An Information-Based Approach to Precision Analysis of Indoor WLAN Localization Using Location Fingerprint

Mu Zhou 1, Feng Qiu 1,*, Zengshan Tian 1, Haibo Wu 2, Qiao Zhang 1 and Wei He 1

Received: 31 October 2015; Accepted: 30 November 2015; Published: 5 December 2015

Academic Editor: Raúl Alcaraz Martínez

1 Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; zhoumu@cqupt.edu.cn (M.Z.); tianzs@cqupt.edu.cn (Z.T.); zhangqiao6120@gmail.com (Q.Z.); hewei@cqupt.edu.cn (W.H.)
2 China Internet Research Lab, Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China; wuhaibo@cstnet.cn
* Correspondence: qiufeng6906@gmail.com

Abstract: In this paper, we proposed a novel information-based approach to precision analysis of indoor wireless local area network (WLAN) localization using location fingerprint. First of all, by using the Fisher information matrix (FIM), we derive the fundamental limit of WLAN fingerprint-based localization precision considering different signal distributions in characterizing the variation of received signal strengths (RSSs) in the target environment. After that, we explore the relationship between the localization precision and access point (AP) placement, which can provide valuable suggestions for the design of the highly-precise localization system. Second, we adopt the heuristic simulated annealing (SA) algorithm to optimize the AP locations for the sake of approaching the fundamental limit of localization precision. Finally, the extensive simulations and experiments are conducted in both regular line-of-sight (LOS) and irregular non-line-of-sight (NLOS) environments to demonstrate that the proposed approach can not only effectively improve the WLAN fingerprint-based localization precision, but also reduce the time overhead.

Keywords: WLAN localization; access point placement; location fingerprint; Fisher information; simulated annealing

1. Introduction

In the recent decade, there has been a growing interest in indoor localization techniques, which are based on existing indoor wireless communication infrastructures and devices. Due to the implementation ease and cost efficiency [1,2], the indoor wireless local area network (WLAN) fingerprint-based localization approach is preferred compared to the conventional trilateration localization approaches, which are easily compromised by the propagation path loss, multi-path fading and environmental shadowing. As far as we know, there are generally two phases involved in WLAN fingerprint-based localization [3–6], namely the off-line phase and on-line phase. In the off-line phase, the fingerprint database is built based on the received signal strength (RSS) measurements, which are associated with the calibrated reference points (RPs) in the target environment. Then, in the on-line phase, the target locations are estimated by the matching from the newly-recorded RSS measurements against the pre-built fingerprint database to estimate the target locations [7].

The multi-path fading, environmental shadowing and channel interference always result in significant temporal and spatial variations of RSS distributions and eventually lead to low localization
Entropy 2015, 17, 8031–8055

precision [8]. To the best of our knowledge, there are very few works focusing on the theoretical analysis of the WLAN fingerprint-based localization precision, but in fact, the fundamental limit of localization precision can help greatly in AP placement optimization to achieve the highly-precise localization purpose.

In this paper, we propose an analytical model to characterize the variation of RSSs under different signal distributions for the sake of deriving the closed-form fundamental limit of the WLAN fingerprint-based localization precision based on the Fisher information matrix (FIM) [9–13]. Furthermore, the impact of the number and geographical locations of APs and environment noise on localization precision is also discussed.

The rest of this paper is structured as follows. The related works are given in Section 2. In Section 3, we introduce the system in detail, including the calculation of the fundamental limit of localization precision and optimization of AP placement. Simulation results are provided in Section 4. In Section 5, some interesting discussions are presented. A case study is presented in Section 6. Finally, we conclude the paper in Section 7.

2. Related Work

In the past decade, the WLAN fingerprint-based localization precision has been studied extensively. The authors in [14] proposed that the number and locations of APs, the physical layout and the mean of RSS measurements at each RP have a significant impact on localization precision. To optimize AP placement, the existing approaches are generally based on the signal coverage, service connectivity, network throughput and transmission rate [15]. The authors in [16] invented a weighted kernel function by which the impact of different APs is differentiated. Specifically, larger weights are assigned to the APs that contribute more to the localization. The authors in [17] developed an approach to find the minimum number of APs that are used to provide the full coverage, as well as to conduct the localization with the errors falling into a given scope. Chen in [18] proposed an AP placement optimization approach by distinguishing every two WLAN RSS fingerprints to guarantee satisfactory signal coverage. The localized local discriminant embedding and AP selection (LLDE-APS) approach in [19] not only optimizes the AP placement, but also saves the power consumption of the mobile terminal. The approach in [19] is featured with the main advantage of extracting the most discriminative RSS features for target localization. However, the existing works in the literature paid very little attention to the theoretical relationship between the AP placement and localization precision. To solve this problem, we propose a novel AP placement optimization approach to improve the localization precision, as well as to preserve the real-time capacity of the system. In [20], the authors proposed the max Euclidean distance (ED)-based AP placement optimization approach, aiming to achieve the maximal sum of the ED between every two location fingerprints in a WLAN environment. However, the existing work from the literature paid very little attention to the theoretical and analytical relations of the AP placement and localization precision. To solve this problem, we proposed a novel AP placement optimization approach based on the relationship between the geographical locations of APs and localization errors with the purpose of enhancing the localization precision, as well as guaranteeing the real-time capacity of localization.

In wireless sensor networks (WSNs), there is a variety of works focusing on the error analysis by using the FIM. The authors in [21] presented the conditions in which the multipath delay can be used to improve the localization accuracy and meanwhile introduced the FIM to investigate the highest achievable localization accuracy. The authors in [22] also studied the relationship between the multipath delay and time of arrival (TOA)-based localization accuracy based on the prior statistics of the errors. The authors in [23] relied on the Cramer–Rao lower bound (CRLB) to estimate the locations of sensors by using the unbiased Gaussian range estimation from the anchor nodes in the angle of arrival (AOA)-based localization. In this paper, we utilize the FIM to derive the fundamental limit of WLAN fingerprint-based localization precision under different signal distributions. This result can provide valuable insights into the improvement of fingerprint-based localization precision, as well as
the overall design of the WLAN localization system. Furthermore, the simulated annealing algorithm (SA) [24,25] is used to search the optimal AP locations, which correspond to the lowest fundamental limit of localization precision. Although the FIM has been significantly considered in localization precision analysis, the existing literature mainly applied it to the situation that the TOA, AOA or RSS following the Gaussian distribution is selected as the location metric. However, in this paper, we aim to reveal the relations of the fundamental limit of localization precision and various signal distributions.

