

基于改进支持向量回归的室内定位算法*

姚英彪,毛伟勇,姚瑞丽,严军荣,冯 维

(杭州电子科技大学通信工程学院 杭州 310018)

摘要:针对室内无线局域网环境中无线信号不稳定,以及传统支持向量回归定位算法在构建位置坐标与信号强度时的单输出导致位置坐标信息之间的关联性降低的问题,提出一种基于改进支持向量回归的室内定位方法。该算法首先对采集到的接收信号强度(RSS)指纹进行对数处理使其更符合正态分布,然后采用高斯滤波过滤掉小概率的指纹值之后构建指纹数据库;其次,为了降低单独构建 x 与 y 坐标模型的误差,提高二维位置信息与RSS之间的关联性,在训练阶段增加训练一个校正坐标 $z = x \cdot y$;最后,根据加权反 K 近邻的方法得到最优的位置坐标。实验结果表明,提出的算法可以减少室内复杂环境带来的噪声干扰,与传统的支持向量回归定位算法相比有更高的定位精度。

关键词:室内定位;过滤;校正坐标;加权反 K 近邻

中图分类号:TN912 TH70 文献标识码:A 国家标准学科分类代码:510.99

Indoor positioning algorithm based on improved support vector regression

Yao Yingbiao, Mao Weiyong, Yao Ruili, Yan Junrong, Feng Wei

(School of Communication Engineering, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: The wireless signal of wireless local area network is unstable in the indoor environment, and the traditional support vector regression (SVR) based positioning method may lead to the reduction of the correlation between the position coordinates and signal strength. Thus, this paper proposes an improved support vector regression (ISVR) based indoor positioning method. Firstly, the logarithmic processing is conducted on the received signal strength (RSS) to make it more consistent with the normal distribution, and then the Gaussian filter is used to filter the small probability of fingerprints before building the fingerprint database. Secondly, in order to reduce the error of constructing X and Y coordinate model separately, a calibration coordinate $z = x \cdot y$ is trained at the training stage, which can improve the correlation relationship between RSS and X - Y position information. Finally, the optimal position coordinates are obtained by weighted inverse K -nearest neighbor (WIKNN) method. The experimental results show that the proposed algorithm can reduce the noise caused by the complicated environment in the room, and has higher positioning accuracy than the traditional support vector regression algorithm.

Keywords: indoor positioning; filter; calibration coordinates; K -nearest neighbor (WIKNN)

0 引 言

随着物联网应用的不断扩大和深入,基于位置的服务(location-based services, LBS)有了越来越广泛的应用,人们对位置信息及时、快速、准确获取的需求也越来越强烈。传统的GPS和蜂窝网技术在室外能达到较高的定

位精度,但在室内环境下,由于建筑物对信号的影响,定位能力受到较大的限制^[1]。无线局域网(wireless local area network, WLAN)由于具有布线简单、价格低廉等特性,被广泛密集部署在室内环境中为用户提供通信服务,基于WLAN的室内定位已经成为当前室内定位的研究热点^[2-5]。

目前WLAN室内定位方法中研究最普遍的是位置

指纹定位方法^[6-7]。位置指纹定位方法一般分为离线阶段和在线阶段两个阶段。离线阶段,首先根据室内环境选择一些参考点,然后在每个参考点上采集周围 WLAN 接入点(access point, AP)的接收信号强度(received signal strength, RSS),以此构建 RSS 指纹,最后利用室内地图确定参考点的坐标,建立由参考点坐标和 RSS 指纹共同组成的位置指纹数据库。在线阶段,用户通过移动终端采集定位位置点的 RSS 指纹,然后将采集到的 RSS 指纹与位置指纹数据库里面的 RSS 指纹匹配,根据匹配结果确定用户当前的位置坐标。针对指纹定位,大量学者对其进行了深入研究。文献[8]提出采用 K 近邻(K nearest neighbor, KNN)算法匹配位置指纹数据库,即用最接近的 K 个 RSS 指纹点的质心作为测试点的估计位置。加权 K 近邻(weighted KNN, WKNN)是 KNN 的改进,核心是将选取的指纹点位置的加权值作为测试点的估计位置。KNN 和 WKNN 容易实现,但是匹配范围大,精度不是很高。文献[9]中提出采用概率统计模型,在位置指纹数据库中存储拟合的 RSS 高斯分布,并在此系统中提出聚类分块的概念。相比于直接利用 RSS 指纹进行计算,既降低系统计算的复杂度,也提高定位精度。文献[10]提出利用支持向量回归算法进行位置指纹匹配,由于该方法具有较好的学习泛化性和全局最优性,因而定位精度得到提高。文献[11]将支持向量机应用到基站定位中,通过设计新的核函数减少丢失信号带来的干扰,提高了支持向量机在复杂环境中的定位精度。文献[12]提出融合卡尔曼滤波和支持向量回归的算法,通过利用前后定位位置之间的关系修正移动轨迹,改善定位效果。文献[13]提出利用核典型相关分析将信号强度空间和物理位置空间的数据转化成各自的典型特征,从而建立信号强度与位置这两个空间之间的非线性映射,并利用核方法将样本间的两两相似性最大关联起来,提高定位的准确性。

