International Journal of Future Generation Communication and Networking
Vol.7, No.5 (2014), pp. 191-206
http://dx.doi.org/10.14257/ijfgcn.2014.7.5.16

# An Improved Indoor Localization of WiFiBased on Support Vector Machines

YuFeng[1,a], JiangMinghua[1,b], LiangJing[1,c], QinXiao[2,d], HuMing[1,e], PengTao[1,f] and HuXinrong[1,g]

*[1]School of Electronic and Electrical Engineering*
*Wuhan Textile University, Wuhan 430200 China;*
*[2]Department of Computer Science and Software Engineering*
*Auburn University, Auburn AL 36849 America;*

*[a]yffly@163.com, [b]jmh@wtu.edu.cn, [c]shinelight@wtu.edu.cn, [d]xqin@eng.auburn.edu*
*[e]mhu@wtu.edu.cn, [f]pt@wtu.edu.cn, [g]hxr@wtu.edu.cn*

***Abstract***

*Indoor localization based on existing WiFi signal strength is becoming increasingly prevalent and ubiquitous. The user-based localization algorithm utilizes the information of the Received Signal Strength(RSS) from the surrounding access points(APs) to determine the user position. In this paper, focusing on the development of a user localization uses existing WiFi environment for its low cost and ease of deployment. We propose an indoor localization of WiFi based on support vector machines(ILW-SVM), and use the bilinear median interpolation method(BMIM) to reduce the calibration effort on creating fingerprint map while still retaining the accuracy of user localization. According to comparison of accuracy of three different kernel functions, choosing the radial basis function(RBF) as kernel function. In addition, we also propose improved ILW-SVM algorithm to solve the indoor localization that nearest neighbor points are not concentrated. At last, overall comparison of kNN, ILW-SVM and improved ILW-SVM in consideration of accuracy. Experimental results indicate that the proposed algorithm can effectively reduce the calibration effort and exhibit superior performance in terms of localization accuracy and stabilization.*

***Keywords:*** *WiFi, indoor localization, ILW-SVM, bilinear median interpolation method(BMIM), radial basis function(RBF), improved ILW-SVM*

## 1. Introduction

Recent advances in mobile computing technology and the increasing availability of WiFi networks have enabled more accurate localization in indoor environments where Global Positioning System(GPS) is less precise. GPS [1-5], has been in service for many years, and applied widely. Additionally, GPS receivers' fundamentals have been matured and very close to standardization, which enables flexible world-wide receivers manufacturing at low cost. In summary, the three main characteristics that make GPS attractive and successful are accuracy, universality, and low cost receivers. However, GPS is not reliable in densely-populated urban areas and almost unavailable in indoor areas. Many indoor localization algorithms have been proposed [6-9], and there has long been interested in the ability to determine the localization of the devices given only WiFi signal strength. In many proposed indoor localization algorithms, the signal strength measurements from different access points (APs) in a wireless local area network (WLAN) are used to infer the localization of indoor mobile devices [10-

ISSN: 2233-7857 IJFGCN

14]. Although there are different choices for the localization measurements methods such as time of arrival (TOA) [15-16], time difference of arrival(TDOA) [17], and angle of arrival(AOA) [18], received signal strength(RSS) [10, 12], is generally the feature of choice for indoor WiFi localization due to its low cost and wide availability without the need for additional hardware.

Indoor localization based on RSS has been extensively studied as an inexpensive solution in recent years. RSS can be easily obtained through a WiFi-integrated mobile device, without any additional hardware. Several RSS-based indoor localization and tracking algorithms have been proposed the utilization of APs' location information [6], [19-20], which may be distributed in the complex environment. Recently the most studied indoor localization algorithms based on RSS are trilateral localization and fingerprint map algorithm.

The trilateral localization algorithm uses the mainstream logarithmic distance path loss model [21-23]. The propagation model points out that whether in indoor or outdoor channel, the average received signal power decreases with the logarithm of distance. This model has been widely used. For any distance, the path loss is expressed as:

$$P_r = P_0 - 10nlog_{10}\frac{d}{d_0} + X_\sigma$$

where $d_0$ is the power measured at a reference distance, $d_0$ assumed to be of 1 m, d is the distance between mobile node and theaccess point, $n$ is the path loss exponent, $X_\sigma$denotes a Gaussian random variable with zero mean caused by shadowing. Thepower is measured at a reference distance, $\sigma$ depends on several factors: fast and slow fading, antenna gain, and transmittedpower.

The major challenge for accurate RSS-based localization comes from the variations of RSS due to the dynamic and unpredictable nature of radio channel, such as shadowing, multipath, the orientation of wireless device, and so on. Thus,instead of using a propagation model to describe the relationship between RSS and position, a fingerprint map is used to localize mobile devices. Some experiments show that the fingerprint map algorithm outperforms the trilateral localization algorithm.

At present, the most popular RSS-based localization technique is based on fingerprint map, [6, 8, 24-25]. When using the fingerprint map algorithm, a commonly used method of estimating user location is to find the nearest reference point, using the Euclidean distance in signal space. Such a technique generally contains of two phases: an offline training phase and an online localization phase. In the offline training phase, fingerprint collected from different APs is built accordingly to characterize the relationship between the RSS measurements and the locations in the deployment area. In the online localization phase, the reference points(RPs) are used to match the location of the mobile devices based on the observed real-time measurements. Various machine learning methods have been proposed to learn the RSS localization relationship and match the indoor localization estimation problem. Most of these methods aim to achieve high localization accuracy. A large number of RPs need to be collected for better performance, implying a great calibration effort.

In this paper, we propose the ILW-SVM based on theory of support vector machines (SVM) to process the signal strength measured by the mobile devices to match the fingerprint database. The SVM classifies data by finding the best hyperplane that separates all data points of one class from those of the other. The best hyperplane for an SVM means the one with the largest margin between the two classes. Margin means the maximal width of the parallel to the hyperplane that has no interior data points. Classification of the fingerprint database can obtain optimal nearest neighbor points.