We clarify that the three main contributions of this paper are that: (i) the FIM is used to derive the fundamental limit of indoor WLAN localization precision; (ii) the theoretical analysis towards the relationship between the localization errors and signal distributions is presented; and (iii) the SA algorithm is selected to search the optimal AP locations, which correspond to the lowest fundamental limit of localization precision.

### 3. System Description

#### 3.1. System Overview

![Flow chart](image)

**Figure 1.** Flow chart of the proposed system.

As discussed in [26], we assume that the RSSs in the non-line-of-sight (NLOS) environment follow the Rayleigh distribution, while the ones in the line-of-sight (LOS) environment follow the
Gaussian distribution. We define that the $i$-th user is in the LOS environment for the $m$-th AP when there is no wall blocking for the rectilinear propagation from the $m$-th AP to the receiver. On the contrary, if there is wall blocking, it is defined that the user is in the NLOS environment for the $m$-th AP. Secondly, we rely on the characteristics of the FIM to derive the fundamental limit of localization precision under the Gaussian signal distribution, Rayleigh signal distribution and mixed signal distribution, respectively. Thirdly, we select all of the candidate AP locations to construct the solution space. Then, we construct the objective function of the SA algorithm for AP placement optimization with the purpose of achieving the lowest fundamental limit of localization precision. Finally, we search the optimal AP locations by the SA algorithm. The proposed system consists of two main modules: (i) the calculation of the fundamental limit of WLAN fingerprint-based localization precision; and (ii) the AP placement optimization by using the SA algorithm to achieve the lowest fundamental limit of localization precision. The flow chart of the system is shown in Figure 1.

3.2. Fundamental Limit of Localization Precision

3.2.1. Localization Precision vs. Signal Distributions

We select the COST231 model [27] to characterize the propagation property in an indoor WLAN environment. The COST231 model considers the large-scale path loss, as well as the signal penetration, like the wall attenuation factor (WAF) and floor attenuation factor (FAF), as described in Equation (1).

$$P = P(d_0) - 10 \beta \log_{10}(\frac{d}{d_0}) - wP_w - \chi$$

where $P$ and $P(d_0)$ stand for the RSSs recorded at the locations with $d$ and $d_0$ meters from the AP, respectively. $d_0$ is the reference distance, and $\beta$ is the path loss exponent. $wP_w$ is the signal fading by the walls with the number of $w$. $\chi$ is the noise with the variance of $\sigma^2$.

By assuming that $\hat{\theta}_i = (\hat{x}_i, \hat{y}_i)^T$ is the estimated location with respect to the $i$-th real location $\theta_i = (x_i, y_i)^T$, we can calculate the covariance matrix of $\hat{\theta}_i$ by:

$$\text{var} [\hat{\theta}_i] = E_{\theta} \left\{ \left( \hat{\theta}_i - \theta_i \right) \left( \hat{\theta}_i - \theta_i \right)^T \right\} = \begin{bmatrix} \sigma_{\xi_i}^2 & \sigma_{\xi_i \hat{y}_i} \\ \sigma_{\hat{y}_i \xi_i} & \sigma_{\hat{y}_i}^2 \end{bmatrix}$$

where $\sigma_{\xi_i}^2$ and $\sigma_{\hat{y}_i}^2$ are the mean square errors (MSEs) of the estimated location in the X and Y coordinates, $\sigma_{\xi_i \hat{y}_i}$ and $\sigma_{\hat{y}_i \xi_i}$ are the covariance between $\hat{x}_i$ and $\hat{y}_i$ and between $\hat{y}_i$ and $\hat{x}_i$, respectively.

Since the value $\hat{\theta}_i$ is calculated from the estimate of $\hat{P}$, the values $\hat{\theta}_i$ and $\hat{P}$ have the same variance $\sigma^2$ under the unbiased estimates of $P$ and $\theta$. Thus, the localization error will increase as the value $\sigma$ increases. By setting $f_{\theta_i}(P)$ as the probability density function (PDF) of $P$ with respect to $\theta$, the expectation of the sharpness of $f_{\theta_i}(P)$ equals $1/\sigma^2$, as shown in Equation (3).

$$E \left\{ -\frac{\partial^2 \ln f_{\theta_i}(P)}{\partial \theta_i^2} \right\} = \frac{1}{\sigma^2}$$

Then, we have:

$$\text{var} [\hat{P}] = \text{var} [\hat{\theta}_i] = \sigma^2 = \frac{1}{E \left\{ -\frac{\partial^2 \ln f_{\theta_i}(P)}{\partial \theta_i^2} \right\}}$$

For the biased estimate of $\hat{\theta}_i$, we easily obtain:

$$\text{var} [\hat{\theta}_i] \geq \frac{1}{E \left\{ -\frac{\partial^2 \ln f_{\theta_i}(P)}{\partial \theta_i^2} \right\}}$$
By setting $J ( \theta_i ) = - \mathbb{E} \left[ \frac{\partial^2 \ln f_{\theta_i} (P)}{\partial \theta_i^2} \right]$ as the FIM of $\theta_i$, we have:

$$\operatorname{var} \left[ \hat{\theta}_i \right] \geq J ( \theta_i )^{-1} \quad (6)$$