在实际的室内环境中,无线信号在传播过程中受到墙壁、门窗以及人员走动等影响,导致无线信号在传播过程中产生显著的多径现象和阴影效应,因而在同一个位置处接收无线 AP 的 RSS 往往表现出复杂的时变特性。文献[14]采用两个传感器装置,分别放置于人体的前后,以此减少室内的多径效应以及人体对视距信号的影响;文献[15]提出利用主成分分析进行定位特征提取,降低环境因素造成的噪声干扰以及定位特征分量之间的相关性,从而提高定位精度;文献[16]提出离线阶段的过滤规则以及在线阶段的多次测量方法,有效的过滤掉噪声信号,以此提高定位精度。

针对室内无线局域网环境中无线信号不稳定,传统支持向量回归(support vector regression, SVR)定位算法在构建位置坐标与信号强度时的单输出导致位置坐标信

息之间的关联性降低的问题,本文提出一种改进支持向量回归(improved support vector regression, ISVR)的室内 WLAN 指纹定位算法。ISVR 通过增加 RSS 指纹预处理、以及模型训练时增加校正坐标,并通过加权反 K 近邻(weighted inverse KNN, WIKNN)得到最终定位位置来改进传统 SVR 的定位精度。具体来说,本文的主要贡献如下。

1) 采用对数高斯滤波的方法对 RSS 指纹进行预处理。对采集到的 RSS 指纹进行对数处理使其更符合正态分布,接着采用高斯滤波方法过滤掉小概率的指纹值构建指纹数据库,提高位置与 RSS 指纹的关联程度。

2) 在传统的支持向量回归训练预测模型训练二维坐标 x 和 y 的基础上,增加训练一个校正坐标 $z = x \cdot y$,进一步提高二维位置与 RSS 指纹之间的关联性。

3) 为了得到更准确的定位坐标,根据计算得到的结果利用 WIKNN 选择得到最优的定位坐标。

本文算法使用的 RSS 数据来自真实 WLAN 环境,通过实验验证了所采用预处理的有效性及增加校正坐标对算法的改进。实验结果表明,本文提出的 ISVR 算法与传统的 SVR 算法相比平均定位误差由 2.11 降低到 1.47 m。

1 基于 ISVR 的室内定位算法

1.1 信号预处理

对于指纹定位而言,指纹库的样本质量对最终的定位效果至关重要,理想环境下的无线信号比较有规律可找,但是由于室内的复杂环境,信号时常会因为环境因素波动很大。由于采集到的 AP 的 RSS 会有噪声的存在,因此对于信号的预处理是有必要。

偏度是描述数据分布对称性的统计量,通过判断偏度大小来确定信号的分布。当偏度为负数时,说明具有左偏分布。当偏度为正数时是右偏分布。从离线参考点中任取 20 个点,观察其偏度分布。从图 1 中可以看出大部分点的偏度较大,说明室内环境的信号强度有着比较严重的左偏现象。如果此时把数据取反再取对数后做偏度统计分析,从图 1 中可以看出偏度值明显降低,更符合正态分布。此时可以根据信号的对数模型建立数学模型进行分析。

设 RSS 的对数正态分布的概率密度函数表示为:

$$f(x) = \frac{1}{-x\sigma\sqrt{2\pi}} \times \exp\left(-\frac{(\ln(-x) - \mu)^2}{2\sigma^2}\right) \quad (1)$$

