Joint Access Point Selection and Local Discriminant Embedding for Energy Efficient and Accurate Wi-Fi Positioning

Zhi-An Deng, Yu-Bin Xu and Lin Ma
Communication Research Center, Harbin Institute of Technology
Harbin, 150080, China
[e-mail: dengzhianan@163.com; {ybxu, malin}@hit.edu.cn]
*Corresponding author: Zhi-An Deng

Received December 10, 2011; revised February 14, 2012; accepted March 6, 2012; published March 25, 2012

Abstract

We propose a novel method for improving Wi-Fi positioning accuracy while reducing the energy consumption of mobile devices. Our method presents three contributions. First, we jointly and intelligently select the optimal subset of access points for positioning via maximum mutual information criterion. Second, we further propose local discriminant embedding algorithm for nonlinear discriminative feature extraction, a process that cannot be effectively handled by existing linear techniques. Third, to reduce complexity and make input signal space more compact, we incorporate clustering analysis to localize the positioning model. Experiments in realistic environments demonstrate that the proposed method can lower energy consumption while achieving higher accuracy compared with previous methods. The improvement can be attributed to the capability of our method to extract the most discriminative features for positioning as well as require smaller computation cost and shorter sensing time.

Keywords: Indoor positioning, pervasive computing, Wi-Fi, energy efficient

This research was supported by the High-Tech Research and Development Program of China (Grant No. 2008AA12Z305), the National Science Foundation for Young Scientists of China (Grant No. 61101122), and China Postdoctoral Science Foundation funded project (Grant No. 20100471057).

DOI: 10.3837/tiis.2012.03.002

1. Introduction

With the rapid development of wireless technologies and pervasive availability of mobile devices, wireless positioning systems have become increasingly important in pervasive computing. Such systems enable a large number of location-based services [1], including emergency caller location identification, people navigation, and healthcare. For indoor use, Wi-Fi positioning system [2][3] is particularly preferred for its cost effectiveness. Most positioning systems deploy client-based architecture for its protection of the privacy of users [4]. However, this architecture processes all computation and sensing operations on power-limited mobile devices. Thus, a major challenge for Wi-Fi positioning is the reduction of energy consumption of mobile devices while ensuring high-level accuracy. A highly accurate positioning system may be useless [5][6] if it requires a frequent recharging of the mobile device. As a result, accuracy is not the only goal of positioning systems. The reduction of energy consumption is equally important.

To save energy while achieving high-level accuracy, existing approaches generally fall into two categories: access point (AP) selection and feature extraction. In AP selection, only a subset of available APs is selected for positioning by measuring and ranking the importance of each AP independently. Several studies [6][7] have shown that a suitable subset of APs may reduce computation cost while ensuring accurate positioning. In feature extraction, a small number of the most significant features are generated and used for positioning by feature extraction methods, including principal component analysis (PCA) [8] and linear discriminant analysis (LDA) [9]. PCA and LDA may discard noisy information and reorganize location information more compactly, hence improving accuracy and reducing computation cost. However, both AP selection and feature extraction have their drawbacks. Current AP selection methods disregard the correlation of discriminant abilities of APs, which renders the previous measures of importance of each AP inaccurate. Furthermore, substantial redundancy still exists among chosen APs. In contrast, feature extraction methods, such as LDA and PCA, are both linear techniques that fail to handle the severe nonlinearity [10] of Wi-Fi positioning.

This paper focuses on improving positioning accuracy while reducing energy consumption of mobile devices. We propose a novel method called LLDE-APS, which stands for the combination of Localized Local Discriminant Embedding (LDE) [11] and AP Selection. The proposed LLDE-APS improves accuracy by extracting location features with the strongest discrimination power. More importantly, LLDE-APS also reduces energy consumption considerably by minimizing the computation cost and sensing time required on mobile devices. Our method presents the following three contributions: First, we develop a joint and intelligent AP selection scheme based on the maximum mutual information (MMI) criterion. This scheme measures and ranks the discriminant ability of APs more accurately than previous schemes. Second, we further propose LDE to extract nonlinear discriminative features from selected APs. LDE effectively adapts the nonlinearity of RSS while maintaining energy efficiency for its linear property. Third, we employ the fuzzy c-means (FCM) clustering algorithm to localize the positioning model into sub-regions, where more compact input signal space and reduced model complexity are achieved.

Experiments in a realistic office environment show that, compared with the weighted K-nearest neighbor (WKNN) and maximum likelihood (ML) methods, the proposed method achieves higher accuracy while reducing computation cost and sensing time by 90.2% and 80.0%, respectively.

2. Related Work

Indoor location information has become increasingly important in pervasive computing because people tend to spend most of their time in indoor environments, such as shopping centers or office buildings. Various indoor positioning systems [12], including radio frequency, ultra wideband, visual sensors, and Wi-Fi, have been developed. Among these systems, Wi-Fi positioning that uses received signal strength (RSS) has attracted the most attention because of its low cost. RSS can be easily obtained by the wireless adapter available in mobile devices, and no additional hardware is needed. Most Wi-Fi positioning systems deploy fingerprinting architecture [13], which has offline and online stages. In the offline stage, RSS fingerprints at reference locations are collected to construct the database called radio map. In the online stage, a real-time RSS signal is compared with the pre-stored radio map to estimate the location. This architecture aims to learn and establish the relationship between RSS and physical locations. Thus, several machine learning algorithms have been applied, including WKNN [14], ML [7], and artificial neural network (ANN) [15].

In Wi-Fi positioning, accuracy and energy efficiency are often contradicting goals. For example, a common method for improving accuracy is to use as many APs as possible during positioning. However, a greater number of APs used result in a higher computation cost required on mobile devices. Another way to improve accuracy is to collect more RSS samples in real-time [16]. However, the addition of sensing time may also increase the energy consumption of mobile devices. To avoid frequent recharging of mobile devices, energy consumption should be reduced as much as possible. Thus, a key challenge in Wi-Fi positioning is the achievement of accurate positioning while ensuring energy efficiency.

To address this challenge, several AP selection methods have been proposed for positioning. Youssef et al. [7] chose a subset of available APs with the largest mean RSS values to reduce computation cost. However, large mean RSS values always result in large RSS variance at a fixed location. In such a case, distinguishing neighbor locations becomes more difficult, thus degrading the positioning accuracy. Chen et al. [6] considered measuring the discriminant ability of each AP and selected the most discriminative APs for positioning. In their work, the discriminant ability of each AP is measured independently based on location information gain value. These studies show that computation cost reduction may be achieved by selecting a proper subset of available APs. However, their approaches measure the discriminant ability of APs independently and disregard the interplay between APs. Furthermore, considerable redundancy still abounds among chosen APs because of their correlation. We thus propose a joint and intelligent AP selection scheme to choose the optimal subset of APs, where feature extraction is further incorporated to reduce redundancy among chosen APs.

Another approach to address the challenge is through feature extraction methods. Fang et al. [8] proposed PCA to extract location features in a decorrelated space. Rather than discarding noisy APs, PCA extracts the feature components representing most of the data variance information in the original RSS space. However, the most descriptive features extracted by PCA are not necessarily suitable for discriminating different locations. Fang et al. [9] reported that LDA showed better accuracy performance than PCA because discriminative information was better captured by the former. However, both LDA and PCA suffer from severe nonlinearity of RSS. Nonlinearity of RSS results from complex signal propagation factors, such as the multipath effect, shadowing, and attenuations attributable to obstructions. Nonlinearity of RSS may be effectively addressed by kernel-based methods, such as kernel canonical correlation analysis [10] and kernel direct discriminant analysis (KDDA) [17]. Our previous work proposed the use of KDDA to extract discriminative features in a kernel-based space, where nonlinear RSS patterns may be captured well. Despite the higher accuracy, high computation cost is required by KDDA to compute kernel functions. Thus, we propose LDE for discriminative feature extraction. LDE is capable of adapting to the nonlinearity of RSS, but computation efficiency is also maintained for its linear property.