Then, we calculate:

$$J ( \theta_i ) = - \mathbb{E} \left[ \frac{\partial^2 \ln f_{\theta_i} (P)}{\partial \theta_i^2} \right] = - \mathbb{E} \left( \int \frac{\partial^2}{\partial \theta_i^2} f_{\theta_i} (P) \, dP - \left[ \frac{\partial \ln f_{\theta_i} (P)}{\partial \theta_i} \right]^2 \right) = - \mathbb{E} \left( \frac{\partial^2}{\partial \theta_i^2} \int f_{\theta_i} (P) \, dP \right) + \mathbb{E} \left[ \frac{\partial \ln f_{\theta_i} (P)}{\partial \theta_i} \right]^2 = E \left[ \frac{\partial \ln f_{\theta_i} (P)}{\partial \theta_i} \right]^2 \quad (7)$$

As discussed in [28, 29], the score function $U ( \theta_i )$ is defined as:

$$U ( \theta_i ) = \nabla \ln f_{\theta_i} (P) = \frac{\partial}{\partial \theta_i} \ln f_{\theta_i} (P) \quad (8)$$

Therefore, based on:

$$E [U ( \theta_i )] = E \left[ \frac{\partial}{\partial \theta_i} \ln f_{\theta_i} (P) \right] = E \left[ \frac{1}{f_{\theta_i} (P)} \cdot \frac{\partial}{\partial \theta_i} (f_{\theta_i} (P)) \right] = \int \frac{\partial}{\partial \theta} (f_{\theta_i} (P)) \cdot \frac{1}{f_{\theta_i} (P)} \cdot f_{\theta_i} (P) \cdot dP = \frac{\partial}{\partial \theta} \int f_{\theta_i} (P) \cdot dP = 0 \quad (9)$$

we have $\operatorname{var}[U ( \theta_i )] = E \left[ \frac{\partial \ln f_{\theta_i} (P)}{\partial \theta} \right]^2 = J ( \theta_i )$.

Analysis with Gaussian signal distribution

As the RSSs follow the Gaussian signal distribution, the joint PDF of RSSs is calculated by:

$$f_{\theta_i} (P) = \prod_{k=1}^{m} \frac{1}{\sqrt{2\pi \sigma_1}} \exp \left( - \frac{\xi_i^2}{2\sigma_1^2} \right) \quad (10)$$

where $\xi_i = P - P (d_0) + 10\beta \log (\frac{d_k}{d_0}) + P_{\text{off}}, \sigma_1^2$ is the deviation of the RSSs collected by the receiver, $d_{ik} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}$, $(x_k, y_k)$ represents the coordinate of the $k$-th AP and $m$ is the AP number. We convert Equation (10) into:

$$f_{\theta_i} (P) = \prod_{k=1}^{m} \frac{1}{\sqrt{2\pi \sigma_1}} \exp \left\{ - \frac{\rho_{1i}^2 (\ln d_{ik})^2}{2} \right\} \quad (11)$$

where $\rho_{1i} = \left( \frac{10\beta}{\sigma_1 \ln 10} \right)^2$. At this point, the geometric relationship between the $i$-th real location and the $k$-th AP is shown in Figure 2.
Based on Equation (7), we have:

\[
J(\theta_i) = \begin{bmatrix}
J_{xx}(\theta_i) & J_{xy}(\theta_i) \\
J_{yx}(\theta_i) & J_{yy}(\theta_i)
\end{bmatrix}
\] (12)

where \(J_{xx}(\theta_i) = -\frac{\partial^2}{\partial x^2}\ln f_{\theta_i}(P), \ J_{xy}(\theta_i) = -\frac{\partial^2}{\partial x \partial y}\ln f_{\theta_i}(P), \ J_{yx}(\theta_i) = -\frac{\partial^2}{\partial y \partial x}\ln f_{\theta_i}(P)\) and \(J_{yy}(\theta_i) = -\frac{\partial^2}{\partial y^2}\ln f_{\theta_i}(P)\). Based on this, we calculate that:

\[
\frac{\partial^2}{\partial x^2}\ln f_{\theta_i}(P) = -\rho_1 \sum_{k=1}^{m} \left\{ \ln d_{ik} \times \frac{\partial}{\partial x_i} \left( \frac{x_i - x_k}{d_{ik}} \right) + \frac{(x_i - x_k)^2}{d_{ik}^2} \right\}
\] (13)

\[
\frac{\partial^2}{\partial x \partial y}\ln f_{\theta_i}(P) = -\rho_1 \sum_{k=1}^{m} \left\{ \ln d_{ik} \times \frac{\partial}{\partial y_i} \left( \frac{x_i - x_k}{d_{ik}} \right) + \frac{(x_i - x_k)(y_i - y_k)}{d_{ik}^2} \right\}
\] (14)

\[
\frac{\partial^2}{\partial y^2}\ln f_{\theta_i}(P) = -\rho_1 \sum_{k=1}^{m} \left\{ \ln d_{ik} \times \frac{\partial}{\partial y_i} \left( \frac{y_i - y_k}{d_{ik}} \right) + \frac{(y_i - y_k)^2}{d_{ik}^2} \right\}
\] (15)

Based on Equations (7) and (9), we can simplify Equations (13–16) into:

\[
\begin{align*}
J_{xx}(\theta_i) &= \rho_1 \sum_{k=1}^{m} \left[ \frac{\cos \alpha_{ik}}{d_{ik}} \right]^2 \\
J_{xy}(\theta_i) &= J_{yx}(\theta_i) \quad \rho_1 \sum_{k=1}^{m} \left[ \frac{\sin \alpha_{ik}}{d_{ik}} \cos \alpha_{ik} \right] \\
J_{yy}(\theta_i) &= \rho_1 \sum_{k=1}^{m} \left[ \frac{\sin \alpha_{ik}}{d_{ik}} \right]^2
\end{align*}
\] (17)