同时可以求出 $-x$ 对数的均值和方差:

$$\mu = \frac{\sum_{k=1}^n \ln(-x_k)}{n} \quad (2)$$

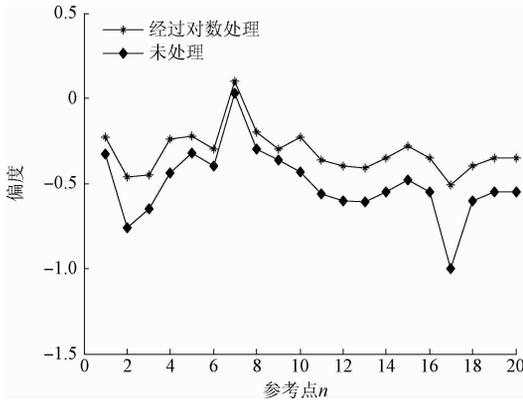


图1 原始数据经过对数处理和未处理的比较
Fig.1 Comparison of the logarithmic processing data and original data

$$\sigma^2 = \frac{\sum_{k=1}^n (\ln(-x_k) - \mu)^2}{n} \quad (3)$$

根据正态分布的特性,如图2所示,均值 μ 是其位置参数,描述正态分布的集中趋势位置。一般而言,RSS在均值附近取值的概率大,有较多的RSS分布在均值附近,而离均值越远的位置取值概率小,有较少的RSS分布在离均值较远的两端。为减少小概率的RSS的影响,本文提出利用高斯滤波对RSS进行预处理。

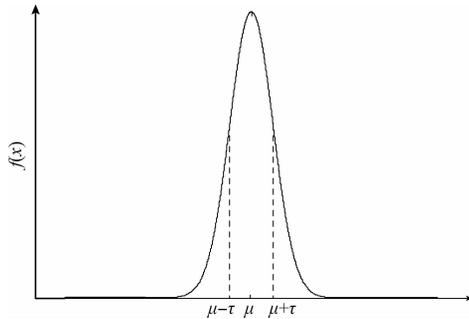


图2 正态分布的特性
Fig.2 Characteristics of normal distribution

具体来说,通过设置阈值 λ ,它表示RSS在 $[\mu - \tau, \mu + \tau]$ 的概率,如式(4)所示, τ 为图2中的范围参数。

$$P(\mu - \tau < \text{RSS} < \mu + \tau) = \int_{\mu - \tau}^{\mu + \tau} f(x) dx = \theta \quad (4)$$

根据设置的阈值 λ ,可以求出范围参数 τ 。也就是说,当RSS在 $[\mu - \tau, \mu + \tau]$ 范围内时,认为该RSS受噪声影响较小,是一个有效值,对最终的定位会产生有益的效果,保留;当RSS不在此范围内时,则认为该RSS受噪声影响较大,对最终定位产生负面效果,作过滤处理。

同时,考虑到在室内环境中各个AP对定位的贡献程度不同,计算定位区域中各个AP的贡献程度,贡献程

度主要由缺失率决定。缺失率定义为 $D(i, j)$,表示在参考点 i 处采集不到第 j 个AP信号强度的概率:

$$D(i, j) = \frac{k}{m} \quad (5)$$

式中: m 为在参考点 i 处采集的总次数, k 为在这 m 次采集中没有采集到第 j 个AP的次数。

缺失率高,相当于是在某参考点处采集RSS指纹的过程中,某一个AP的RSS信息比较不容易被采集到,可以认为是距离较远或者干扰严重影响该AP在参考点 i 处提供定位有效信息,可靠程度低;反之,则认为缺失率低,可靠程度高。

基于缺失率的预处理机制如下。

首先,根据在每个参考点 i 采集到的各个AP的RSS,经过对数高斯滤波预处理,计算第 j 个AP的缺失率 $D(i, j)$ 。

其次,设置阈值 θ ,如果缺失率 $D(i, j) > \theta$,说明在参考点 i 处,第 j 个AP的RSS缺失率过大,可靠性不大,作过滤处理;如果缺失率 $D(i, j)$ 小于阈值 θ ,说明在这个参考点处该AP对应的RSS缺失率低,信号值可利用。需要指出的是, $D(i, j) < \theta$ 时,仍然可能在参考点 i 处采集不到第 j 个AP的RSS,此时对于缺失的RSS,可以用该AP信号值的平均值来代替。