In addition, clustering analysis methods have been introduced to reduce computation cost in a large-scale positioning region. Clustering analysis divides the entire region into several sub-regions and independently develops the positioning model in each sub-region. Youssef et al. [7] first used a joint clustering technique to group reference locations. Locations sharing the same AP subset with the largest RSS values are clustered into one sub-region. However, due to uncertainty of RSS [18], the AP subset with the largest RSS values may vary with time and degrade the classification performance. Chen et al. [6] applied k-means clustering to generate clusters automatically. The k-means compresses the input space by maximizing within-class similarity and minimizing between-class similarity. However, the k-means algorithm also suffers from uncertainty and nonlinearity of RSS. We thus propose FCM for clustering analysis, which performs better than k-means through the use of fuzzy principles.

3. The Proposed LLDE-APS Method

This section first outlines the proposed LLDE-APS method. Clustering analysis using the FCM clustering algorithm is then introduced to localize the positioning model into sub-regions. Then, we propose a novel joint and intelligent AP selection scheme based on the MMI criterion. Finally, we further propose LDE to extract the nonlinear discriminative features.

3.1 Overview of LLDE-APS

The proposed LLDE-APS method includes offline and online stages, as seen in **Fig. 1**. The offline training stage can be performed on a powerful server with sufficient computation resources, whereas the online positioning stage is performed on source-weak mobile devices.

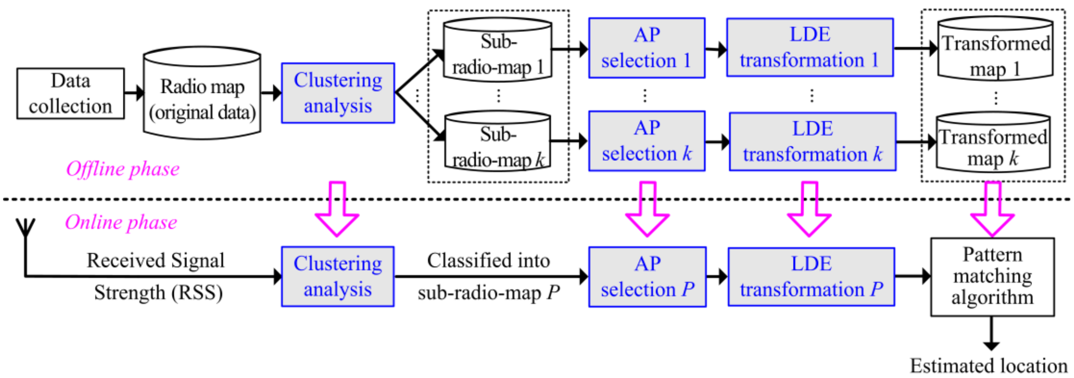


Fig. 1. Architecture of the proposed LLDE-APS method

During the offline stage, we first collect RSS values at pre-defined reference locations to build the radio map. The entire region is then divided into k sub-regions using clustering analysis. Each sub-region corresponds to a sub-radio-map. We then develop an independent AP selection strategy based on the MMI criterion and obtain the transformation matrix of LDE in each sub-radio-map. Finally, we build the related transformed map for each sub-radio-map. The transformed maps include the average feature vector at each reference location, which is computed by averaging the transformed feature samples at each reference location.

During the online stage, we first classify the real-time RSS signal into a related sub-radio-map P. The real-time feature vector is then obtained based on AP selection and LDE transformation P. The Euclidean distances between the real-time feature vector and pre-stored average feature vectors at each reference location are computed. Finally, the location is estimated by linearly combining the coordinate vectors of the K nearest reference locations, the weights of which are corresponding inverse distances. In fact, the pattern matching algorithm used in LLDE-APS is the widely used WKNN algorithm. However, the input signals are not the original RSS signals. We explore AP selection to discard noisy APs and then extract discriminative features by LDE. As a result, the input singals become the location features with enhanced discrimination power and reduced dimension.

3.2 Clustering Analysis for Localizing the Positioning Model

3.2.1 Necessity for Localizing the Positioning Model

Localizing Wi-Fi positioning for a large-scale positioning region has two advantages. First, a localized positioning model improves positioning accuracy. The optimal AP selection and subsequent feature extraction solutions are determined by associated statistical properties of RSS signals, such as RSS variance at a fixed location and RSS differences among different locations. However, these statistical properties always vary with physical locations. For example, RSS variances at reference locations nearer to the AP are always larger than those farther from the AP. Therefore, a monolithic positioning model for the entire region is suboptimal. By localizing the positioning into sub-regions, more suitable AP selection and feature extraction solutions are obtained, thus improving accuracy.

Second, a localized positioning model reduces the computation cost significantly. For pattern-matching-based positioning methods, such as WKNN and ML, the computation cost is proportional to the size of the radio map. By localizing the positioning into the sub-region, the computation cost significantly decreases because of the smaller number of reference locations.

3.2.2 Clustering Analysis Using the Fuzzy C-Means Algorithm

For clustering analysis in Wi-Fi positioning, a high probability of classifying the user into the right sub-region should be ensured because a wrong classification may yield a large positioning error. However, the uncertainty and nonlinearity of RSS signals degrade the classification performance of existing clustering methods, such as k-means. We adopt the FCM clustering algorithm [19] in this paper. FCM has been widely used and has a highly robust classification performance even with noisy and nonlinear data.

FCM partitions N' reference locations $z_i\left(i=1,\ldots,N'\right)$ into k clusters $S_l\left(l=1,\ldots,k\right)$. Each cluster S_l is associated with a cluster center $\mathbf{C}_l$. The relationship between every reference location and the cluster is fuzzy. That is, the degree of z_i belonging to $\mathbf{C}_j$ is represented by a membership $u_{ij}=[0,\ 1]$. Each reference location z_i is represented by the related average RSS vector $\overline{\mathbf{x}}^i$. All reference locations is denoted as $S=\left\{\overline{\mathbf{x}}^i,\ i=1,\ldots,N'\right\}$. FCM aims to minimize the objective function: $J=\sum_{j=1}^{k}\sum_{i=1}^{N'}u_{ij}^m D_{ij}^2$ with the constrained function: $\sum_{j=1}^{k}u_{ij}=1,$ for $i=1$ to N', where $m=2$ is the fuzzifier parameter; k is the number of clusters; N' is the number of reference locations; and D_{ij} is the Euclidean distance between the average RSS vector $\overline{\mathbf{x}}^i$ and the cluster center $\mathbf{C}_j$. FCM divides the reference locations in such a way that RSS signals assigned to the same cluster should be as similar as possible, whereas two signals from different clusters should be as dissimilar as possible.

FCM involves the iterative generation of a set of fuzzy cluster centers and associated memberships of each reference location. FCM starts by randomly selecting k reference locations as initial cluster centers and is processed as follows:

1. Input a set of k initial cluster centers $SC_0 = \left\{ \mathbf{C}_j\left(0\right), j = 1\ldots,k \right\}$. Set $p = 1$.

2. For cluster centers $SC_p = \left\{ \mathbf{C}_j\left(p\right), j = 1,\ldots,k \right\}$ at the p th step, compute the Euclidean distance $\left\{ D_{ij},\ i = 1,\ldots,N',\ j = 1,\ldots,k \right\}$. Update memberships u_{ij} : $u_{ij} = \left(\sum_{l=1}^{k} \left(D_{ij} / D_{il} \right)^{1/(m-1)} \right)^{-1}$.

3. Update the cluster center: $\mathbf{C}_j\left(p+1\right) = \sum_{i=1}^{N'} u_{ij}^m \overline{\mathbf{x}}^i \Big/ \sum_{i=1}^{N'} u_{ij}^m$ to obtain SC_{p+1}.

4. Stop the iterations if $\left\| C_j\left(p\right) - C_j\left(p-1\right) \right\|_2 < \varepsilon, j = 1,\ldots,k$ is satisfied, where $\varepsilon > 0$ is a small value; otherwise, set $p+1 \rightarrow p$, and go to step 2.