Therefore, we can calculate \(\{ J(\theta_i) \}^{-1}\) by:

\[
\{ J(\theta_i) \}^{-1} = \frac{1}{|J(\theta_i)|} \begin{bmatrix} J_{yy}(\theta_i) & -J_{yx}(\theta_i) \\
-J_{xy}(\theta_i) & J_{xx}(\theta_i)\end{bmatrix}
\] (18)

where \(|J(\theta_i)| = J_{xx}(\theta_i)J_{yy}(\theta_i) - J_{xy}(\theta_i)^2\). Based on Equation (6), we have:

\[
\sigma_{\hat{x}_i}^2 = E(\hat{x}_i - x_i)^2 \geq J_{xx}^{-1}, \sigma_{\hat{y}_i}^2 = E(\hat{y}_i - y_i)^2 \geq J_{yy}^{-1}
\] (19)
Finally, the fundamental limit of localization precision with respect to $\theta_i$ equals:

$$V_m(\theta_i) = J_{xx}^{-1} + J_{yy}^{-1} = \frac{J_{xx} + J_{yy}}{J_{xx} \cdot J_{yy} - J_{xy}^2} = \frac{\frac{1}{\rho^2} \cdot \sum_{k=1}^{m} \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2 + \sum_{k=1}^{m} \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2}{\sum_{k=1}^{m} s \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2 + \sum_{k=1}^{m} s \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2}^2$$ (20)

Analysis with Rayleigh signal distribution

As the RSSs follow the Rayleigh signal distribution, the joint PDF of RSSs is calculated by:

$$f_{\theta_i}(P) = \prod_{k=1}^{m} \left( \frac{x_i}{\sigma^2} \right) \times \exp \left( -\frac{x_i^2}{2\sigma^2} \right)$$ (21)

where $\frac{4-\pi}{8} \sigma^2$ is the deviation of RSSs. We convert Equation (18) into:

$$f_{\theta_i}(P) = \prod_{k=1}^{m} \left( \frac{\rho_2 \ln(d_{ik})}{10\beta \ln 10} \right) \times \exp \left\{ -\frac{\alpha_1}{2} (\ln d_{ik})^2 \right\}$$ (22)

where $\rho_2 = \left( \frac{10\beta}{\sigma^2 \ln 10} \right)^2$. Similar to the previous discussion under the Gaussian signal distribution, we have:

$$\frac{\partial^2}{\partial x_i^2} \ln f_{\theta_i}(P) = \sum_{k=1}^{m} \left\{ (\rho_2 \ln d_{ik} - \frac{1}{\ln^2 d_{ik}}) \times \frac{\partial^2}{\partial x_i^2} \left( \frac{x_i - x_k}{d_{ik}^2} \right) - (\rho_2 + \frac{1}{\ln^2 d_{ik}}) \times \left( \frac{x_i - x_k}{d_{ik}^2} \right)^2 \right\}$$ (23)

$$\frac{\partial^2}{\partial y_i \partial x_i} \ln f_{\theta_i}(P) = \sum_{k=1}^{m} \left\{ (\rho_2 \ln d_{ik} - \frac{1}{\ln^2 d_{ik}}) \times \frac{\partial}{\partial y_i} \left( \frac{x_i - x_k}{d_{ik}^2} \right) - (\rho_2 + \frac{1}{\ln^2 d_{ik}}) \times \left( \frac{x_i - x_k}{d_{ik}^2} \right)^2 \right\}$$ (24)

$$\frac{\partial^2}{\partial y_i^2} \ln f_{\theta_i}(P) = \sum_{k=1}^{m} \left\{ (\rho_2 \ln d_{ik} - \frac{1}{\ln^2 d_{ik}}) \times \frac{\partial^2}{\partial y_i^2} \left( \frac{y_i - y_k}{d_{ik}^2} \right) - (\rho_2 + \frac{1}{\ln^2 d_{ik}}) \times \left( \frac{y_i - y_k}{d_{ik}^2} \right)^2 \right\}$$ (25)

Based on Equations (7) and (9), we obtain:

$$\begin{align*}
J_{xx} & = \sum_{k=1}^{m} \left( \rho_2 + \frac{1}{\ln^2 d_{ik}} \right) \times \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2 \\
J_{xy} & = \sum_{k=1}^{m} \left( \rho_2 + \frac{1}{\ln^2 d_{ik}} \right) \times \frac{\sin \frac{\alpha_{ik}}{d_{ik}} \cos \frac{\alpha_{ik}}{d_{ik}}}{\sin \frac{\alpha_{ik}}{d_{ik}}} \\
J_{yy} & = \sum_{k=1}^{m} \left( \rho_2 + \frac{1}{\ln^2 d_{ik}} \right) \times \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2
\end{align*}$$ (27)

Finally, the fundamental limit of localization precision with respect to $\theta_i$ is:

$$V_m(\theta_i) = J_{xx}^{-1} + J_{yy}^{-1} = \frac{J_{xx} + J_{yy}}{J_{xx} \cdot J_{yy} - J_{xy}^2} = \frac{\sum_{k=1}^{m} s \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2 + \sum_{k=1}^{m} s \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2}{\sum_{k=1}^{m} s \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2 \times \sum_{k=1}^{m} s \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2 - \sum_{k=1}^{m} s \times \left[ \sin \frac{\alpha_{ik}}{d_{ik}} \right]^2 \left[ \cos \frac{\alpha_{ik}}{d_{ik}} \right]^2}^2$$ (28)

where $s = \rho_2 + \frac{1}{\ln^2 d_{ik}}$.  

8037
For simplicity, we define the average of the fundamental limit of localization precision for the target environment as:

\[
V_{\text{ave}}(\theta) = \frac{1}{n} \sum_{i=1}^{n} V_m(\theta_i)
\]  

(29)

where \(n\) is the number of RPs.