We have applied the proposed algorithm based on Libsvm, which is a fast and efficient general SVM software package, designed and developed by the National Taiwan University Lin Chih-Jen. The implementation has shown that the proposed algorithm is able to estimate the location in real time with more precise localization.

The remainder of this paper is organized as follows: A brief description of related work is presented in Section 2. Section 3 describes the ILW-SVM algorithm in detail. Reconstruction of the fingerprint map by bilinear median interpolation method and the improved ILW-SVM theory are detailed in Section 4. The experiment results and evaluation through implementations are listed in Section 5. Finally, Section 6 concludes the paper.

## 2. Related Work

The SVM theory is a new and promising technique for data classification and regression. It is a tool for statistical analysis and machine learning, and it performs very well in many classification and regression applications. SVM has been used extensively for a wide range of applications in science, medicine, and engineering with excellent empirical performance [26].The theory of SVM is found in [27-28]. Support vector classification(SVC) of multiple classes and support vector regression(SVR) have been used successfully in localization estimation [13, 29-31]. Most of these methods aim to achieve high localization accuracy.

In [13], Zhili Wu tackled the GSM localization problems through machine learning approaches. The paper considers the direct application of the GSM system for location estimation that relies only on signal strength and there is no need for implementing new special equipment in the GSM network to fulfill the localization task. Considering the estimation as a missing value learning problem, they have tested the kernel-based support vector regression with different kernels.

In [29], Ling Pei presents an indoor navigation solution by combining physical motion recognition with wireless positioning. Twenty-seven simple features are extracted from the built-in accelerometers and magnetometers in a smartphone. Eight common motion states used during indoor navigation are detected by a Least Square Support Vector Machines (LS-SVM)classification algorithm, e.g., static, standing with hand swinging, walking normally while holding the phone in hand, walking normally with hand swinging, fast walking, U-turning, going up stairs, and going down stairs. Despite the fact that the motion recognition solution proposed in this paper provides correct motion recognition up to 95.53%. In the test cases, the motion behavior varies from person to person. More people will be involved to test the motion recognition algorithms and determine the most useful features for classification.

In [30], Duc A. Tran proposes LSVM, a novel solution that satisfies the requirements above, offers fast localization, and alleviates the border problem significantly. LSVM is also effective in networks with the existence of coverage holes or obstacles. LSVM localizes the network using the learning concept of Support Vector Machines. The LSVM algorithm is suitable for networks of larger scale because it is based on connectivity rather than direct signal strength. They define the kernel function and categorize the sensor nodes, a strategy to apply the classifiers, and a theoretical bound analysis on the localization error.

In [31], they presents a wireless network location determination system based on SVM, and this system has been applied in laboratory. The working procedure of the localization can be divided into two phases: offline phase and online phase. In the offline phase, storing the information into the database, the location inference system will retrieve the relevant information as the training data for SVM to train the model for inference purpose. Because SVM requires that each data instance is represented as a vector of real numbers, the client identity will he converted into numeric data when being stored as the training data for SVM. In the online phase, converting the received information as the input format of SVM, and SVM will use the model trained in the offline phase to infer the location of the wireless client according to these information.

In this paper, the propose ILW-SVM based on the fingerprint map for localization. In the online localization phase, classification of fingerprint database uses the theory oh SVM. When the same nearest neighbor points are concentrative, the ILW-SVM can achieve high accurate. Besides sometimes the same nearest neighbor points are dispersive, the localization accuracy is unstable. In order to solve this problem, the improved ILW-SVM algorithm is proposed, which provides a novel framework for WiFi localization with high accuracy and stabilization. The proposed algorithms are exhibited in the simulation study.

Focusing on how to significantly reduce the offline calibration effort while still achieving high accuracy in localization estimation. Many researchers try to find the best method to overcome the problem of calibration effort. The accuracy goes down with reduced calibration data, so they have to make trade-off between accuracy and effort. In [6], they try to reconstruct the fingerprint map form measurements at only a small number of fingerprints, and in [8], a hybrid generative/discriminative semi-supervised learning algorithm is proposed that utilizes a large number of unlabeled points to supplement a small number of labeled points.

In this paper, we propose the bilinear median interpolation method (BMIM) to solve this problem, the BMIM is based on fingerprint map with initial cell grid. Empirical experiments show that proposed algorithm gain high accuracy while improving the fingerprint map grid size with BMIM, which reduces the effort on offline calibration effort. Furthermore it is simpler than [6, 8], algorithm.

## 3. Indoor Localization System

The proposed ILW-SVM algorithm based on the traditional fingerprint method is developed. As mentioned, the proposed algorithm consists of two phases: an offline training phase, and an online localization phase. In the offline training phase, RSS readings are collected on a grid of RPs, the received signal strength information is interpolated with the BMIM, and eventually stored in fingerprint database. The online localization phase includes: measuring real-time RSS with mobile devices, finding out the facing orientation of mobile device by compass sensor, which is owned by most smart mobile devices, and then estimating localization with the ILW-SVM algorithm.

### 3.1. Offline Training Phase

#### 3.1.1 Building Fingerprint Database

During an offine training phase, the time samples of RSS readings are collected at known locations, referred to as the reference points, by pointing the mobile device to different orientations, (*i.e.*, north, south, east, west). The raw set of RSS time samples collected from AP i at RP j with orientation o is denoted as $\{r_{i,j}^{(o)}(t), t=1,\cdots,q, q>1\}$, with the q representing the total number of time samples collected. Then, the average of the RSS time samples is computed and stored in a database, known as the radio map. Such fingerprint map can be represented by $\Psi^{(o)}$:

$$\Psi^{(o)} = \begin{pmatrix} \gamma_{1,1}^{(o)} & \gamma_{1,2}^{(o)} & \cdots & \gamma_{1,N}^{(o)} \\ \gamma_{2,1}^{(o)} & \gamma_{2,2}^{(o)} & \cdots & \gamma_{2,N}^{(o)} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{L,1}^{(o)} & \gamma_{L,2}^{(o)} & \cdots & \gamma_{L,N}^{(o)} \end{pmatrix}$$