During the offline stage, we generate the fuzzy clusters and the related fuzzy cluster centers. During the online stage, we first compute the Euclidean distance between the real-time RSS vector and the pre-stored fuzzy cluster centers, and the user is then classified into the cluster, the cluster center of which is nearest to the real-time RSS vector.

3.3 Joint and Intelligent Access Point Selection

3.3.1 Motivation

First, we present the motivation for AP selection.The rapid development of Wi-Fi and the increasing demand for wideband mobile communications have resulted in the deployment of a high-density of APs at many indoor environments, such as offices and universities [20]. Dozens of APs are always available at a fixed location. Thus, a simple way to improve positioning accuracy is to use as many APs as possible because each AP may provide a unique view of the user's physical location. However, a high-dimensional input space comprising a large number of APs may increase computation cost and thus give rise to the overfitting problem. Furthermore, not all APs improve positioning accuracy. Some APs with poor discriminant ability may introduce noise and degrade accuracy. Therefore, selecting the most discriminative APs and discarding noisy ones are necessary.

We then present the motivation for the proposed joint and intelligent AP selection scheme. Current AP selection schemes measure the discriminant ability of each AP individually. However, they all ignore the correlation of the discriminant abilities of APs, thus incurring an inaccurate measurement of the joint discriminant ability of APs. We propose a joint and intelligent AP selection scheme based on the MMI criterion (called JointMMI) [21]. Mutual information measures the amount of uncertainty reduction given knowledge about the RSS values from APs. The subset of available APs with the MMI is chosen for positioning.

3.3.2 AP Selection Based on MMI Criterion

We first introduce the computation of individual mutual information gained from each AP and then generalize the computation into the joint mutual information gain. The individual mutual information gained from AP_i is obtained by computing the two entropy values as follows:

$$MuInfo(AP_i) = H(\mathbf{L}) - H(\mathbf{L} | AP_i) \tag{1}$$

where $\mathbf{L}$ denotes the user's coordinate variable in target positioning region; $H(\mathbf{L})$ denotes the entropy of the coordinate variable without knowing any information from AP; and

$H(\mathbf{L}|AP_i)$ denotes the conditional entropy of the coordinate variable knowing RSS from AP_i. In fingerprinting positioning, we collect RSS values from APs at discrete reference locations. Thus, we compute the entropy values approximately over the reference locations:

$$\begin{aligned} H(\mathbf{L}) &= -\sum_{j=1}^{N} Pr(\mathbf{L}_j)\log Pr(\mathbf{L}_j) \\ H(\mathbf{L}|AP_i) &= -\sum_{v}\sum_{j=1}^{N} Pr(\mathbf{L}_j, AP_i = v)\log Pr(\mathbf{L}_j|AP_i = v) \end{aligned} \tag{2}$$

where $\mathbf{L}_j$ indicates the coordinate vector of the j th reference location; v stands for all possible RSS values from AP_i; N is the number of reference locations in the sub-region; $Pr(\mathbf{L}_j)$ is the prior probability of the j th reference location, which is assumed to be uniformly distributed; and $Pr(\mathbf{L}_j|AP_i = v)$ is the conditional probability computed as follows:

$$Pr(\mathbf{L}_j|AP_i = v) = Pr(AP_i = v|\mathbf{L}_j)Pr(\mathbf{L}_j)\big/Pr(AP_i = v) \tag{3}$$

In (3), $Pr(AP_i = v|\mathbf{L}_j)$ is computed through the statistical result at the reference location.

We then show how to compute the joint mutual information gained from APs subsets. Using $(AP_1,\ldots,AP_t)$ for positioning, associated joint mutual information is computed as follows:

$$\begin{aligned} &MuInfo(AP_1,\ldots,AP_t) = H(\mathbf{L}) - H(\mathbf{L}|AP_1,\ldots,AP_t) \\ &H(\mathbf{L}|AP_1, AP_2,\ldots,AP_t) = -\sum_{v_1}\cdots\sum_{v_t}\sum_{j=1}^{N} Pr(\mathbf{L}_j, AP_1 = v_1,\ldots,AP_t = v_t)\cdot\log\left(Pc\right) \end{aligned} \tag{4}$$

where $Pc{=}Pr(\mathbf{L}_j|AP_1 = v_1,\ldots,AP_t = v_t)$ is computed as given:

$$Pr(\mathbf{L}_j|AP_1 = v_1,\ldots,AP_t = v_t) = Pr(AP_1 = v_1,\ldots,AP_t = v_t|\mathbf{L}_j)Pr(\mathbf{L}_j)\big/Pr\left(AP_1 = v_1,\ldots,AP_t = v_t\right) \tag{5}$$

RSS values from different APs at fixed locations can be assumed to be independent from one another [7]. Then, (5) can be computed as follows:

$$\begin{aligned} Pr(AP_1 = v_1,\ldots,AP_t = v_t|\mathbf{L}_j) &= \prod_{i=1}^{t} Pr(AP_i = v_i|\mathbf{L}_j) \\ Pr(AP_1 = v_1,\ldots,AP_t = v_t) &= \sum_{j=1}^{N} Pr(AP_1 = v_1,\ldots,AP_t = v_t|\mathbf{L}_j)Pr(\mathbf{L}_j) \end{aligned} \tag{6}$$

The computation cost of the proposed AP selection scheme is substantive and increases with the number of reference locations and APs. However, the number of reference locations is always limited in a sub-radio-map after clustering analysis. More importantly, we may perform it during offline stage, which does not incur any online computation cost. In contrast, online computation cost may be reduced by AP selection because the number of APs involved in online positioning is reduced significantly.

3.4 Feature Extraction Using LDE

3.4.1 Advantages of Using LDE

Wi-Fi positioning can be viewed as a pattern classification problem. Each reference location is considered a class labeled by related RSS samples. Classification performance is primarily determined by two factors, namely, ratio of training sample size to feature dimensions and discrimination power of the features. Although AP selection may reduce feature dimension by discarding noisy APs, substantive redundancy still exists among chosen APs. Thus, feature extraction methods, including LDA and PCA, have been used to reduce dimensionality. However, both of them become less efficient with the severe nonlinearity of RSS.

Nonlinear dimensionality reduction for Wi-Fi positioning can be addressed based on the manifold assumption. Manifold-based dimensionality reduction methods [22][23] have been successfully applied in numerous machine learning problems. The manifold assumption involves high-dimensional data that lie close to a low-dimensional nonlinear manifold embedded in ambient space. In the presence of little noise during Wi-Fi positioning, each two-dimensional physical location vector will uniquely determine a high-dimensional RSS vector. Thus, the original RSS signal space will be mapped onto a two-dimensional manifold. Although RSS always appears to be noisy, RSS data are always assumed to lie close to a nonlinear manifold embedded in ambient space.

Under the manifold assumption, LDE is a promising technique for nonlinear dimensionality reduction. Manifold is approximately represented by constructing a neighborhood graph. Thus, nonlinearity is handled well by preserving locality geometry. Simultaneously, class information is explored to enhance the discrimination power. Moreover, several nonlinear techniques can be applied for nonlinear dimensionality reduction. However, they are always computationally expensive. Conversely, LDE may effectively extract low-dimensional nonlinear features, while retaining low computation cost for its linear property.

3.4.2 LDE

Given n-dimensional RSS samples $\mathbf{x}_i \quad (i=1,\ldots,M)$, LDE extracts features as given: $\mathbf{y}_i = \mathbf{A}^T\mathbf{x}_i$, where $\mathbf{A}=[\mathbf{a}_1,\mathbf{a}_2,\ldots,\mathbf{a}_d]$ is the transformation matrix, $\mathbf{y}_i = \left[y_i^1, y_i^2,\ldots,y_i^d\right]^T$ is the d-dimensional feature vector of $\mathbf{x}_i$, y_i^r is the r th feature component, l_i is the class label of $\mathbf{x}_i$, and M is the number of RSS samples in the sub-region. LDE aims to keep neighbor samples of the same class close and to distance neighbor samples of different classes.