### 3.2.2. Impact of the AP Number

Based on Equations (20) and (28), the relationship between the fundamental limit of localization precision with \(m\) and \(m + 1\) APs respectively is described as:

\[
V_m(\theta_i) - V_{m+1}(\theta_i) = \sqrt{\eta_{1m} \sum_{k=1}^{m} \left( \frac{\cos a_{ik}}{d_k} \right)^2 - \omega_{1m}^2} - \sqrt{\eta_{1m+1} \sum_{k=1}^{m+1} \left( \frac{\cos a_{ik}}{d_k} \right)^2 - \omega_{1m+1}^2}
\]

(30)

\[
\eta_{1m} = \sum_{k=1}^{m} \left( \frac{\cos a_{ik}}{d_k} \right)^2 + \sum_{k=1}^{m} \left( \frac{\sin a_{ik}}{d_k} \right)^2
\]

\[
\omega_{1m}^2 = \sum_{k=1}^{m} \left( \frac{\cos a_{ik}}{d_k} \right)^2 \times \sum_{k=1}^{m} \left( \frac{\sin a_{ik}}{d_k} \right)^2 - \sum_{k=1}^{m} \left( \frac{c \sin a_{ik} \cos a_{ik}}{d_k} \right)^2
\]

(31)

c = \rho_1 and s for the Gaussian and Rayleigh signal distributions, respectively. Based on Equations (30) and (31), we observe that the increase of the AP number reduces the fundamental limit of localization precision. Furthermore, the relationship \(V_m(\theta_i) = V_{m+1}(\theta_i)\) holds as the AP locations are collinear (i.e., \(a_{i(m+1)} = a_{ik}, k \in \{1, \ldots, m\}\)). Figures 3 and 4 show the cumulative distribution functions (CDFs) of errors with different AP numbers in the LOS and NLOS environments, respectively. As can be seen from these figures, the variation of the AP number has a slight impact on the localization errors when the AP number is larger than three. Furthermore, compared to the Rayleigh signal distributions, the Gaussian signal distribution generally results in higher localization precision.

![CDFs of errors with different AP numbers in the LOS environment. (a) Under the Gaussian signal distribution; (b) under the Rayleigh signal distribution.](image)
3.2.3. Impact of Noise Variance

Based on Equations (20) and (28), we can observe that the variation of $\sigma$ has a significant impact on the fundamental limit of localization precision. To illustrate this observation more clearly, we present the CDFs of errors with different variances of noise in the LOS and NLOS environments, respectively, as shown in Figures 5 and 6. From these figures, the increase of noise variance ($i.e., \sigma^2$) results in the decrease of localization precision, as expected.
3.2.4. Fundamental Limit with a Mixed Signal Distribution

In the actual indoor environment, the RSSs from different APs cannot be guaranteed to follow a unique signal distribution. Therefore, without losing generality, we assume that the RSSs from the former \( m_1 \) APs follow the Gaussian signal distribution, while the RSSs from the latter \( m_2 \) APs follow the Rayleigh signal distribution. Based on this, we have:

\[
\begin{align*}
    f_{\theta_i}(P) &= f_{\theta_i}(P_1) \cdot f_{\theta_i}(P_2) \\
    f_{\theta_i}(P_1) &= \prod_{k=1}^{m_1} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left(-\frac{x_i^2}{2\sigma_1^2}\right) \\
    f_{\theta_i}(P_2) &= \prod_{k=1}^{m_2} \frac{1}{\sqrt{2\pi\sigma_2^2}} \cdot \exp\left(-\frac{x_i^2}{2\sigma_2^2}\right)
\end{align*}
\]

Thus, we have:

\[
\begin{align*}
    J_{xx} &= \sum_{k_1=1}^{m_1} \rho_1 \left[ \frac{\cos a_1}{d_{a_1}} \right]^2 + \sum_{k_2=1}^{m_2} s \left[ \frac{\cos a_2}{d_{a_2}} \right]^2 \\
    J_{xy} = J_{yx} &= \sum_{k_1=1}^{m_1} \rho_1 \left[ \frac{\sin a_1}{d_{a_1}} \cdot \frac{\cos a_1}{d_{a_1}} \right]^2 + \sum_{k_2=1}^{m_2} s \left[ \frac{\sin a_2}{d_{a_2}} \cdot \frac{\cos a_2}{d_{a_2}} \right]^2 \\
    J_{yy} &= \sum_{k_1=1}^{m_1} \rho_1 \left[ \frac{\sin a_1}{d_{a_1}} \right]^2 + \sum_{k_2=1}^{m_2} s \left[ \frac{\sin a_2}{d_{a_2}} \right]^2
\end{align*}
\]

Therefore, the fundamental limit of localization precision with respect to \( \theta_i \) under the mixed signal distribution is calculated by:

\[
V(\theta_i) = J_{xx}^{-1} + J_{yy}^{-1} = \frac{J_{xx} + J_{yy}}{J_{xx} \cdot J_{yy} - J_{xy}^2}
\]

3.3. AP Placement Optimization

As one of the most representative heuristic optimization algorithms, the SA algorithm applies the concept of the annealing process in metallurgy to conduct the optimization search. In the SA algorithm, an initial temperature is set before the annealing process. As the temperature drops, the
SA algorithm iteratively searches for the optimal solution. Different from many existing optimization searching algorithms, like the hill climbing (HC) algorithm [30], the SA algorithm distributes an acceptance probability to each newly-obtained solution, rather than discarding the new solution, which has a higher cost compared to the previous one. Based on this, the SA algorithm is featured with the high probability of achieving the global optimum. Furthermore, the SA algorithm has been proven to consume much lower time overhead compared to the widely-used brute-force searching (BFS) [31] algorithms.