Where $\{\gamma_{i,j}^{(o)} = \frac{1}{q}\sum_{t=1}^{q}\gamma_{i,j}^{(o)}(t)$ isthe average of RSS readings(in dBm scale) over time domain from $AP_i$ at $RP_j$ withorientation o, for $\mathrm{i} = 1,2,\cdots,\mathrm{L}, \mathrm{j} = 1,2,\cdots,\mathrm{N}$, and $\mathrm{o} \in \sigma = \{\mathrm{north}, \mathrm{south}, \mathrm{east}, \mathrm{west}\}$.L is the total number of APs thatcan be detected, and N is the number of RPs. The columnsof $\Omega^{(o)}$,radio map vectors, represent the RSS readings ateach RP with a particular orientation o, which can be referredto as $\gamma_j^{(o)} = [\gamma_{1,j}^{(o)}, \gamma_{2,j}^{(o)}, \cdots, \gamma_{L,j}^{(o)}]^T, j = 1,2,\cdots,N$,where thesuperscript T denotes transposition.

### 3.1.2. The Bilinear Median Interpolation Method (BMIM)

In order to reduce the offline calibration effort while still achieving high accuracy in localization estimation. We interpolate and determine the value in the blank spot areas. As expected, the accuracy goes down with reduced calibration data, the BMIMis proposed to overcome this problem, the BMIM not only reduces the calibration effort, but also improves the accuracy as being proved in Section 5.5.

The main idea of the BMIM is a linear interpolation in two directions. In the paper, using interpolation in the vertical at first, and then interpolation in the horizontal next. There are four fingerprint maps representing the four directions of RSS. The interpolation function $\Delta^{(o)}$ is denoted as:

$$\Delta^{(o)} = \begin{pmatrix} 0 & \frac{\gamma_{1,1}^{(o)} + \gamma_{1,2}^{(o)}}{2} & 0 & \cdots & 0 \\ \frac{\gamma_{1,1}^{(o)} + \gamma_{2,1}^{(o)}}{2} & \frac{\gamma_{1,1}^{(o)} + \gamma_{1,2}^{(o)} + \gamma_{2,1}^{(o)} + \gamma_{2,2}^{(o)}}{2} & \frac{\gamma_{1,2}^{(o)} + \gamma_{2,2}^{(o)}}{2} & \cdots & \frac{\gamma_{1,N}^{(o)} + \gamma_{2,N}^{(o)}}{2} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{\gamma_{L-1,1}^{(o)} + \gamma_{L,1}^{(o)}}{2} & \frac{\gamma_{L-1,1}^{(o)} + \gamma_{L-1,2}^{(o)} + \gamma_{L,1}^{(o)} + \gamma_{L,2}^{(o)}}{2} & \frac{\gamma_{L-1,2}^{(o)} + \gamma_{L,2}^{(o)}}{2} & \ddots & \frac{\gamma_{L-1,N}^{(o)} + \gamma_{L,N}^{(o)}}{2} \\ 0 & \frac{\gamma_{L,1}^{(o)} + \gamma_{L,2}^{(o)}}{2} & 0 & \cdots & 0 \end{pmatrix}$$

Since the initial fingerprint map with the initial grid size is not accurate enough, after interpolation by the BMIM, we just get the ultimate fingerprint map, and the concise representation is as follows:

$$\Omega^{(o)} = \begin{pmatrix} M_{1,1}^{(o)} & M_{1,2}^{(o)} & \cdots & M_{1,2N-1}^{(o)} \\ M_{2,1}^{(o)} & M_{2,2}^{(o)} & \cdots & M_{2,2N-1}^{(o)} \\ \vdots & \vdots & \ddots & \vdots \\ M_{2L-1,1}^{(o)} & M_{2L-1,2}^{(o)} & \cdots & M_{L,2N-1}^{(o)} \end{pmatrix}$$

Where $\mathrm{M}_{i,j}^{(o)}$ is the RSS of $RP_{i,j}$ with the orientation o, for $\mathrm{i} = 1,2,\cdots,2\mathrm{L}-1, \mathrm{j} = 1,2,\cdots,2\mathrm{N}-1$, and $\mathrm{o} \in \sigma = \{\mathrm{north}, \mathrm{south}, \mathrm{east}, \mathrm{west}\}$.

### 3.2. Online Localization Phase

The actual localization of the mobile devices take place in the online localization phase. During the online localization phase,the RSS measurement vector is denoted as:

$$\omega_r^0 = [\omega_{1,r}^o, \cdots, \omega_{L,r}^o]^T$$

Where$\{\omega_k^0, k = 1,2,\cdots,L\}$, is collected by the mobile device in o orientation. We only select the maximum RSS value as localization evaluation standard, and determine the orientation of the mobile device by compass sensor, and match data base using ILW-SVM algorithm. According to ILW-SVM algorithm, the database should be divided into two categories by optimal hyperplane. The algorithm will be described in details in the following Section 4.1.

In majority situation, there are some same value reference points in the fingerprint map when the fingerprint database is large. It is unstable for localization by the ILW-SVM algorithm. So we propose the improved ILW-SVM algorithm, and collecting the real-time RSS $\omega_r$ by mobile devices. Since of$\omega_r$is easily influenced by environmental factors, before matching fingerprint map, we expand the real-time data as$\omega_r \pm \xi$, separate $\omega_r \pm \xi$ form fingerprint database $\Omega$ by ILW-SVM algorithm, and then obtain the nearest neighbor data $\omega_i$, and $\Omega \cap \omega_i \in (\omega_r \pm \xi)$, where $\xi$ is signal fluctuation factor.

At last, on the basis of$\omega_i$, using centroid algorithm to achieve more accuracy localization, the detailed introduction in Section 4.2.

## 4. Increasing Localization Accuracy by Using ILW-SVM

In order to enhance accuracy of indoor localization, we empirically study the effect of reducing the calibration effort on the fingerprint map by interpolating new map with smaller grid size, which uses the BMIM. And according to collected real-time RSS to match fingerprint database by the ILW-SVM algorithm. In the following content, the ILW-SVM algorithm will be introduced in detail.