To achieve this objective, two undirected neighborhood graphs G and G' are first constructed over all samples. The graph G adds the edge between $\mathbf{x}_i$ and $\mathbf{x}_j$ if $l_i = l_j$ and $\mathbf{x}_i$ is a neighbor sample of $\mathbf{x}_j$. For G', an edge between $\mathbf{x}_i$ and $\mathbf{x}_j$ is added if $l_i \neq l_j$ and $\mathbf{x}_i$ is a neighbor sample of $\mathbf{x}_j$. $\mathbf{W}$ and $\mathbf{W}'$ are sparse and symmetric affinity matrices of G and G', respectively. LDE aims to address the following optimization problem:

$$\text{Maximize } J(\mathbf{A}) = \sum_{i,j}\left\|\mathbf{A}^T\mathbf{x}_i - \mathbf{A}^T\mathbf{x}_j\right\|^2 w'_{ij}, \text{ Subject to } \sum_{i,j}\left\|\mathbf{A}^T\mathbf{x}_i - \mathbf{A}^T\mathbf{x}_j\right\|^2 w_{ij} = 1 \tag{7}$$

where, if $\mathbf{x}_i$ connects $\mathbf{x}_j$ in G', $w'_{ij} = \exp\left(-\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2 / \sigma'\right)$; otherwise, $w'_{ij} = 0$. If $\mathbf{x}_i$ connects $\mathbf{x}_j$ in G, $w_{ij} = \exp\left(-\left\|\mathbf{x}_i - \mathbf{x}_j\right\|^2 / \sigma\right)$; otherwise, $w_{ij} = 0$.

The optimization in (7) is solved as a generalized eigenvalue problem:

$$\mathbf{X}(\mathbf{D}'-\mathbf{W}')\mathbf{X}^T\mathbf{a}_i=\lambda_i\mathbf{X}(\mathbf{D}-\mathbf{W})\mathbf{X}^T\mathbf{a}_i,\ i=1,\ldots,d \tag{8}$$

where $\mathbf{X}=[\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_M]$, $\lambda_1\geq\lambda_2\geq\ldots\geq\lambda_d$, w'_{ij} and w_{ij} are the elements of $\mathbf{W}'$ and $\mathbf{W}$, respectively, and the diagonal elements of diagonal matrices $\mathbf{D}'$ and $\mathbf{D}$ are $d'_{ii}=\sum_j w'_{ij}$ and $d_{ii}=\sum_j w_{ij}$, respectively. Each feature component corresponds to an eigenvalue:

$$\lambda_i=\mathbf{a}_i^T\mathbf{X}(\mathbf{D}'-\mathbf{W}')\mathbf{X}^T\mathbf{a}_i/\mathbf{a}_i^T\mathbf{X}(\mathbf{D}-\mathbf{W})\mathbf{X}^T\mathbf{a}_i,\ i=1,\ldots,d \tag{9}$$

where the numerator of the right side in (9) measures the between-class distance of neighbor samples, and the denominator measures the within-class distance of neighbor samples. A larger between-class distance and smaller with-class distance render stronger discrimination power of the feature component. Thus, the discrimination power of a feature vector can be indicated by the accumulative percentage of related eigenvalues. We determine the dimension of the extracted features by setting a threshold β^*:

$$\beta(d)=\sum_{i=1}^{d}\lambda_i\Big/\sum_{j=1}^{n}\lambda_j\geq\beta^* \tag{10}$$

The smallest integer value satisfying (10) is set as the feature dimension. In the offline training stage, we find that a practical approach is to choose β^* between 80% and 90%.

4. Analysis on Energy Savings in the LLDE-APS Method

Energy efficiency in wireless networks has attratced considerable attention [24][25]. This paper focuses on reducing the energy consumption of mobile devices while achieving high-level Wi-Fi positioning accuracy. We explore three approaches in LLDE-APS. First, rather than broadcasting a probe frame, we collect RSS signals from APs in a passive scanning mode, which saves a significant amount of energy. Second, we reduce the computation cost on mobile devices by minimizing the dimension of location features involved in positioning. The third approach is the maximization of the sleep time of the mobile device by minimizing the required number of RSS samples to be collected in real-time.

A network interface card (NIC) has two modes in collecting RSS signals from APs, namely, passive scanning and active scanning. In passive scanning, NIC receives the beacons periodically sent by APs. The signal strength and identification of the AP contained in the beacons are collected and used for positioning. In active scanning, rather than receiving the beacons from the APs, NIC broadcasts a probe frame, and all APs within range respond with a probe response. NIC then obtains RSS signals from the probe response frames. The major advantage of the active scanning mode is that it may obtain RSS signals from APs immediately, without waiting for a beacon period. However, the default beacon period is always 100 ms, which is sufficiently small for real-time positioning. More importantly, the active scanning mode consumes substantially more power than the passive mode. As shown in [5][26], the energy consumption mode of NIC in a Wi-Fi system includes the Sleep mode, Idle mode, and the two transmission modes. The two transmission modes comprise transmitting packages (TX) and receiving packages (RX). Ebert et al. [26] reported that, in the Sleep, Idle, RX, and TX modes, the average energy consumption values were approximately 20, 110, 900, and 2500 MW, respectively. The adopted passive scanning mode does not need to send probe frames actively, thus decreasing energy consumption compared with the active mode.

Another approach for saving energy is the reduction of the computation cost on mobile devices. In our method, we first localize the positioning model into a sub-region. In each sub-region, the number of reference locations is considerably smaller than that in the entire positioning region. We then reduce the dimension of input signal space further by discarding noisy and useless APs. Finally, we deploy LDE to extract a small set of the most discriminative features for positioning. As will be seen in Section 5.5, compared with the other methods, the dimension of location features is minimized by our method, thus decreasing the computation cost on mobile devices.

The third approach is to lengthen sleep time of mobile devices by minimizing sensing time. Our method ensures this condition by minimizing the number of real-time RSS samples required while achieving the same accuracy. Ergen et al. [27] reported that most of the energy was wasted in overhearing data intended for other stations. Without a transmission task, NIC may revert to the sleep mode to save energy until a “wake up” packet arrives. Thus, reducing energy consumption requires a reduction in sensing time. You et al. [28] and Zhuang et al. [29] reported that users may trade higher energy consumption reduction for lower positioning accuracy by adapting sensing frequency. In contrast, we optimize the positioning method itself to make the algorithm robust to the number of reduced real-time samples. Our method effectively extracts the most discriminative features while discarding noisy and redundant information. In Section 5.6, we will show that our method outperforms other methods in terms of accuracy despite using only a small fraction of real-time samples. Thus, our method facilitates a longer sleep time during online positioning.

5. Experimental Results and Discussions

In this section, we first describe the experimental setup. Next, clustering analysis results are explored. We then compare the performance of the proposed AP selection and feature extraction methods with previous approaches (Sections 5.3 and 5.4). We subsequently demonstrate the online computation cost and sensing time reduction of the proposed method (Sections 5.5 and 5.6). Finally, we validate the improvement in accuracy of the proposed method in two classic indoor positioning environments.

5.1 Experimental Setup

To evaluate the proposed LLDE-APS, we performed the experiments in a classic indoor office environment (i.e., part of the building of our department), as shown in **Fig. 2**. We collected realistic RSS samples within the shading target region, the size of which was $25.3\ \mathrm{m} \times 24.7\ \mathrm{m}$. A total number of 24 APs were available in the region, and an average number of 15 APs were covered at each location, with 8 APs seen in the target region and the other APs in the adjacent regions or other floors. A total of 5,500 training samples at 55 reference locations were collected to build the radio map, with 100 samples per location. We also collected 5500 testing samples at 55 test locations to evaluate the performance. Reference locations were separated from one another by about 2 m, and test locations were uniformly selected in the target region. The sensing rate is 2 samples per second. We obtained a real-time RSS sample vector by averaging two testing samples and using it to estimate the location.