1.2 SVR算法模型建立

基于SVR算法的定位首先需要构建指纹数据库。指纹数据库是在待定位区域按照一定的间隔选取若干参考点,然后在每个参考点位置采集 T 次全部AP的信号强度。记参考点 i 的位置坐标为 $\mathbf{P}_i = (x_i, y_i)$,参考点 i 的RSS指纹向量为 $\mathbf{R}_i = [R_{i,1} \ \cdots \ R_{i,j} \ \cdots \ R_{i,N}]$, $R_{i,j}$ 表示在参考点 i 采集到的第 j 个AP的信号强度。原始RSS指纹经过1.1节所述的预处理,得到新的指纹数据库 $[\mathbf{P}_i, \mathbf{R}'_i]$, $\mathbf{R}'_i \in \mathbf{R}^N$ 。然后利用SVR算法^[17]构建指纹点的信号强度和其坐标之间的非线性关系。首先通过一个非线性映射 $\Phi: \mathbf{R}^N \rightarrow \mathbf{F}$ 将输入空间 \mathbf{R}^N 映射到高维的特征空间 \mathbf{F} 中,然后在 \mathbf{F} 中构建位置坐标 \mathbf{P} 与RSS指纹的最优回归函数:

$$\mathbf{P} = \mathbf{W}^T \cdot \Phi(\mathbf{R}) + b \quad (6)$$

式中: \mathbf{W} 为权重系数, $\mathbf{W} \in \mathbf{F}$, b 为偏置系数。为了最终确定式(6)中的参数 \mathbf{W} 和 b ,可以依据结构风险化最小原则,求解如下凸二次规划问题:

$$\min \frac{1}{2} \|\mathbf{W}\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \quad (7)$$

并满足:

$$\begin{cases} \mathbf{P} - \mathbf{W}^T - b \leq \varepsilon + \xi_i \\ \mathbf{W}^T - b - \mathbf{P} \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (8)$$

式中: C 为惩罚参数, ξ_i, ξ_i^* 为松弛变量。

式(8)的凸二次规划问题的拉格朗日多项式为:

$$L = \frac{1}{2} \|\mathbf{W}\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) - \sum_{i=1}^N (\eta_i \xi_i + \eta_i^* \xi_i^*) -$$

$$\sum_{i=1}^N \alpha_i (\varepsilon + \xi_i - \mathbf{P} + \mathbf{W}^T \cdot \Phi(\mathbf{R}) + b) - \sum_{i=1}^N \alpha_i' (\varepsilon + \xi_i + \mathbf{P} - \mathbf{W}^T \cdot \Phi(\mathbf{R}) - b) \quad (9)$$

式中: $\xi_i, \xi_i^*, \alpha_i, \alpha_i^*$ 为拉格朗日乘子, 且需要满足 $\xi_i^*, \alpha_i^* \geq 0$ 。根据最优解存在的条件, 可得:

$$\begin{cases} \frac{\partial L}{\partial \mathbf{W}} = \mathbf{W} - \sum_{i=1}^N (\alpha_i - \alpha_i^*) \Phi(\mathbf{R}_i) = 0 \\ \frac{\partial L}{\partial b} = \sum_{i=1}^N (\alpha_i^* - \alpha_i) = 0 \\ \frac{\partial L}{\partial \xi_i^*} = C - \alpha_i^* - \eta_i^* = 0 \end{cases} \quad (10)$$

通过对偶优化求解, 可得:

$$\max - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i - \alpha_i^*) (\alpha_j - \alpha_j^*) k(\mathbf{R}_i, \mathbf{R}_j) - \varepsilon \sum_{i=1}^N (\alpha_i + \alpha_i^*) + \sum_{i=1}^N P_i (\alpha_i - \alpha_i^*) \quad (11)$$

并满足 $\sum_{i=1}^N (\alpha_i - \alpha_i^*) = 0, \alpha_i, \alpha_i^* \in [0, C]$ 。