In our system, we rely on the SA algorithm to search the optimal AP locations, which correspond to the smallest value of the objective function $f = V_{\text{ave}}(\theta)$, as shown in Figure 1. Specifically, we first define the parameters used in the SA algorithm, i.e., the starting temperature $T_0$, cooling factor $\alpha$, iteration number $N$ and ending temperature $T_s$. Second, we select all of the candidate AP locations to construct the solution space $W$. Third, we construct the initial solution $W_0 = W_{\text{current}}$, where $W_{\text{current}}$ is the solution in the current iteration, and then conduct the solution updating to obtain the new solution $W_{\text{new}} = W_{\text{current}}$. Fourth, we do the searching of the optimal solutions in an iterative manner. The acceptance probability of each newly-obtained solution, $p$, is calculated by:

$$p = \min \left[ 1, \exp \left( -\frac{l (f(W_{\text{new}}) - f(W_{\text{current}}))}{T} \right) \right]$$  \hspace{1cm} (35)

where $l$ is a constant. $T$ is the temperature for the current iteration. $f(W_{\text{new}})$ and $f(W_{\text{current}})$ are the values of the objective function when the AP is located at $W_{\text{new}}$ and $W_{\text{current}}$, respectively. The temperature for the $z + 1$-th iteration equals $T_{z+1} = \alpha T_z$. To illustrate this process more clearly, the flow chart of the SA algorithm used for AP placement optimization is shown in Figure 7.

**Figure 7.** Flow chart of the SA algorithm used for AP placement optimization.
4. Simulation Results

We conduct the simulations in two typical environments, the regular LOS and irregular NLOS environments. For simplicity, we select the $k$-nearest neighbor ($k$-NN) algorithm [2] to examine the error performance of the proposed approach. The $k$-NN algorithm computes the Euclidean distances in signal space between the newly-recorded RSSs and pre-stored RSSs in the fingerprint database and then calculates the geometrical center of the $k$-nearest neighbors as the estimated location. The parameters that are also used in [32] are shown in Table 1.

Table 1. Parameters. RP, reference point.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Platform</td>
<td>MATLAB 7.0</td>
</tr>
<tr>
<td>Carrier frequency</td>
<td>2.4 GHz</td>
</tr>
<tr>
<td>$P(d_0)$</td>
<td>$-28$ dBm</td>
</tr>
<tr>
<td>$\beta$</td>
<td>2.2</td>
</tr>
<tr>
<td>Number of candidate AP locations</td>
<td>144 (LOS), 176 (NLOS)</td>
</tr>
<tr>
<td>Number of RPs</td>
<td>144 (LOS), 176 (NLOS)</td>
</tr>
<tr>
<td>Number of test points</td>
<td>100</td>
</tr>
<tr>
<td>Dimensions of the target environment</td>
<td>$12$ m $\times$ $12$ m (LOS), $36$ m $\times$ $21$ m (NLOS)</td>
</tr>
<tr>
<td>$k$</td>
<td>3</td>
</tr>
<tr>
<td>$T_0$</td>
<td>200</td>
</tr>
<tr>
<td>$N$</td>
<td>500</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.95</td>
</tr>
<tr>
<td>$T_S$</td>
<td>0.1</td>
</tr>
<tr>
<td>$P_w$</td>
<td>10</td>
</tr>
</tbody>
</table>

4.1. Regular LOS Environment

Figure 8 shows the regular LOS environment with dimensions of 12 m by 12 m. The 144 RPs (with •'s) with an interval of 1 m are uniformly calibrated in this environment.

Figure 9 shows the CDFs of errors achieved by the LLDE-APS [19], conventional max ED-based [20], symmetric and proposed fundamental limit-based AP placement optimization approaches. As the simplest approach, the symmetric AP placement optimization approach divides the target environment into a batch of subareas with the approximately same dimensions, and meanwhile, an AP is placed at the geometrical center of each subarea. As can be seen from Figure 9, we observe that the proposed approach provides smaller localization errors compared to the conventional ones when the AP number is larger than two.
4.2. Irregular NLOS Environment

Figure 10 shows the irregular NLOS environment with dimensions of 36 m by 21 m. We uniformly calibrated 176 RPs (with •’s) with an interval of 1 m in the lobby. In this environment, there are in total six office rooms, two washrooms, three corridors, one lobby, one staircase and two elevators. The six areas where the candidate APs can be placed are notated as Subareas 1–6.
Figure 10. The irregular NLOS environment.

Figure 11 shows the CDFs of errors by using the previously-introduced four AP placement optimization approaches as the AP number increases from one to four. Based on the results in Figures 9 and 11, it can be proven that in both regular LOS and irregular NLOS environments, the proposed approach can guarantee the high WLAN fingerprint-based localization precision.

Figure 11. CDFs of errors in the irregular NLOS environment. (a) With one AP; (b) with two APs; (c) with three APs; (d) with four APs.
4.3. Computation Overhead

We conduct all of the computation on a desktop with an Inter(R) core(TM) i3-3220 processor, 4 GB RAM and Windows 8 operating system. We compare the time overhead by the max ED-based AP placement optimization and the proposed AP optimization as the AP number increases from one to four. By using the SA algorithm, the time overhead involved in the max ED-based AP placement optimization is much higher than the proposed AP optimization, since the ratios of time overhead are much larger than one, as shown in Figure 12. From this figure, we observe that the proposed approach consumes much lower time overhead compared to the max ED-based AP placement optimization approach. Furthermore, the increase of the AP number enlarges the ratios of the time overhead. Based on this, the proposed approach is proven to be more efficient for the scenario with a large number of APs used for localization.

![Figure 12. Ratios of time overhead by the SA algorithm.](image_url)

Figure 13 compares the time overhead required by the BFS, SA and HC algorithms under different numbers of APs. From Figure 13, we observe that the time overhead by the HC algorithm is the lowest among these three algorithms, whereas it very easily falls into the local optimum. Different from the HC algorithm, the SA algorithm is based on the universal searching scheme and, meanwhile, requires lower time overhead compared to the BFS algorithm. Therefore, it is proven that the SA algorithm generally performs better than the HC and BFS algorithms, especially under a large number of APs.

![Figure 13. Time overhead by the brute-force searching (BFS), SA and hill climbing (HC) algorithms.](image_url)
5. Discussion

The interesting questions about the difference in the placement of the optimal AP locations, as well as the localization errors under different signal distributions remain to be answered.