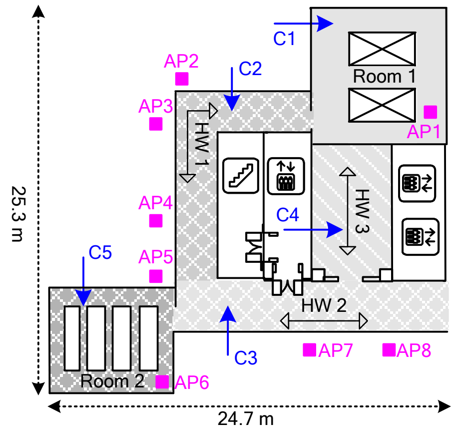


Fig. 2. Experimental environment includes two rooms (Rooms 1 and 2) and three hallways (HW 1, HW 2, and HW 3). The entire region is divided into five clusters, namely, C1, C2, C3, C4, and C5.

The entire positioning region is divided into five clusters: C1, C2, C3, C4, and C5. The number of reference locations is 16, 12, 9, 8, and 10 from C1 to C5, respectively. The same is true for the number of test locations. In each cluster, an independent AP selection scheme and LDE transformation solution are developed. For simplicity, the AP selection and feature extraction results are only presented and discussed in C1, although similar improvements are obtained in the other clusters. We define positioning accuracy as the cumulative error distance distribution. The error distance is the Euclidean distance between the actual coordinate of the user and the estimated result. We split the training samples into 10 folds and use 10-fold cross validation to obtain the optimal parameters of LLDE-APS. The real parameter values of the positioning model are set by the best validation accuracy.

5.2 Clustering Analysis Result

In this section, we first compare the classification accuracy of the proposed FCM with the widely used k-means algorithm. Classification accuracy is defined as the probability of classifying the user into the correct sub-region. **Fig. 3** compares the offline classification accuracy between FCM and k-means versus the number of clusters. The classification accuracy decreases gradually as the number of clusters increases because the smaller clusters are more similar to one another, making them difficult to distinguish from one another. However, the classification accuracy of FCM is higher than that of k-means algorithm. FCM effectively handles the uncertainty of RSS and generates more representative cluster centers by introducing fuzzy principles. The effectiveness of FCM can also be verified by the clustering result shown in **Fig. 2**. The physical adjacent reference locations are all divided into the same sub-regions, rendering more compact signal space. We only deploy 13 APs with maximum RSS values for clustering analysis because the classification accuracy is comparable to that using all APs, while reducing computation cost.

We then show the effect of the number of clusters on positioning accuracy. Both the FCM and k-means algorithms are combined with the proposed AP selection scheme and LDE algorithm. As shown in **Fig. 4**, clustering analysis using FCM results in significant accuracy improvement. Compared with the positioning model built over the entire region, the localized model with cluster number $k = 5$ increases the testing accuracy within 2 m from 67.2% to 73.7%. In each sub-region, a more suitable AP selection scheme and LDE transformation matrix are obtained in the more compact RSS signal space. However, the smaller clusters have lower classification accuracy. The best balance is kept when the number of clusters $k=5$. **Fig. 4** also shows that FCM achieves higher positioning accuracy than k-means because the classification accuracy of the former is higher than that of the latter. Furthermore, little online computation cost is required on mobile devices because the generation of fuzzy cluster centers in FCM can be performed offline. The online computation cost of FCM is the same as that of k-means, which merely needs to compute the Euclidean distance between the real-time RSS sample and the k cluster centers to identify the nearest cluster center.

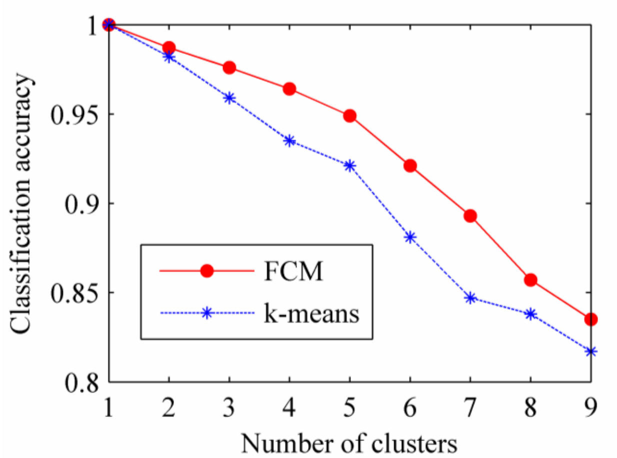


Fig. 3. Classification accuracy versus the number of clusters

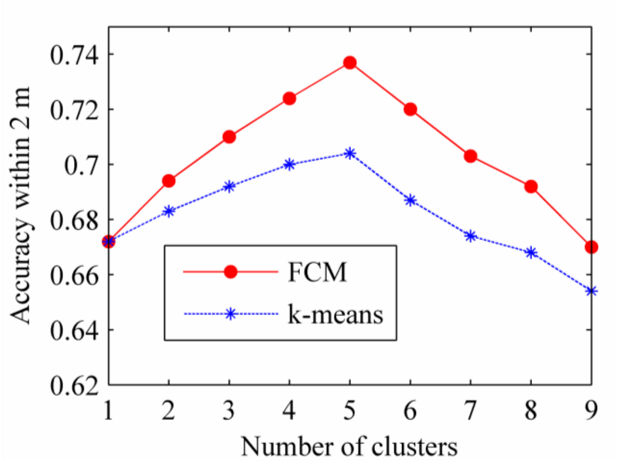


Fig. 4. Accuracy within 2 m versus the number of clusters

5.3 AP Selection Result

In this section, we study the effect of various AP selection schemes on the testing accuracy of LLDE-APS in cluster 1, which has 21 APs. Various AP selection schemes are combined with related optimal LDE transformation. **Fig. 5** compares the accuracy within 2 m among four AP selection schemes: InfoGain, in which the APs are ranked in descending order of the individual information gain values; MaxMean, in which the APs are ranked in descending order of the mean RSS values; Random, in which the APs are ranked randomly; and the proposed JointMMI, in which the subsets of APs are ranked in descending order of the joint mutual information gain. Two conclusions are drawn as follows:

1. JointMMI chooses the optimal subset of APs. JointMMI jointly and intelligently measures and ranks the importance of AP subsets via the MMI criterion. A more accurate measurement is taken by JointMMI because it considers the correlation among APs. This fact is supported by **Fig. 5**. Compared with the other AP selection schemes, JointMMI has the best accuracy. The accuracy within 2 m of JointMMI is 74.0%, whereas those of InfoGain, MaxMean, and Random are 69.9%, 69.0%, and 69.0%, respectively. In addition, the accuracy of JointMMI monotonically increases and reaches the highest value at the fastest rate. The optimal number of APs used for JointMMI is only 13, smaller than the other schemes. Hence, JointMMI identifies the best combination of APs with the minimum number of APs.

2. Initial AP selection by JointMMI improves the accuracy of LDE significantly. Compared with using the full APs, JointMMI for initial AP selection increases accuracy within 2 m from 69.0% to 74.0%. JointMMI also identifies and discards noisy or redundant APs effectively. As shown in **Fig. 5**, the accuracy of LDE does not necessarily increase with more APs used. When the number of APs used is larger than 13, APs with little location information introduce significant noise, so that the accuracy decreases gradually. Hence, initial AP selection enhances the performance of LDE by avoiding introducing noise.