由此, 可以得到:

$$\mathbf{W} = \sum_{i=1}^N (\alpha_i - \alpha_i^*) \Phi(\mathbf{R}_i) \quad (12)$$

$$\mathbf{P} = \sum_{i=1}^N (\alpha_i - \alpha_i^*) k(\mathbf{R}_i, \mathbf{R}) + b \quad (13)$$

其中, 选择的核函数为高斯核函数:

$$k(\mathbf{R}_i, \mathbf{R}) = \exp\left(-\frac{1}{2\delta^2} \|\mathbf{R}_i - \mathbf{R}\|^2\right) \quad (14)$$

基于 SVR 算法的定位, 这里只考虑室内平面定位, 需要输出用户的二维坐标 (x, y) , 也就是需要用到多输出的预测, 而传统 SVR 算法的输出都是一维的, 此时可以通过多个单输出代替多输出, 来实现多输出 SVR 算法。但是室内坐标系是一个二维相关联的信息, 所以单独训练每个坐标得到的二维坐标向量会在一定程度上降低训练模型的准确度。

为了降低单独构建 x 与 y 坐标模型的误差, 提高二维位置信息与信号强度之间的关联性, 在训练阶段增加训练一个校正坐标 $z = x \cdot y$, z 同时包含 x 和 y 的信息。在训练阶段, 针对 x, y, z 分别构建 3 个训练模型, 因此得到的预测结果中会有 3 组坐标值, 分别是 (x', y') 、 $(x', z'/x')$ 、 $(z'/y', y')$ 。通过校正坐标得到的 3 组坐标值后, 如果得到的坐标值相同, 则可以直接输出结果; 反之, 如果得到的坐标值不同, 那么需要在这 3 组坐标中选择一组坐标输出。针对这个问题, 本文在 1.3 节采用加权反 K 近邻的方法进行选择。

本文对测试点位置的回归预测利用了支持向量机工具箱 LIBSVM^[18], 进行网格搜索获取最优惩罚参数 C 和核函数宽度 δ 的值。

1.3 WIKNN 算法

WKNN 算法通过计算指纹库中各个位置的信号强度向量与预测位置的信号强度向量之间的欧氏距离, 选择欧氏距离最近的 k 个坐标, 然后给所选择的 k 个坐标乘上一个加权系数后求和, 得到预测位置的坐标。而 WIKNN 算法是 WKNN 算法的逆运算, 根据 1.2 节得到的预测位置坐标组合, 通过计算指纹库中各个位置的坐标与预测位置坐标之间的欧氏距离, 选择欧氏距离最近的 k 个位置坐标, 然后给所选择的 k 个坐标对应的信号强度乘上一个加权系数后再求和, 得到预测位置所对应的信号强度向量。具体的实现步骤如下。

1) 假如选择的坐标是 (x', y') , 首先利用式 (15) 计算 (x', y') 和指纹数据库中的坐标 (X_i, Y_i) 之间的距离 d_i , 然后按从小到大的顺序选择前 k 个坐标所对应的信号强度向量 \mathbf{R}_i , 最后利用式 (16) 对这 k 个信号强度向量进行加权, 得到坐标 (x', y') 对应的信号强度向量 \mathbf{R} 。

$$d_i = \sqrt{(X_i - x_i)^2 + (Y_i - y_i)^2} \quad (15)$$

$$\mathbf{R} = (r_{ss_1}, \dots, r_{ss_N}) = \frac{1}{\sum_{i=1}^N \omega_i} \sum_{i=1}^N \omega_i \mathbf{R}_i =$$

$$\frac{1}{\sum_{i=1}^N \omega_i} \sum_{i=1}^N \omega_i (R_{ss_1}, R_{ss_2}, \dots, R_{ss_N}) \quad (16)$$

式中: $\omega_i = \frac{1/d_i}{\sum_{i=1}^k 1/d_i}$ 。

2) 根据在 (x', y') 处采集得到的信号强度向量 \mathbf{R}' , 利用式 (17) 计算信号强度向量 \mathbf{R} 与 \mathbf{R}' 之间的欧氏距离 $D(1)$ 。

$$D = \sqrt{\left(\sum_{j=1}^N |\mathbf{R}' - \mathbf{R}|^2\right)} \quad (17)$$

3) 对 $(x', z'/x')$ 、 $(z'/y', y')$ 的计算重复步骤 1) 和 2), 分别得到欧式距离 $D(2), D(3)$ 。

4) 将上述得到的欧氏距离 $D(1), D(2), D(3)$ 排序, 选择欧氏距离最小的 $D(i)$ 所对应的坐标, 并以此坐标作为最终的位置坐标。

1.4 定位算法步骤

基于改进 SVR 算法定位的总体框架如图 3 所示。定位主要步骤如下, 其中步骤 1) ~ 3) 为离线训练阶段, 步骤 4) ~ 6) 为在线定位阶段。

1) 在定位区域选定参考位置点, 在每个参考位置点处采集来自不同 AP 的一定数量的 RSS 指纹值, 组成 RSS 位置指纹数据库;