Figure 14 shows the results of the placement of the optimal AP locations under the Gaussian, Rayleigh and mixed signal distributions, respectively, in the NLOS environment. From these figures, we can find that the optimal AP locations, which correspond to the lowest fundamental limit of localization precision, are generally non-collinearly distributed. To illustrate this result clearly, we take the mixed signal distribution as an example. By connecting every two most adjacent optimal AP locations with a purple line, the purple lines should be collinear when the optimal AP locations are collinearly distributed.

Figure 14. Placement of the optimal AP locations. (a) With three APs; (b) with four APs; (c) with five APs; (d) with six APs; (e) with seven APs; (f) with eight APs; (g) with nine APs; (h) with 10 APs.
Figure 15. CDFs of errors under different signal distributions. (a) With three APs; (b) with four APs; (c) with five APs; (d) with six APs; (e) with seven APs; (f) with eight APs; (g) with nine APs; (h) With 10 APs.
Figure 15 compares the CDFs of errors with respect to the Gaussian, Rayleigh and mixed signal distributions, respectively, in the NLOS environment. As can be seen from these figures, we observe that the Gaussian signal distribution generally brings the smallest localization errors, while the largest errors result from the Rayleigh signal distribution. Based on this, we can make a reasonable conjecture that a high localization precision is most likely to be provided when the APs with the LOS property to the receiver are used for localization.

6. Case Study

Based on the previous discussion, we find that the proposed approach can help much in designing a highly-precise localization system. In this section, we will continue to focus on the cases study in a real indoor WLAN environment. Specifically, the positioning precision and the corresponding time overhead are investigated under different AP numbers, AP optimization approaches and signal distributions.

As shown in Figure 16, all of the experiments are conducted in a real indoor WLAN environment with dimensions of 57 m × 25 m on the same floor in the Yi Fu building at Chongqing University of Posts and Telecommunications (CQUPT). The 10 candidate AP locations are notated as 1⃝, 2⃝, ···, 10⃝. The Samsung S7568 smartphone, which has our developed WLAN RSS scanner installed, and D-link DAP-2310 AP are selected as the receiver and transmitter, respectively, as shown in Figure 17. To construct the fingerprint database, we uniformly calibrate 73 RPs in two straight corridors and one office room, namely Areas 1, 2 and 3.
6.1. CDFs of Errors with Different AP Numbers

Figure 18 shows the CDFs of errors with different AP numbers under the Gaussian, Rayleigh and mixed signal distributions. As can be seen from this figure, the increase of the AP number improves the positioning precision, which is in accordance with the previous simulation results. However, when the AP number is larger than five, the impact of the variation of the AP number on the positioning precision is slight.

![Figure 18](image1.png)

Figure 18. CDFs of errors with different AP numbers. (a) Under the Gaussian signal distribution; (b) under the Rayleigh signal distribution; (c) under the mixed signal distribution.

6.2. CDFs of Errors by Using Different AP Optimization Approaches

To further study the effectiveness of the proposed approach, Figure 19 compares the CDFs of errors by using the proposed fundamental limits, existing LLDE-APS and max ED approaches. In Figure 19, the lower and upper bounds of errors indicate the groups of candidate AP locations corresponding to the lowest and highest mean of positioning errors, respectively. From this figure, we can find that although the proposed approach fails to achieve the highest mean of positioning errors, its error performance is still better than the existing LLDE-APS and max ED approaches.
6.3. Positioning Errors under Different Signal Distributions

We compare the mean of positioning errors corresponding to the groups of optimal candidate AP locations under the Gaussian, Rayleigh and mixed signal distributions, respectively, in Table 2. From this table, we observe that the group of optimal candidate AP locations under the mixed signal distribution achieves the lowest positioning errors compared to the Gaussian and Rayleigh signal distributions. Furthermore, the mixed signal distribution is featured with the same error performance as the Rayleigh signal distribution, which relies on the fact that the RSSs in the target environment can be fitted well by the Rayleigh signal distribution, especially when the AP number is larger than three.

Table 2. Mean of positioning errors under different signal distributions.

<table>
<thead>
<tr>
<th>AP Number</th>
<th>Gaussian Signal Distribution</th>
<th>Rayleigh Signal Distribution</th>
<th>Mixed Signal Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>3.884 m</td>
<td>5.360 m</td>
<td>3.225 m</td>
</tr>
<tr>
<td>3</td>
<td>2.911 m</td>
<td>2.974 m</td>
<td>2.701 m</td>
</tr>
<tr>
<td>4</td>
<td>2.494 m</td>
<td>2.210 m</td>
<td>2.210 m</td>
</tr>
<tr>
<td>5</td>
<td>2.255 m</td>
<td>1.966 m</td>
<td>1.966 m</td>
</tr>
<tr>
<td>6</td>
<td>2.032 m</td>
<td>1.884 m</td>
<td>1.954 m</td>
</tr>
<tr>
<td>7</td>
<td>1.8998 m</td>
<td>1.819 m</td>
<td>1.819 m</td>
</tr>
</tbody>
</table>
6.4. Time Overhead by Using Different AP Optimization Approaches

To examine the efficiency of the proposed approach, Figure 20 illustrates the time overhead involved in the proposed fundamental limits, LLDE-APS and max ED approaches, respectively, as the AP number increases from three to six. From this figure, we can find that the time overhead required by the proposed approach is significantly lower than the one by the LLED-APS or max ED approach. This result is due to the reason that the LLED-APS or max ED approach requires traversing all of the RPs in each round during the iterative process, whereas the proposed approach optimizes the AP locations by calculating the closed-form solution to the fundamental limit of localization precision only once.