### 4.1. Support Vector Classification Model

In this paper, based on the real-time measured RSS, the proximate real-time measured data is classified from the fingerprint database. Assume that the fingerprint database is defined as $\Omega$, the measured real-time RSS date is defined as $\omega_r^o$ , we separate $\omega_r^o$form fingerprint database $\Omega$ by SVM.

Support Vector Machines is an efficient method to solve classification problem. In the case of finite data space, steps taken typically in SVM are as follows:

If we carry out nonlinear transformation characteristics of x, the new feature is described as$z = \varphi(x)$, and the new featurespace structure of SVM decision function is:

$$f(x) = \mathrm{sgn}(\omega^{\varphi} \cdot z + b) = \mathrm{sgn}(\sum_{i=1}^{n} a_i y_i \big(\varphi(x_i) \cdot \varphi(x)\big) + b)$$

Define the kernel function $K(x,z) = \exp\left(-\frac{||x-z||^2}{\delta^2}\right) = \varphi(x_i) \cdot \varphi(x_j)$

Support vector machines in transformation space can be written as:

$$f(x) = \mathrm{sgn}(\sum_{i=1}^{n} a_i y_i K(x_i \cdot x) + b)$$

Among them, the coefficient $\alpha$ is the optimal solution of the problem:

$$max_a Q(a) = \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i,j=1}^{n} a_i a_j y_i y_j \mathrm{K}(x_i \cdot \mathrm{x}))$$

$$\text{s.t.} \quad \sum_{i=1}^{n} y_i a_i = 0$$

$$0 \leq a_i \leq C, i = 1, \cdots, n$$

b is obtained through support vector formula:

$$y_i\left(\sum_{i=1}^{n} a_i \mathrm{K}(x_i \cdot \mathrm{x}) + b\right) - 1 = 0$$

From computational perspective, regardless of how much transform space dimension of$\varphi(\mathrm{x})$ is, the linear support vectormachines can be carried out in the original space by kernel function, which avoids the computation of high dimensional space, and complexity of kernel function computing and inner product does not increase substantially.

### 4.2. Localization Based on Improved ILW-SVM

The major challenge for accurate RSS-based localization comes from the variations of RSS due to the dynamic and unpredictable nature of radio channel, such as shadowing, multipath, the orientation of wireless device. In order to develop accuracy of the real-time RSS, we expand the scope of RSS value, and then classify them with ILW-SVM.

After classification by SVM, the fingerprint database is divided into two parts, each represents the nearest neighbors of the real-time measured data in the whole database and is described as $\omega_i = R_1,\ R_2,\ \cdots, R_i$, There is one-to-one correspondence between each fingerprint and location. Online measurement of RSS is $\omega_r$.

Given a calculation space distance function in the signal Dist(; ), the process of the nearest neighbor method can bedescribed as selecting a shortest distance signal reference fingerprint j:

$$\mathrm{Dist}(\omega_i, R_j) \leq \mathrm{Dist}(\omega_i, R_i), \forall\ \mathrm{i} \neq \mathrm{j}$$

While the signal distance can be described as $\mathrm{L}_p$ by a weighted distance:

$$\mathrm{L}_p = \frac{1}{v_a}(\sum_{i=1}^{N_a}\frac{1}{\beta_i}||\gamma_i - N_i||^p)^{\frac{1}{p}}$$

Where $v_a$ is the dimension of search space (or the number of APs system is deployed), $\beta_i$ is weighted factor, $p$ is normparameter, while $p = 2$, this distance is called the Euclidean distance.

According to the nearest neighbor data$\omega_i, i = 1,2, \cdots, n$is obtained form $\Omega$, we can further improve the localization accuracy by centroid algorithm , and the localization formula is as follows:

$$\begin{cases} \hat{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \end{cases}$$

Where $(x_i, y_i)$ is the coordinate of i-st reference point that is one of the nearest neighbors of the real-time measured data.$(\hat{x}, \hat{y})$is the final result of localization.

## 5. Experiment Results and Evaluation

In this section, we evaluate the effectiveness of the proposed algorithm. To demonstrate the advantage of the ILW-SVM algorithm, and compare it with the based-kNN(k nearest neighbor) fingerprint localization method, the accuracy descriptions of which can be found in Section 5.6. The purpose of comparing ILW-SVM with the kNN is to evaluate the performance has been improved in this experimental platform. The goal of comparing ILW-SVM with improved ILW-SVM is to experimentally reveal that improved ILW-SVM can obtain higher accuracy than ILW-SVM. In Section 5.5, we also analyze the error of BMIM. In addition, the numerical results presented in Section 5.5 and 5.6 validate the correctness of improved ILW-SVM.

### 5.1. Experimental Platform Building

We perform the experiment in the 2-D simulated area. The total dimension of area is 100 ×100 meters. Assume that thearea is equipped with WiFi environment. To structure the fingerprint map, the environment is modeled as a space of 25 grid locations, each grid is $4 \times 4$ meters(besides the edge of grids). This size is chosen as starting cell grid size, then interpolateit to make more detailed map with the BMIM. Since we think that training data on too many locations on offline phase is impractical, attempt to calibrate in few locations and interpolate all the other data on grid points with our model. In the experiment environment, we collect RSS in the center of each grid, and consider them as RPs.

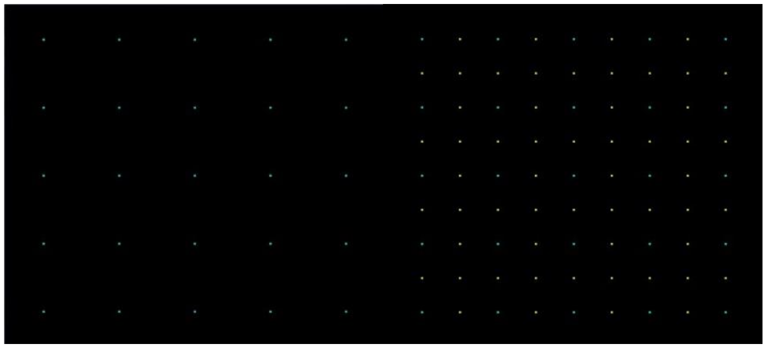
In this paper, we classify the fingerprint database by Libsvm, the general steps of the Libsvm is:

- In accordance with the requirements of Libsvm software package format the data sets.
- Scale the data simply.
- Select the kernel function.
- Use the cross validation to choose the best parameters.
- Obtain the support vector machine model by training the whole training data sets with the optimum parameters.
- Test and predict with the obtained model.