2) 对采集到的原始 RSS 指纹进行预处理, 得到新的 RSS 位置指纹数据库;

3) 利用 ISVR 构建参考点的信号强度与其位置坐标之间的非线性关系, 得到离线训练模型;

4) 实时采集信号强度 RSS, 同时对在线 RSS 进行预处理, 得到在线阶段的信号强度向量;

5) 利用离线阶段得到的训练模型进行定位, 得到定位坐标组合;

6) 根据定位坐标组合, 应用 WIKNN 的方法得到用户最终的位置。

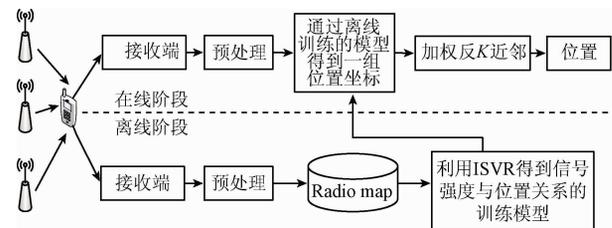


图3 基于 ISVR 的室内定位总体框架

Fig.3 Framework of ISVR-based indoor location

1.5 算法复杂度分析

在一个实际的基于指纹的定位系统中, 离线阶段的算法复杂度与系统定位实时性无关, 因此本文在这里主要分析本文提出算法的在线阶段的时间复杂度。设 N 为 AP 的个数, T 为每个位置参考点处每个 AP 采集的次数。ISVR 主要包括预处理、支持向量回归和加权反 K 近邻 3 个步骤。预处理的时间复杂度为 $O(T \cdot N)$ 。在一个大型的实际定位场景中, 只会需要增加 N , T 是可以固定, 因此预处理的时间复杂度可以近似为 $O(N)$ 。支持向量回归实质上将预处理后的 RSS 代入式(13)中, 其时间复杂度为 $O(N)$ 。加权反 K 近邻步骤中, 在指纹数据库中得到定位点的最近邻参考点及计算加权合成 RSS 都是常数项执行次数, 因而其时间复杂度为 $O(1)$ 。综上所述, ISVR 在线阶段的时间复杂度为 $O(N)$ 。

需要指出的是, 传统 SVR 也有支持向量回归过程, 因而其时间复杂度也为 $O(N)$ 。因此, 本文提出的 ISVR 和传统 SVR 的时间复杂度数量级相同。但是, 由于本文提出的 ISVR 相比传统 SVR 增加了更多的步骤, 因而其实际运行时间必然高于传统 SVR。

2 实验结果与分析

2.1 实验环境的建立

为验证该文提出的 ISVR 算法的室内定位效果, 在杭州电子科技大学第一教学楼 2 楼进行了实验, 局部平面图如图 4 所示, 长 45 m、宽 18 m, 面积约为 810 m²。定位区域包括过道和办公室, 办公室内有办公设备等办公用品, 而且人员走动频繁, 干扰复杂多变, 因此这个环境下的数据比较有代表性。