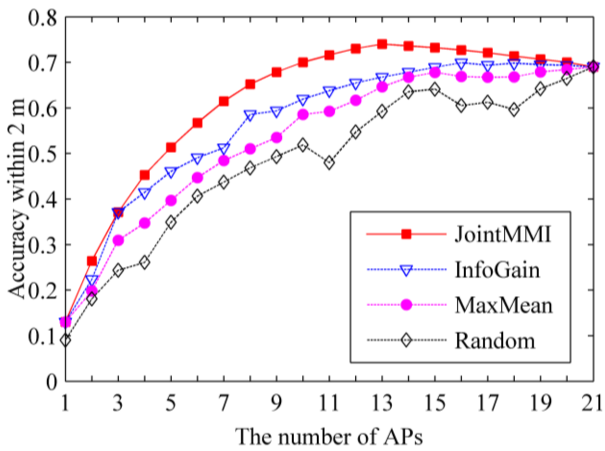


Fig. 5. Accuracy within 2 m versus the number of APs for different AP selection schemes in cluster 1

5.4 Feature Extraction Result

This section first compares the best testing accuracy of four feature extraction methods: LDE, KDDA, LDA, and PCA in cluster 1. All these methods deploy the optimal AP subset chosen by JointMMI. The original RSS signal space using the optimal AP subset and WKNN is also compared. **Fig. 6** shows the cumulative error distribution of different feature extraction methods with optimal feature dimensions. Accuracy of LDE within 2 m is 74.0%, whereas those of KDDA, LDA, PCA, and RSS are 75.3%, 63.5%, 59.4% and 57.0%, respectively. Accuracy of LDE within 3 m is 85.2%, whereas those of KDDA, LDA, PCA, and RSS are 84.9%, 76.0%, 72.7%, and 70.7%, respectively. Thus, LDE obtains comparable accuracy with KDDA, while significantly improving accuracy compared with LDA, PCA, and RSS.

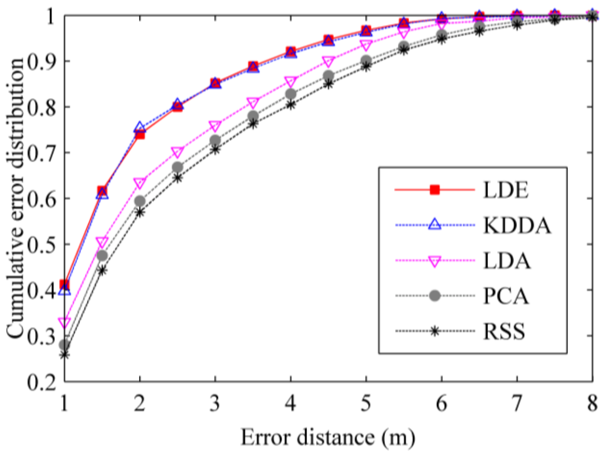


Fig. 6. Accuracy comparison among different feature extraction methods in cluster 1, with the optimal dimension $d = 3$ in LDE, $d = 6$ in KDDA, $d = 5$ in LDA, $d = 7$ in PCA, and $d = 13$ in RSS space

We further compare these feature extraction methods. RSS without feature extraction has the worst performance because considerable noise or redundancy are directly introduced, thus degrading accuracy. As shown in **Fig. 6**, PCA obtains little accuracy improvement compared with RSS because the most descriptive, rather than discriminative features, are extracted. In contrast, LDA performs better than PCA by extracting the features that best separate different locations. However, both PCA and LDA suffer from severe nonlinearity of RSS. They are both linear techniques and are incapable of extracting nonlinear features. In comparison, LDE is capable of capturing nonlinear features by considering RSS signal space as a manifold and preserving the locality geometry of the manifold. LDE explores neighbor relations and class information of RSS samples, thus enhancing the discrimination power of location features effectively. KDDA also extracts nonlinear features effectively by employing the kernel-based technique. However, KDDA requires excessive online computation cost, as will be shown in Section 5.5 and Table 1. Furthermore, without clustering analysis and AP selection, the accuracy of KDDA [17] (called KDDA-only) becomes significantly lower than that of the proposed LLDE-APS algorithm, as will be shown in Section 5.7.

Finally, we study the effect of feature dimensions on the accuracy of different feature extraction methods. Each feature component corresponds to an eigenvalue, which indicates the information quantity obtained. For PCA, the eigenvalue is proportional to the feature variance that each component retains. for KDDA and LDA, the eigenvalue is proportional to the discriminative information quantity of each component, whereas for LDE, the eigenvalue is proportional to the ratio of the local between-class distance to the local within-class distance of neighbor samples. LDE further considers the local geometry of RSS signal space to adapt to the nonlinearity of RSS. As shown in **Fig. 7**, setting $\beta^* = 85\%$ in (10) for LDE, the cumulative percentage of the eigenvalues of the first three components already exceeds β^*, whereas KDDA, LDA, and PCA need to retain the first six, five, and seven components to exceed β^*. As a result, the optimal feature vector extracted by LDE not only has enhanced discrimination power, but also has the minimum dimension among the compared methods.

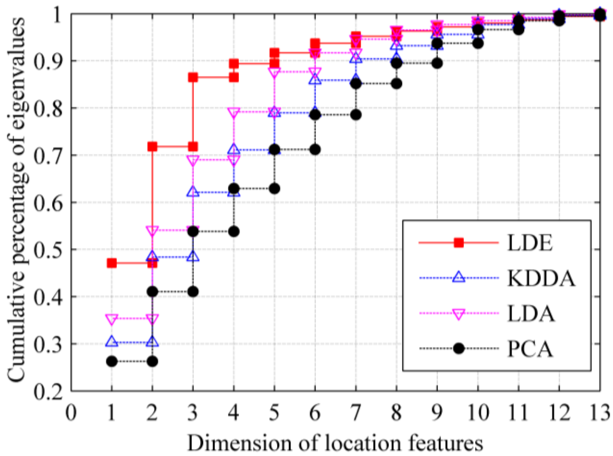


Fig. 7. Cumulative percentage of eigenvalues versus the dimension of location features in cluster 1

5.5 Online Computation Cost Reduction

As shown in **Fig. 8**, we compare the computation cost of LLDE-APS with the widely used WKNN, WKNN with clustering analysis (WKNN-CA), and WKNN-CA with AP selection (WKNN-CAAPS). We measure the computation cost by the average number of multiplications required for a single online positioning procedure. The computation cost over different clusters is not the same because the number of reference locations, the related optimal AP subset and LDE transformation are all different in each cluster.

Fig. 8 shows the number of multiplications in each cluster and the total average number for the entire positioning region. Compared with WKNN, the other approaches reduce the computation cost via clustering analysis. WKNN computes the Euclidean distances between the real-time RSS sample vector and RSS sample vectors of all reference locations over the entire region, whereas the other approaches compute the distances over the cluster of significantly smaller size. In our experiments, the enire region was first divided into five clusters, where the numbers of reference locations are 16, 12, 9, 8, and 10 from C1 to C5, respectively. Hence, WKNN-CA decreases the number of multiplications from 1385 to 470, 370, 295, 270, and 320, respectively. For WKNN-CAAPS, the computation cost is further reduced by discarding noisy APs. Rather than using all APs, the optimal number of APs selected is 13, 12, 14, 13 and 11 in each cluster, respectively. However, the proposed LLDE-APS requires the smallest computation cost. After feature extraction via LDE, the ultimate feature dimensions for each cluster are only 3, 3, 2, 2, and 3, respectively. As a result, LLDE-APS reduces the expected computation cost by 90.2%, 62.8%, and 37.5% against WKNN, WKNN-CA, and WKNN-CAAPS, respectively.