![Figure 20. Time overhead by using different AP optimization approaches.](image)

6.5. Extension to a Multi-Floor Environment

We continue to conduct the experiments in a typical indoor multi-floor environment (including the fourth the fifth floors in the Yi Fu building at CQUPT), as shown in Figure 21. These two floors have the same planestructure depicted in Figure 16. The difference from the previously discussed single-floor environment can be summarized in three ways: (i) the propagation model in the multi-floor environment considers not only the signal fading by the walls, but also the one by the floors; (ii) the estimated location is a two-dimensional coordinate in the single-floor environment, while the one in the multi-floor environment is a three-dimensional coordinate; and (iii) the FIM of the estimated location in the single-floor environment is a symmetric two by two matrix, while the one in the multi-floor environment is a symmetric three by three matrix. Based on this, the previous results under the single-floor condition can be easily extended into the multi-floor scenario. In addition, to perform the testing, the nine candidate AP locations are notated as 1, 2, ..., 9, and meanwhile, the target environment is divided into five subareas, namely Areas 1, 2, 3, 4 and 5. To construct the fingerprint database, we uniformly calibrate 53 and 73 RPs on the fourth and fifth floors, respectively.
6.5.1. CDFs of Errors by Using Different AP Optimization Approaches

To investigate the performance of the proposed approach in the indoor multi-floor environment, Figure 22 compares the CDFs of errors by using the proposed fundamental limits, existing LLDE-APS and max ED approaches. From this figure, it is proven that the proposed approach can generally achieve the highest positioning accuracy compared to the existing AP optimization approaches in the indoor multi-floor environment.

Figure 22. CDFs of errors by using different AP optimization approaches. (a) With three APs; (b) with four APs; (c) with five APs; (d) with six APs.
6.5.2. Positioning Errors under Different Signal Distributions

Table 3 compares the mean of positioning errors corresponding to the groups of optimal candidate AP locations under the Gaussian, Rayleigh and mixed signal distributions, respectively. In this table, we can find that by using the groups of optimal candidate AP locations, the mixed signal distribution always achieves the smaller mean of positioning errors than the one by the Gaussian or Rayleigh signal distribution. Furthermore, the groups of optimal candidate AP locations under different AP numbers are illustrated in Table 4.

Table 3. Mean of positioning errors under different signal distributions.

<table>
<thead>
<tr>
<th>AP Number</th>
<th>Gaussian Signal Distribution</th>
<th>Rayleigh Signal Distribution</th>
<th>Mixed Signal Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>5.07 m</td>
<td>3.65 m</td>
<td>3.65 m</td>
</tr>
<tr>
<td>4</td>
<td>3.55 m</td>
<td>3.59 m</td>
<td>3.26 m</td>
</tr>
<tr>
<td>5</td>
<td>3.25 m</td>
<td>3.31 m</td>
<td>3.19 m</td>
</tr>
<tr>
<td>6</td>
<td>3.12 m</td>
<td>3.32 m</td>
<td>3.12 m</td>
</tr>
<tr>
<td>7</td>
<td>3.08 m</td>
<td>2.93 m</td>
<td>2.87 m</td>
</tr>
<tr>
<td>8</td>
<td>2.87 m</td>
<td>3.02 m</td>
<td>2.78 m</td>
</tr>
<tr>
<td>9</td>
<td>2.77 m</td>
<td>2.77 m</td>
<td>2.77 m</td>
</tr>
</tbody>
</table>

Table 4. Groups of optimal candidate AP locations under different signal distributions.

<table>
<thead>
<tr>
<th>AP Number</th>
<th>Gaussian Signal Distribution</th>
<th>Rayleigh Signal Distribution</th>
<th>Mixed Signal Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>1 3 5 9</td>
<td>1 6 7</td>
<td>1 6 7</td>
</tr>
<tr>
<td>4</td>
<td>1 3 4 5 7</td>
<td>1 3 4 5 8</td>
<td>4 5 6 7</td>
</tr>
<tr>
<td>5</td>
<td>3 4 6 7 8 9</td>
<td>1 3 6 8 9</td>
<td>2 4 6 7 8</td>
</tr>
<tr>
<td>6</td>
<td>2 3 4 6 8 9</td>
<td>1 3 5 6 7 9</td>
<td>2 3 4 6 7 9</td>
</tr>
<tr>
<td>7</td>
<td>2 3 5 6 7 9</td>
<td>3 4 5 6 7 9</td>
<td>1 2 3 6 7 8</td>
</tr>
<tr>
<td>8</td>
<td>1 2 3 4 5 6 7 9</td>
<td>1 3 4 5 6 7 9</td>
<td>1 2 3 4 5 6 7 9</td>
</tr>
<tr>
<td>9</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td>1 2 3 4 5 6 7 8 9</td>
<td>1 2 3 4 5 6 7 8 9</td>
</tr>
</tbody>
</table>

7. Conclusions

In this paper, we proposed a novel information-based approach to analyze the localization precision, as well as to optimize the AP placement for indoor WLAN localization using the location fingerprint. We derived the fundamental limit of WLAN fingerprint-based localization precision by using the FIM and then relied on the SA algorithm to search the optimal AP locations, which correspond to the lowest fundamental limit of localization precision. Compared to the widely-used max ED-based and symmetric AP placement approaches, the proposed approach performs better in the aspects of localization errors and time overhead. Moreover, experimental results are presented in order to support our claims. For the future work, the information-based precision analysis of WLAN fingerprint-based localization in a multi-floor environment forms an interesting topic.

Acknowledgments: This work was supported in part by the Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), the National Natural Science Foundation of China (61301126), the Special Fund of Chongqing Key Laboratory (CSTC), the Fundamental and Frontier Research Project of Chongqing (cstc2013jcyjA40041, cstc2012jjiA40040), the Science Foundation for Young Scientists of CQUPT (A2012-77) and “Wenfeng Top-level Talent Leading Project” of QUPPT.

Author Contributions: The work presented in this paper corresponds to a collaborative development by all authors. Mu Zhou defined the research line and was involved in writing the paper. Feng Qiu developed the proposed system, performed the data analysis and wrote the paper. Zengshan Tian and Haibo Wu developed the application of the collection data. Qiao Zhang and Wei He were involved in performing the experiments. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.
References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).