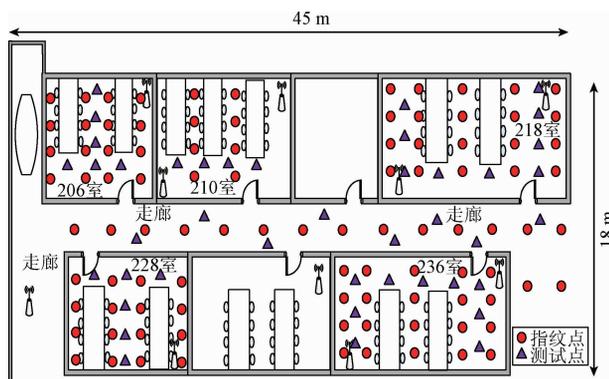


图4 实验环境

Fig.4 Experimental environment

实验软件平台是基于 Android 开发的 WiFi Scanner 采集各个 AP 的信号值, 硬件平台是 HUAWEI P8 和 10 个型号为 MERCURY MW325R 的无线路由设备。整个实验区域选取了 96 个指纹点和 42 个测试点, 其中 218 室选取 20 个指纹点和 9 个测试点, 236 室选取 20 个指纹点和 9 个测试点, 228 室选取了 16 个指纹点和 6 个测试点, 206 室选取了 16 个指纹点和 6 个测试点, 走廊区域选取了 16 个指纹点和 8 个测试点, 210 室选取了 8 个指纹点和 4 个测试点。指纹点在这 6 个区域中均匀分布, 相邻指纹点间距 1~3 m, 测试点的位置为随机选取, 在每个指纹点处能接收 5~8 个 AP 的 RSS 信号。在一天的不同时段进行 RSS 信号的采集, 在指纹点上采集 60 次作为一组, 每次采集间隔 1 s, 总共采集 3 组。在测试点上连续采集 180 次进行定位。

在实验中分别实现了传统的 SVR 算法^[10]、 K 近邻算法^[8]以及改进的 SVR 算法 ISVR。KNN 和 WKNN 中近邻数为 5, SVR 的核函数选择泛化能力较强的高斯核函数。定义定位误差 err 和平均定位误差 $avgerr$ 分别为:

$$err_i = \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2}$$

$$avgerr = \frac{1}{n} \sum_{i=1}^n err_i \quad (18)$$

式中: x_i, y_i 是第 i 个测试点的真实坐标, x'_i, y'_i 是第 i 个测试点的估计位置坐标, n 为测试点的总数量。

2.2 阈值参数 λ 对定位算法的影响

参数 λ 是一个在 (0, 1) 范围内的常数。图 5 所示为算法的平均定位误差随 λ 值的变化情况, 在 λ 趋近于 1 时, 由于没有对原始数据进行过滤, 因此定位误差较大; 随着 λ 的减小, 对样本的过滤逐渐严格, 可以有效过滤掉噪声信号, 定位误差变小, 在 $\lambda = 0.6$ 时, 平均误差达到最小; λ 小于 0.6 趋近于 0 时, 对样本的过滤过于严格导致部分有用信息丢失, 因此平均误差会增大。因此本文中选择的阈值 λ 为 0.6。

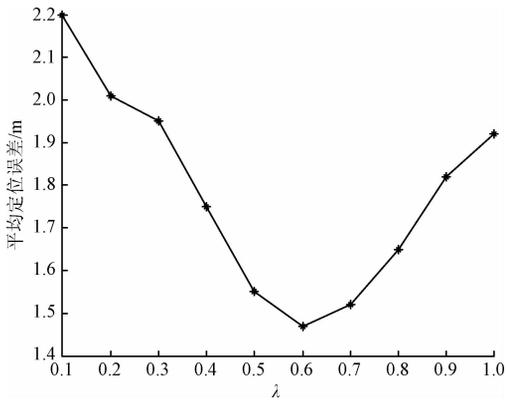


图5 平均定位误差随参数 λ 变化情况
 Fig.5 The average positioning error varies with the parameter λ

2.3 阈值参数 θ 对定位算法的影响

从图6中可以看出,当 θ 趋近于0时,只保留了样本中少部分的有用信息,此时的定位误差最大;随着 θ 的增大,样本中的信息变多,定位误差变小,在 $\theta=0.8$ 时,平均误差达到最小; $\theta>0.8$ 时,会有一部分噪声信息无法过滤,因此平均误差会增大。因此本文中选择的阈值 $\theta=0.8$ 。

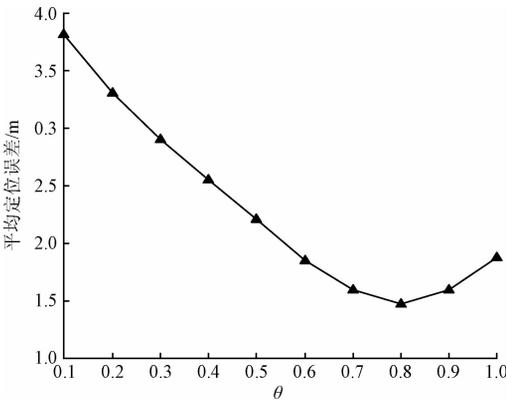


图6 平均定位误差随参数 θ 变化情况
 Fig.6 The average positioning error varies with the parameter θ