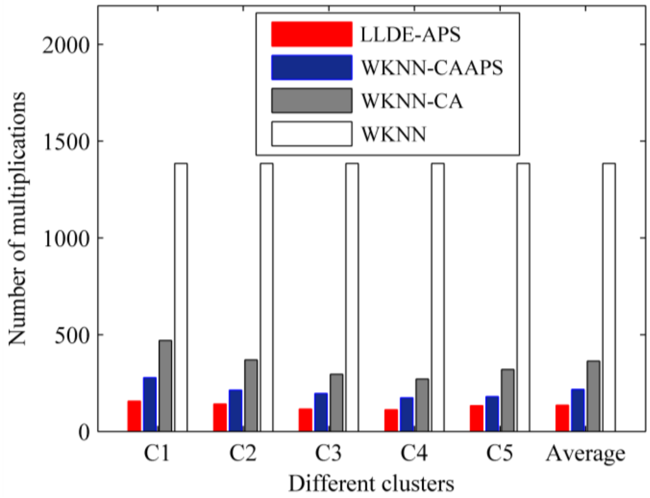


Fig. 8. Online computation cost comparison of LLDE-APS, WKNN, WKNN-CA, and WKNN-CAAPS

In addition, we also compare the online computation cost among different feature extraction methods, as seen in **Table 1**. We denote the optimal feature dimension of LDE, KDDA, LDA, and PCA as d_1, d_2, d_3, and d_4, respectively. Despite the higher accuracy, KDDA requires considerably more computation cost than the other methods because all M samples in the target region are involved. In contrast, LDE requires the smallest computation cost because it is linear and has the minimum optimal feature dimension. We take cluster 1 as an example, in which $n = 13$, $d_1 = 3$, $d_2 = 6$, $d_3 = 5$, $d_4 = 7$, and $M = 1600$. Hence, the number of multiplications for LDE, KDDA, LDA, and PCA are 39, 32000, 65 and 91, respectively.

Table 1. Online computation cost comparison among different feature extraction methods

Methods	Multiplication/division	Addition	Exp
LDE	nd_1	$(n-1)d_1$	0
KDDA	$(n+d_2+1)M$	$(2n+d_2-1)M-d_2$	M
LDA	nd_3	$(n-1)d_3$	0
PCA	nd_4	$(n-1)d_4$	0

5.6 Online Sensing Time Reduction

In Section 4, we show that the third approach saves energy by reducing online sensing time. If we can achieve the same accuracy with a sensing time that is shorter than that of previous methods, more sleep time is allowed during online positioning, which significantly reduces energy consumption. In this section, we verify that the proposed LLDE-APS can achieve higher accuracy while requiring a shorter sensing time.

Fig. 9 shows the accuracy within 2 m versus different online sensing time. We compare five methods: WKNN, ML, LDA, KDDA, and the proposed LLDE-APS. As shown in **Fig. 9**, using only 1 s sensing time (i.e., two real-time RSS samples), our method and KDDA already outperform the other methods using 5 s, indicating that we can save 4 s for each positioning procedure while achieving higher positioning accuracy. A smaller number of real-time samples may degrade positioning accuracy because of the increased uncertainty of RSS. However, our method can tolerate the uncertainty of RSS to some extent. As shown in Sections 5.3 and 5.4, LLDE-APS achieves the high-level accuracy with a significantly shorter sensing time by discarding noisy APs and maximizing the separability of location features.

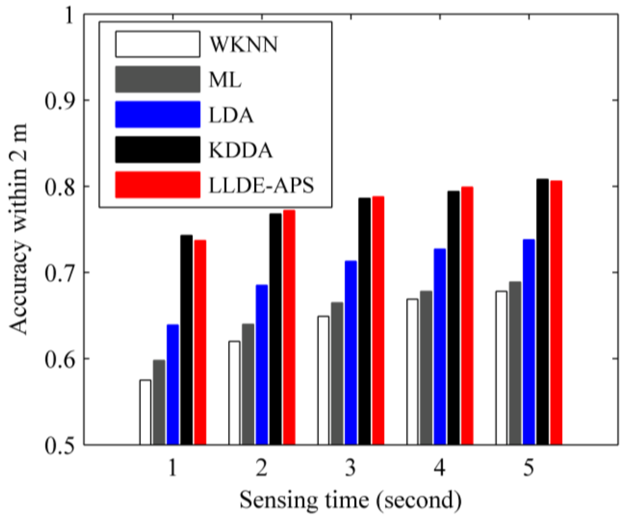


Fig. 9. Accuracy within 2 m versus online sensing time of LLDE-APS, KDDA, LDA, ML, and WKNN

5.7 Accuracy Comparison in Two Classic Environments

To validate the accuracy of the proposed LLDE-APS, we performed the experiments in two classic environments. First is the office environment shown in Section 5.1 and **Fig. 2**, which includes hallways, walls, and rooms. Second is the open hall of our building. The main difference between the two classic environments is that, in the former, most of the RSS signals propagate in a non-line-of-sight way, whereas no obstructions exist between APs and the target region in the latter. In the open-hall environment, we deploy 12 APs within a size of $20\ \mathrm{m} \times 16\ \mathrm{m}$. We collect 6,300 training samples at 63 reference locations and 5,000 testing samples at 50 test points, with 100 samples per location.

Table 2 and **3** compare positioning accuracy in the indoor office and open environments, respectively. All compared methods incorporate FCM and the proposed JointMMI, except for the KDDA-only method [17], which only deploys KDDA for feature extraction. The accuracy of WKNN is lowest because only the average RSS sample vectors at each reference location are explored. The accuracy of ANN is slightly higher than that of WKNN. In contrast, given ANN's disadvantage of local optimization, a number of large, unexpected positioning errors may arise. Thus, the standard deviation of error of ANN is largest. ML performs better than ANN and WKNN through the use of statistical RSS information. Moreover, LDA further improves accuracy by extracting discriminative features. The KDDA-only method performs worse than the LDA and PCA methods because the monolithic model over the entire region is suboptimal, and noisy APs may degrade the accuracy. However, LLDE-APS and KDDA significantly improve accuracy in both environments, despite the different clustering analysis, AP selection, and feature extraction models used in the two environments. In particular, LLDE-APS reduces the mean positioning error and standard deviation of error by 21.9% and 28.6% in the office environment while exhibiting 20.5% and 35.1% error improvement in the open environment, compared with WKNN and ML, respectively.

Table 2. Accuracy and positioning error (m) comparison in the indoor office environment

Methods	Mean error	Standard deviation	Accuracy within 2 m	Accuracy within 3 m
LLDE-APS	1.71	1.45	73.7%	84.9%
KDDA	1.68	1.49	74.3%	84.7%
KDDA-only	2.23	2.14	58.7%	72.4%
WKNN	2.37	2.12	57.5%	71.3%
ML	2.19	2.03	59.8%	74.2%
ANN	2.34	2.42	58.3%	71.2%
LDA	2.04	1.96	63.9%	76.5%
PCA	2.21	2.08	59.6%	73.0%

Table 3. Accuracy and and positioning error (m) comparison in the indoor open environment

Methods	Mean error	Standard deviation	Accuracy within 2 m	Accuracy within 3 m
LLDE-APS	1.67	1.35	74.3%	85.6%
KDDA	1.63	1.40	75.8%	85.2%
KDDA-only	2.21	2.15	59.7%	73.7%
WKNN	2.26	2.14	59.2%	73.4%
ML	2.10	2.08	61.6%	75.5%
ANN	2.24	2.38	59.1%	73.8%
LDA	1.98	1.95	64.7%	77.3%
PCA	2.16	2.02	61.3%	74.4%

6. Conclusion

In this paper, we propose a novel method for improving Wi-Fi positioning accuracy while reducing energy consumption of mobile devices. The contributions of the proposed LLDE-APS method are threefold. First, we develop a joint AP selection scheme to select the optimal subset of APs. Considering the correlation of the discriminant abilities of the APs, we measure the importance of APs more accurately than previous AP selection schemes. Furthermore, we show that initial AP selection may enhance the performance of subsequent feature extraction because noisy APs with little location information are discarded.

Second, we propose LDE for nonlinear discriminative feature extraction. Unlike previous linear feature extraction methods that cannot handle the nonlinearity of RSS, LDE adapts well to nonlinearity. Furthermore, the location features extracted by LDE have the strongest discrimination power and minimum feature dimension. Our previously proposed KDDA method obtains comparable accuracy with LDE. However, the online computation cost of LDE is considerably smaller than that of KDDA. As a result, our method not only improves accuracy significantly, but also reduces energy consumption greatly, owing to the much smaller computation cost and shorter sensing time required.