### 5.2. Reconstruction Fingerprint Map by BMIM

In order to reduce the number of RPs, we propose the BMIM to reconstruct the fingerprint map, the detailed algorithm is in Section 3.1.2. After interpolation, the reference points from 25 to 81 are shown in Figure 1.



(a) 25 RFs (b) 81 RFs after interpolation

**Figure 1. Reconstruction of the Fingerprint Map by Interpolation**

### 5.3. Performance of ILW-SVM algorithm

The concept of the ILW-SVM algorithm, which is originally developed for classification problems, is the use of hyperplanes to define decision boundaries, separates data into different classes. The ILW-SVM is able to handle both simple and linear classification tasks, as well as more complex, i.e., nonlinear classification problem. The idea of ILW-SVM is to map the original data from the input space to a high dimensional, or even infinite-dimensional, the classification problem becomes simpler in the feature space.

How to change the original training data into a linear space in the feature space, is the basic function of kernel function in SVM. So the kernel function is the key point of the ILW-SVM algorithm. The basic kernel functions are studied in the following three forms:

- Polynomial SVM of degree d: $\mathrm{K}(\mathrm{x},\mathrm{z}) = (\zeta + x^T \mathrm{z})^d, \zeta \geq 0$
- Radial basis function (RBF): $\mathrm{K}(\mathrm{x},\mathrm{z}) = \exp(-\frac{\left||x-z|\right|^2}{\delta^2})$
- sigmoid function: $\mathrm{K}(\mathrm{x},\mathrm{z}) = \tanh(k_1 x^T \mathrm{z} + k_2)^d$

Where K(,) is positive definite for all $\delta$ values in the RBF kernel case, and$\zeta > 0$ values is in the polynomial case, butnot for all possible choices of $k_1$, $k_2$ in the sigmoid function case. Such kernel transform the original data into a higherdimensional feature space, where they are separable by a hyperplane in this feature space.

According to the pattern recognition theory, the low dimensional space linearly inseparable pattern through nonlinear mapping into a high dimensional feature space, it may realize linearly separable, but if we use this technique for classification or regression in high dimensional space directly, there exists nonlinear mapping function form, parameters, the dimension of the feature space and other problems, and the largest disorder is the curse of dimensionality in high dimensional feature space. The technology of kernel function can solve this problem efficiently.

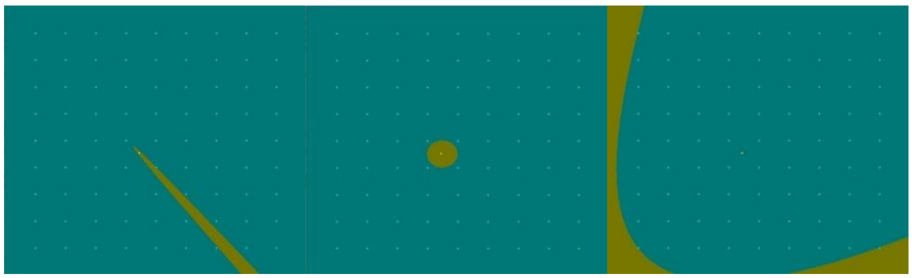
As a modular method, kernel function can be divided into two parts-kernel function design and algorithm design. See the details below:

- Collect, collate and standardize the sample data.
- Select or construct kernel function.
- Transform the sample data into kernel matrix with kernel function. This step is equivalent to input of data into a high dimensional feature space with a nonlinear mapping function.
- Apply various linear algorithm based on kernel matrix in the feature space.

- Obtain a nonlinear model of the input space.

#### 5.3.1. Performance of different kernel functions

Classification of data uses three different kernel functions by Libsvm as the following Figure 2.



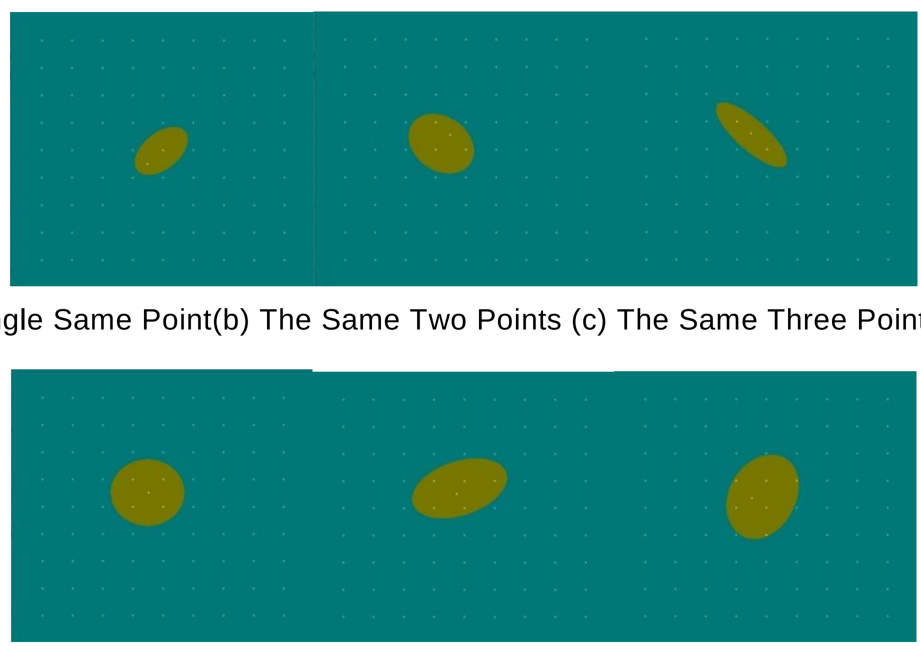
(a) Polynomial function (b) Radial Basis Function(RBF) (c) Sigmoid Function

**Figure 2. Performance Different Kernel Function**

We can get that the radial basis function (RBF) achieve the best classification form the above figure easily. So we choose the RBF as the classification of function in this study.

