light on the insight into the scalability of the fingerprinting localization system: First, the localization reliability drops dramatically before the number of users increases to a critical point and then decreases smoothly, where the critical point tends to appear when the number of users equals the number of access points (APs) deployed in the service region; second, even if the number of users approaches to infinity, the fingerprinting localization system still retains certain level of reliability.

IX. Acknowledgement

References

[1] Z. Yang, Z. Zhou and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Comput. Surv.*, vol. 46, no. 2, pp.1–32, 2013.

[2] S. He and S. H. G. Chan, "Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons," in *IEEE Communications Surveys & Tutorials*, 2016, pp. 466-490.

[3] P. Bahl and V. N. Padmanabhan, "Radar: An in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, 2000, pp. 775–784.

[4] A. Haeberlen, E. Flannery, A. M. Ladd, A. Rudys, D. S. Wallach and L. E. Kavraki, "Practical robust localization over large-scale 802.11 wireless networks," in *Proc. ACM MobiCom*, 2004, pp. 70–84.

[5] M. Youssef and A. Agrawala, "The horus wlan location determination system," in *Proc. ACM MobiSys*, 2005, pp. 205–218.

[6] K. Chintalapudi, A. Padmanabha Iyer and V. N. Padmanabhan, "Indoor localization without the pain," in *Proc. ACM MobiCom*, 2010, pp. 173–184.

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

[8] H. Liu, Y. Gan, J. Yang, S. Sidhom, Y. Wang, Y. Chen and F. Ye, "Push the Limit of WiFi based Localization for Smartphones," in *Proc. ACM MobiCom*, 2012, pp. 305–316.

[9] H. Liu, J. Yang, S. Sidhom, Y. Wang, Y. Chen and F. Ye, "Accurate WiFi Based Localization for Smartphones Using Peer Assistance," *IEEE Transactions on Mobile Computing* , vol. 13, no. 10, pp.2199–2214, Oct. 2013.

[10] C. Wu, Z. Yang and Y. Liu, " Smartphones based Crowdsourcing for Indoor Localization," *IEEE Transactions on Mobile Computing* , vol. 13, no. 10, pp.2199–2214, Oct. 2013.

[11] Y. Wen, X. Tian, X. Wang and S. Lu, "Fundamental limits of RSS fingerprinting based indoor localization," in *Proc. IEEE INFOCOM*, 2015.

[12] X. Tian,R. Shen,D. Liu,Y. Wen and X. Wang, "Performance Analysis of RSS Fingerprinting Based Indoor Localization," *IEEE Transactions on Mobile Computing*, 2016. Online:http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7797481.

[13] Z. Yang, C. Wu, Z. Zhou, X. Zhang, Y. Wang and Y. Liu, "Mobility Increases Localizability: A Survey on Wireless Indoor Localization using Inertial Sensors," *ACM Comput. Surv.*, vol. 47, no. 3, Article 54, Apr. 2015.

[14] M. Kotaru, K. Joshi, D. Bharadia and S. Katti, "SpotFi: Decimeter Level Localization Using WiFi," in *Proc. ACM SIGCOMM*, 2015, pp. 269-282.

[15] S. Kumar, S. Kumar S and D. Katabi, "Decimeter-level localization with a single WiFi access point," in *Proc. USENIX NSDI*, 2016, pp. 165-178.

[16] L. Yang, Y. Chen, X. Li, C. Xiao, M. Li and Y. Liu, "Tagoram: Real-Time Tracking of Mobile RFID Tags to High Precision Using COTS Devices," in *Proc. ACM MobiCom*, 2014.

[17] K. Liu, X. Liu and X. Li, "Guoguo: Enabling Fine-Grained Smartphone Localization via Acoustic Anchors," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp.1144–1156, May 2016.

[18] KAILOS, https://kailos.io/.

[19] G. Shen, Z. Chen, P. Zhang, T. Moscibroda and Y. Zhang, "Walkie-Markie: Indoor pathway mapping made easy," in *Proc. USENIX NSDI*, 2013, pp. 85–98.