Third, we employ the FCM clustering algorithm to localize the positioning model. In each sub-region, more compact RSS patterns are obtained. Thus, more suitable AP selection and feature extraction solutions are achieved. In addition, computation cost is also reduced due to the smaller number of reference locations.

Experimental results in the classic office environment show that, compared with the WKNN and ML methods, LLDE-APS reduces the mean positioning error and standard deviation of error by 21.9% and 28.6%, respectively, while reducing the computation cost by 90.2%. Furthermore, our method achieves higher accuracy than other methods while saving 80% sensing time (4 s) for a single positioning process. Experiments in an indoor open environment also demonstrate the effectiveness of the proposed LLDE-APS method.

References

[1] K. Petrova, and B. Wang, "Location-Based services deployment and demand: a roadmap model," *Electron Commerce Research*, vol.11, pp.5–29, Nov.2011. Article (CrossRef Link)

[2] N. Chang, et al. "Robust indoor positioning using differential Wi-Fi access points," *IEEE Transactions on Consumer Electronics*, vol.56, no.3, pp.1860-1867, Aug.2010. Article (CrossRef Link)

[3] A. Kushki, et al., "Intelligent dynamic radio tracking in indoor wireless local area networks," *IEEE Transactions on Mobile Computing*, vol.9, no.3, pp.405-419, Mar.2010. Article (CrossRef Link)

[4] T. Hashem, and L. Kulik, "Don't trust anyone: Privacy protection for location based services," *Pervasive and Mobile Computing*, vol.7, no.1, pp.44-59, Mar.2011. Article (CrossRef Link)

[5] B. Delaney, "Reduced energy consumption and improved accuracy for distributed speech recognition in wireless environments," *PhD dissertation*, 2004.

[6] K. Y. Chen, et al. "Power efficient access point selection for indoor location estimation," *IEEE Transactions on Knowledge and Data Engineering*, vol.18, no.7, pp.877-888, Jul.2006. Article (CrossRef Link)

[7] M. Youssef, et al. "WLAN location determination via clustering and probability distributions," in *Proc. of First Int. Conf. on Pervasive Computing and Communications*, pp.143–150, Mar.2003.

[8] S. H. Fang, et al. "Location fingerprinting in a de-correlated space," *IEEE Transactions on Knowledge and Data Engineering*, vol.20, no.5, pp.685-691, May.2008. Article (CrossRef Link)

[9] S. H. Fang, and T. N. Lin, "Projection based location system via multiple discriminant analysis in wireless local area networks," *IEEE Transactions on Vehicular Technology*, vol.58, no.9, pp.5009-5019, Nov.2009. Article (CrossRef Link)

[10] J. J. Pan, et al., "Multidimensional vector regression for accurate and low cost location estimation in pervasive computing," *IEEE Transactions on Knowledge and Data Engineering*, vol.18, no.9, pp.1181–1193, Sept.2006. Article (CrossRef Link)

[11] H. T. Chen, et al. "Local discriminant embedding and its variants," in *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol.2, pp.846-853, Jun.2005.

[12] Z. Song, et al. "A survey on indoor positioning technologies," *Theoretical and Mathematical Foundations of Computer Science*, vol.164, pp.198-206, May.2011. Article (CrossRef Link)

[13] Y. Jie, Y. Qiang, and N. Lionel, "Learning adaptive temporal radio maps for signal strength based location estimation," *IEEE Transactions on Mobile Computing*, vol.7, no.7, pp.869-883, Jul.2008. Article (CrossRef Link)

[14] P. Bahl, and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *Proc. of IEEE INFOCOM*, pp. 775–784, Mar. 2000. Article (CrossRef Link)

[15] M. Borenovic, et al. "Space Partitioning Strategies for Indoor WLAN Positioning with Cascade-Connected ANN Structures," *International Journal of Neural Systems*, vol. 21, no. 1, pp. 1-15, Feb. 2011. Article (CrossRef Link)

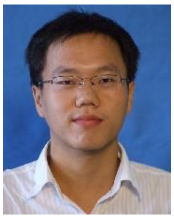
[16] H. Lim, et al. "Zero configuration, robust indoor localization: Theory and Experimentation," in *Proc. of IEEE INFOCOM 2006*, pp. 1-12, Apr. 2006.Article (CrossRef Link)

[17] Y. B. Xu, et al. "An indoor positioning algorithm with Kernel direct discriminant analysis," in *Proc. of IEEE Globecom 2010*, pp.1-5, Dec.2010. Article (CrossRef Link)

[18] K. Kaemarungsi, et al. "Properties of indoor received signal strength for WLAN location fingerprinting," in *Proc. of MOBIQUITOUS 2004*, pp.14-23, Aug.2004. Article (CrossRef Link)

[19] T. Kanungo, et al. "An efficient k-Means clustering algorithm: Analysis and implementation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.24, no.7, pp.881-892, Jul.2002. Article (CrossRef Link)

[20] A.P. Jardosh, et al. "GreenWLANs: On-Demand WLAN infrastructures," *Mobile Networks and Applications*, vol.14, no.6, pp.798-814, Dec.2009. Article (CrossRef Link)
[21] H. Peng, et al. "Feature selection based on mutual information criteria of max dependency, max-relevance, and min redundancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.27, no.8, pp.1226-1238, Aug.2005. Article (CrossRef Link)
[22] S. T. Roweis, and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol.290, no.5500, pp.2323-2326, Dec.2000. Article (CrossRef Link)
[23] C. Hou, C. Zhang, Y. Wu, and Y. Jiao, "Stable local dimensionality reduction approaches," *Pattern Recognition*, vol.42, no.9, pp.2054-2066, Sep.2009. Article (CrossRef Link)
[24] X. Ge, C. Cao, M. Jo, and M. Chen, "Energy efficiency modeling and analyzing based on multi-cell and multi-antenna cellular networks," *KSII Transactions on Internet and Information Systems*, vol.4, no.4, pp.560-574, Aug.2010. Article (CrossRef Link)
[25] I. Humar, et al. "Rethinking energy efficiency models of cellular networks with embodied energy," *IEEE Network Magazine*, vol.25, no.2, pp.40-49, Mar.2011.
[26] J. P. Ebert, et al. "A trace based approach for determining the energy consumption of a WLAN network interface," in *Proc. of European Wireless* 2002, pp.230-236, Feb.2002.
[27] M. Ergen, and P. Varaiya, "Decomposition of energy consumption in IEEE 802.11," in *Proc. of IEEE ICC 2007*, pp.403-408, Jun.2007. Article (CrossRef Link)
[28] C. W. You, et al. "Sensor enhanced mobility prediction for energy efficient localization," in *Proc. of IEEE SECON 2006*, pp.565-574, Sep.2006. Article (CrossRef Link)
[29] Z. Zhuang, et al. "Improving energy efficiency of location sensing on smartphones," in *Proc. of ACM MobiSys 2010*, pp.315-330, Aug.2010. Article (CrossRef Link)



Zhi-An Deng received the B.S. and M.S. degree in Information and Communication Engineering at Harbin Institute of Technology (HIT), Harbin, China, in 2006 and 2008, respectively. Currently, he is pursuing his Ph.D. degree in Information and Communication Engineering, at HIT. His main research interests include wireless positioning, machine learning and signal processing.

Yu-Bin Xu received the B.S. degree in radio measurement, in 1986, the M.S. degree in Electronics and Communication System, in 1993, and Ph.D. degree in Electronics and Communication System, in 2005, all from HIT, Harbin, China. Currently, he is a Professor and Vice Director of Communication Research Center. His research interests include wireless positioning, radio-wave propagation, collective and mobile communications.

Lin Ma received the B.S. degree, in 2003, the M.S. degree, in 2005, and Ph.D. degree, in 2009, all in Electronics and Communication System from HIT, Harbin, China. Currently, he is a Lecturer in Communication Research Center. His research interests include wireless positioning and tracking, UWB microwave imaging and collective communications.