#### 5.3.2. Performance of RBF

Since there exist multiple same RSS in the fingerprint database when the database is large enough, we will discuss more of the same value problem later. In Figure 3, it is easy to achieve high accuracy when the same value reference points are concentrative.



(a) Single Same Point(b) The Same Two Points (c) The Same Three Points

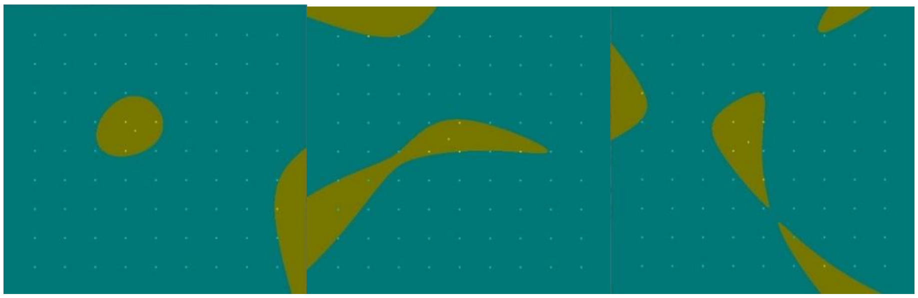
(d) The Same Four Points (e) The Same Five Points (f) The Same Six Points

**Figure 3. Performance of RBF**

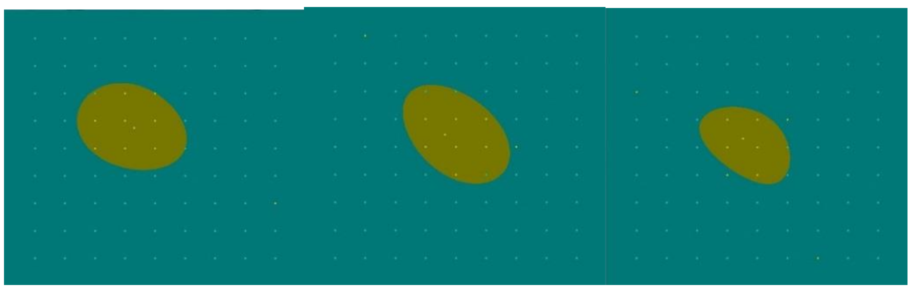
 

### 5.4. Improved ILW-SVM: Expanded Real-time Measurement of RSS

In addition to the above situation that all the same value RPs around the testing point, sometimes the same value finger print points are not around the testing point as is shown in Figure 4 (a)(b)(c). Base on Section 4.2. We classify the database into two parts with improved ILW-SVM algorithm as following Figure 4 (e)(f)(g).



(a) The Same Four Points (b) The Same Five Points (c) The Same Six Points



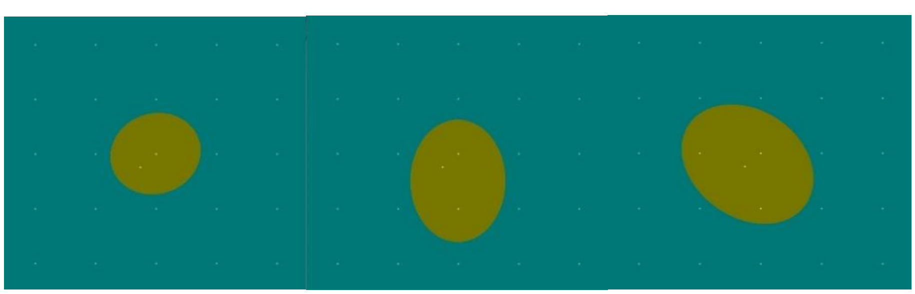
(d) The Same Four Points (e) The Same Five Points (f) The Same Six Points

**Figure 4. Improved SVM Modified Real-time Measurement of RSS**

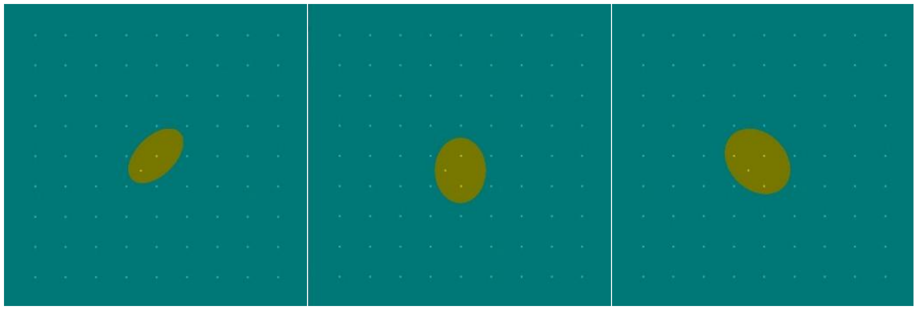
### 5.5. Error Analysis of the BMIM

In the paper, we use the BMIM algorithm to reconstruct the fingerprint map as introducing in Section 3.1.2. According to Section 5.3.2, we know that it is most accurate using RBF than other kernel functions, and based on RBF, the accuracy of comparison is made before(Figure 5(a),(b),(c)) and after(Figure 5(e),(f),(g)) using the BMIM, and the results are shown in Figure 5.

The localization area using BMIM is smaller than the area without using BMIM seen obviously from Figure 5. So it iseasy to demonstrate the BMIM which is very important to improve the localization accuracy.