[20] C. Luo, H. Hong and M. C. Chan, "PiLoc: a self-calibrating participatory indoor localization system," in *Proc. IEEE IPSN*, 2014, pp. 143–153.

[21] E. Elnahrawy, X. Li and R. P. Martin, "The limits of localization using signal strength: A comparative study," in *Proc. IEEE SECON*, 2004, pp. 406–414.

[22] G. Chandrasekaran, M. A. Ergin, J. Yang, S Liu, Y. Chen, M. Gruteser and R. P. Martin, " Empirical evaluation of the limits on localization using signal strength," in *Proc. IEEE SECON*, 2009.

[23] A. M. Hossain and W. S. Soh, "Cramer-rao bound analysis of localization using signal strength difference as location fingerprint," in *Proc. IEEE INFOCOM*, 2010.

[24] T. V. Haute, E. D. Pooter, F. Lemic, V. Handziski, N. Wirström, T. Voigt, A. Wolisz and I. Moerman, "Platform for benchmarking of RF-based indoor localization solutions," *IEEE Communication Magazine*, vol. 53, no. 9, pp. 126–133, 2015.

[25] T. V. Haute, E. D. Pooter, I. Moerman, F. Lemic, V. Handziski, A. Wolisz, N. Wirström and T. Voigt "Comparability of RF-based Indoor Localization Solutions in Heterogeneous Environments: An Experimental Study," *Int. J. Ad Hoc and Ubiquitous Computing*, vol. 23, nos. 1/2, pp. 92–114, 2016.

[26] S. Obayashi, J. Zander, "A body-shadowing model for indoor radio communication environments View Document", *IEEE Trans. Antennas Propag.*, vol. 46, no. 6, pp. 920-927, June 1998.

[27] M. Ayadi, A. Zineb, "Body Shadowing and Furniture Effects for Accuracy Improvement of Indoor Wave Propagation Models", *IEEE Trans. Wireless Commun.*, vol. 13, no. 11, pp. 5999-6006, Nov. 2014

[28] I. Kashiwagi, T. Taga, T. Imai, "Time-varying path-shadowing model for indoor populated environments view document", *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 16-28, Jan. 2010.

[29] H. Hashemi, "Impulse response modeling of indoor radio propagation channels," *IEEE J. Sel. Areas Commun.*, vol. 11, no. 7, pp. 967–978, Sept. 2006.

[30] M. Youssef and A. Agrawala, "Handling samples correlation in the Horus system, in *Proc. IEEE INFOCOM*, 2004.

[31] M. Youssef, M. Abdallah and A. Agrawala, "Multivariate analysis for probabilistic WLAN location determination systems, in *Proc. the Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, 2005.

[32] International Telecommunication Union(ITU), "Guidelines for evaluation of radio interface technologies for IMT-Advanced", in Report ITU-RM.2135-1, pp. 30–33, Dec.2009, Online:https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2135-1-2009-PDF-E.pdf.

[33] WINNER ,"WINNER Channel Models", IST-4-027756,pp.44–45,Sept.2007, Online:http://www.cept.org/files/1050/documents/winner2%20-%20final%20report.pdf.

[34] J. Wang, D. Fang, Z. Yang, H. Jiang, X. Chen, T. Xing and L. Cai, "E-HIPA:An Energy-Efficient Framework for High-Precision Multi-Target-Adaptive Device-free localization,"*IEEE Transactions on Mobile Computing*, 2017. Online:http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=7469386

[35] J. Wang, H. Jiang, J. Xiong, K. Jamieson, X. Chen, D. Fang and B. Xie, "LiFS: Low Human Effort, Device-Free Localization with Fine-Grained Subcarrier Information," in *Proc. ACM MobiCom*, 2016.

[36] Technical Report, online https://www.dropbox.com/s/u83yyj56mnhe6sh/Secon%202018-Capacity-TechRep.pdf?dl=0.