(a) Original Fingerprint-1 (b) Original Fingerprint-2 (c) Original Fingerprint-3



(d) Interpolation of Fingerprint-1 (e) Interpolation of Fingerprint-2 (f) Interpolation of Fingerprint-3

**Figure 5. Error Analysis of the BMIM**

### 5.6. Comparison of Localization Error

The goal of conducting this experiment is to compare the proposed improved ILW-SVM algorithm against the ILW-SVM and kNN algorithm to understand the accuracy of the three algorithms in different number of same value RPs. We compare the accuracy with different algorithms in same data. The same value of different numbers of reference points are using for localization in Figure 6.

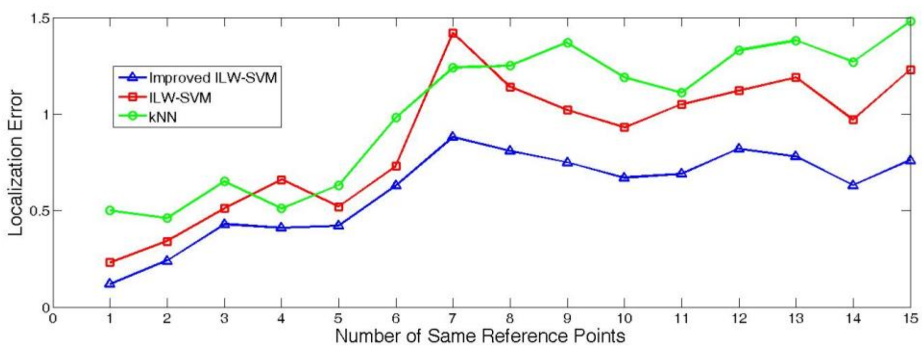



**Figure 6. Comparison of Localization Error**

We can learn that, based on Figure 6, the improved ILW-SVM algorithm achieves the highest accuracy of three algorithms. The average error of ILW-SVM is lower than kNN, in addition to the situation of five and seven same value of reference points. Sometimes the nearest neighbor points are not concentrated, the stabilization of the ILW-SVM is bad, to solve above situation, the improved ILW-SVM is proposed, and it can solve this problem well as knowing in Figure 6 distinctly. Comprehensive performance comparison, the improved ILW-SVM achieves the best performance.

## 6. Conclusion

In this paper, we propose the ILW-SVM algorithm for indoor localization. First of all, according to the fingerprint map and real-time measured RSS, classification of data uses the theory of SVM that finds out the best hyperplane, separating the nearest neighbor points from all data points, and then localization uses centroid algorithm. In the ILW-SVM experience, three kernel functions are used in classification of fingerprint database, RBF obtain highest accuracy of them, and RBF is used in the ILW-SVM. In addition, we use the BMIM to reduce the offline calibration effort for the indoor localization. To solve the situation of the same value of reference points are not concentrative, the improved ILW-SVM algorithm has been presented. Experimental results show that our proposed algorithm can greatly reduce the calibration effort but still achieves superior localization accuracy.

In the future, we plan to improve this algorithm to apply it in more fingerprint database, and selecting parameter optimization of kernel function is also our key work. Moreover, we intend to test the validity of proposed algorithm in other complex building with many rooms and floors.

## Acknowledgements

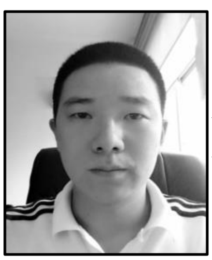
This work was supported by National Natural Science Foundation of China under grant project No.61232004 and No.61103085.

## References

[1] H. Akcan, V. Kriakov, H. Bronnimann and A. Delis, "Gps-free node localization in mobile wireless sensor networks in Proceedings of the 5th ACM international workshop on Data engineering for wireless and mobile access", ACM, **(2006)**, pp. 35-42.
[2] N. Bulusu, J. Heidemann and D. Estrin, "Gps-less low-cost outdoor localization for very small devices", Personal Communications, IEEE, vol. 7, no. 5, **(2000),** pp. 28-34.
[3] A. S. Wong, T. K. Woo, A.-L. Lee and X. Xiao, *et.al*., "An agps-based elderly tracking system", First International Conference, IEEE, **(2009),** pp. 100-105.
[4] A. Khodabandeh, A. Amiri-Simkooei and M. Sharifi, "Gps position time-series analysis based on asymptotic normalityof m-estimation," Journal of Geodesy, vol. 86, no. 1, **(2012),** pp. 15-33.
[5] S. Choi, K. Kim, Y. Kim and J. Lee, "Outdoor precision position estimation system using multi-gps receivers," In Intelligent Robotics and Applications, Springer, **(2013)**, pp. 66-76.
[6] C. Feng, W. Au, S. Valaee and Z. Tan, "Received signal strength based indoor positioning using compressive sensing", **(2011)**.
[7] J. Koo and H. Cha, "Unsupervised locating of wifi access points using smartphones", **(2012)**.
[8] R. Ouyang, A. Wong, C. Lea and M. Chiang, "Indoor location estimation with reduced calibration exploiting unlabeleddata via hybrid generative/discriminative learning", **(2011)**.
[9] A. S. Paul and E. A. Wan, "Wi-fi based indoor localization and tracking using sigma-point kalman filtering methods in Position, Location and Navigation Symposium", **(2008)**, pp. 646-659.
[10] Y. Kim, Y. Chon and H. Cha, "Smartphone-based collaborative and autonomous radio fingerprinting Systems, Man and Cybernetics, Part C: Applications and Reviews", vol. 42, no. 1, **(2012)**, pp. 112-122.
[11] A. Kushki, K. N. Plataniotis and A. N. Venetsanopoulos, "Kernel-based positioning in wireless local area networks,"Mobile Computing, IEEE Transactions on, vol. 6, no. 6, **(2007),** pp. 689-705.
[12] V. Moghtadaiee and A. G. Dempster, "Wifi fingerprinting signal strength error modeling for short distances in IndoorPositioning and Indoor Navigation (IPIN)", IEEE, **(2012)**, pp. 1-6.
[13] Z.-l. Wu, C.-H. Li, J. K.-Y. Ng and K. R. Leung, "Location estimation via support vector regression", Mobile Computing, IEEE Transactions, vol. 6, no. 3, **(2007)**, pp. 311-321.
[14] J. Yin, Q. Yang and L. M. Ni, "Learning adaptive temporal radio maps for signal-strength-based location estimation Mobile Computing", vol. 7, no. 7, **(2008)**, pp. 869-883.
[15] T.-Y. Chin, "Rss-based doa indoor location estimation system and method", **(2013)** January.
[16] C. Yang, Y. Huang and X. Zhu, "Hybrid tdoa/aoa method for indoor positioning systems in Location Technologies", **(2007)**, pp. 1-5.
[17] M. Bocquet, C. Loyez and A. Benlarbi-Delai, "Using enhanced-tdoa measurement for indoor positioning Microwaveand Wireless Components Letters", IEEE, vol. 15, no. 10, **(2005)**, pp. 612-614.
[18] C. Wong, R. Klukas and G. G. Messier, "Using wlan infrastructure for angle-of-arrival indoor

user location in Vehicular Technology Conference", **(2008)**, pp. 1-5.
[19] A. Kushki, K. Plataniotis and A. Venetsanopoulos, "Kernel-based positioning in wireless local area networks", IEEETrans, **(2007)** June.
[20] Y. Chen, Q. Yang, J. Yin and X. Chai, "Power-efficient access-point selection for indoor location estimation," IEEETrans, Knowledge and Data Eng, **(2006)** July.
[21] Z. Sen, Z. Xiaojia and Y. Bin, "Trilateral-centroid-localization in wireless sensor network with mobile anchor node," Computer Measurement & Control, vol. 7, **(2009)**, p. 068.
[22] C. Xiaoyan, H. Ting and G. Mingjie, "Improved dv-hop algorithm based on self-localization of wsn," in InformationScience and Engineering (ICISE), 2009 1st International Conference", IEEE, **(2009)**, pp. 4026-4029.
[23] J. Zhang, W. Li, D. Cui, X. Sun and F. Zhou, "Study on improved dv-hop node localization algorithm in wireless sensornetwork in Industrial Electronics and Applications (ICIEA) the 5th IEEE Conference", IEEE**, (2010)**, pp. 1855-1858.
[24] Y. Luo, O. Hoeber and Y. Chen, "Enhancing wi-fi fingerprinting for indoor positioning using human-centric collaborativefeedback", vol. 3, no. 1, **(2013)**, pp. 1-23.
[25] C. Sapumohotti, M. Y. Alias and S. W. Tan, "Wilocsim: Simulation testbed for wlanlocation fingerprinting systems", Progress in Electromagnetics Research B, vol. 46, **(2013)**, pp. 1-22.
[26] N. Cristianini and J. Shawe-Taylor, "An introduction to support vector machines and other kernel-based learning methods", Cambridge University press, **(2000)**.
[27] V. Kecman, "Learning and soft computing: support vector machines, neural networks, and fuzzy logic models", MIT press, **(2001)**.
[28] V. Vapnik, "The nature of statistical learning theory", Springer, **(2000)**.
[29] L. Pei, J. Liu, R. Guinness, Y. Chen, *et.al*., "Using ls-svm based motion recognition for smart phone indoor wireless positioning", Sensors, vol. 12, no. 5, **(2012)**, pp. 6155-6175.
[30] D. A. Tran and T. Nguyen, "Localization in wireless sensor networks based on support vector machines Parallel and Distributed Systems", IEEE Transactions on, vol. 19, no. 7, **(2008)**, pp. 981-994.
[31] C.-L. Wu, L.-C. Fu and F.-L. Lian, "Wlan location determination in e-home via support vector classification in Networking, sensing and control", IEEE international conference, IEEE, vol. 2, **(2004)**, pp. 1026-1031.

## Authors



**Yu Feng** received the BE and B. Admin degrees in Wuhan Textile University, China, in 2008. Currently, he is working toward the MS degree in School of Electronic and Electrical Engineering from the Wuhan Textile University in 2012. His current research interests include wireless localization technologies, internet of thing and machine learning.

**Jiang Minghua**, received the MS degree and PhD degrees in Computer Science from the Huazhong University of Science and Technology in 1999 and 2011 respectively. Currently, he is a Professor in School of Mathematics and Computer Science, Wuhan Textile University. His research interests include storage systems, service computing, distributed computing and parallel computing.

**Liang Jing** received the BS and MS degrees in Computer Science from the Huazhong University of Science and Technology. His research interests include embedded System and artificial intelligence.

**Qin Xiao** received the BS and MS degrees in computer science from Huazhong University of Science and Technology, China, in 1996 and 1999, respectively, and the PhD degree in computer science from the University of Nebraska, Lincoln, in 2004. He is a currently an assistant professor of computer science at Auburn University. Prior to joining Auburn University in 2007, he was with the New Mexico Institute of Mining and Technology for three years. He has been on the program committees of various international conferences, including IEEE Cluster, IPCCC, and ICPP. From 2000 to 2001, he was a subject area editor of the IEEE Distributed Systems Online. His research interests include parallel and distributed systems, real-time computing, storage systems, and fault tolerance. He is a member of the IEEE and the IEEE Computer Society. He received a US National Science Foundation Computing Processes and Artifacts (CPA)Award in 2007.

**Hu Ming** received the BS and PhD degrees in computer science from Huazhong University of Science and Technology, China, and the MS degree in Chinese people's Liberation Army Command College. His research interests include storage systems, service computing, distributed computing and parallel computing.

**Peng Tao** received the MS degree and PhDdegree in Computer Science from the Huazhong University of Science and Technology in 2006 and 2011 separately. He is currently an associate professor in School of Mathematics and Computer Science, Wuhan Textile University. His research interests include Data Mining, pattern recognition, network security.

**Hu Xinrong** received the BS degree from Wuhan Textile University in 1994, and the MS degree form Wuhan University of Technology in 1997, and the PhD degree from Huazhong University of Science and Technology in 2008. Her research interests include computer image processing, virtual reality technology, and computer vision.