2.4 预处理及增加校正坐标对算法的影响

图7所示为不同的参考点采样样本数情况下,各个算法的平均定位误差比较。当采样样本数量一样的情况下,经过预处理的平均定位误差明显小于未处理的平均定位误差,在采样样本数为120时,ISVR算法的平均定位误差为1.51 m,明显低于采样样本数为120时的其他算法,SVR、SVR+校正坐标算法的平均定位误差分别为2.18、1.95 m。随着采样样本数的增长,ISVR算法相比于SVR+校正坐标算法的平均定位误差分别降低了0.47、0.51、0.43、0.44、0.38、0.39 m。通过增加校正坐

标,可以提高位置与RSS指纹之间的关联性,提高训练模型的准确性。随着采样样本数的增长,ISVR算法相比于SVR+预处理算法的平均定位误差分别降低了0.23、0.29、0.25、0.22、0.2、0.21 m。

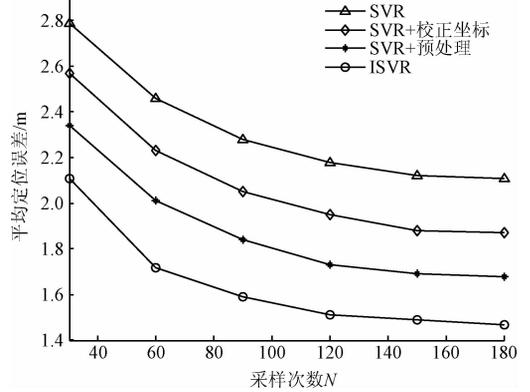


图7 预处理及增加校正坐标对SVR的影响
 Fig.7 Influence of preprocessing and increasing correction coordinate on SVR

2.5 不同定位算法的结果

图8所示为在测试点RSS指纹采集次数为180时,KNN、WKNN、传统SVR、预处理SVR以及ISVR算法的平均定位误差累计概率分布,可以看出当平均定位误差小于3 m时,ISVR、预处理SVR、传统SVR、WKNN、KNN算法的累计概率分别为93%、89%、82%、80%、75%。表1所示为不同算法的定位误差比较,可以看出基于支持向量回归的ISVR和SVR算法的定位精度明显高于KNN和WKNN,说明SVR算法在构建位置坐标与信号强度之间的非线性关系时具有优势,而且本文所提出的ISVR算法的定位精度明显高于其他定位算法,主要原因是在复杂的室内环境中,ISVR通过预处理过滤掉噪声信号,从而建立更准确的训练模型,并且通过增加校正坐标以及利用WKNN选择最优的位置坐标,提高了定位精度。

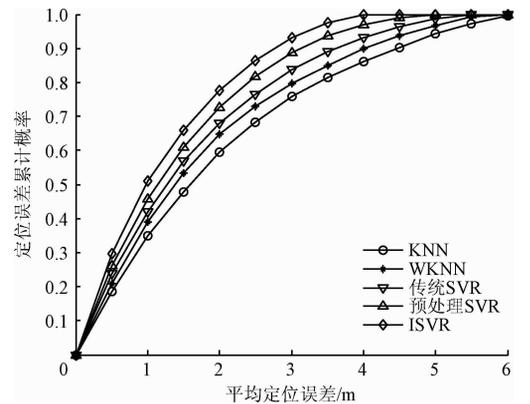


图8 不同算法的定位误差累计概率分布
 Fig.8 The probability distribution of positioning error for different algorithms

表1 不同算法的定位误差

Table 1 Positioning error of different algorithms

算法	平均定位 误差/m	≤3 m 的误差 概率/%	75% 的 误差/m	90% 的 误差/m
KNN	2.51	75	2.87	4.36
WKNN	2.32	80	2.61	3.92
传统 SVR	2.11	82	2.36	3.55
预处理 SVR	1.68	89	2.18	3.13
ISVR	1.47	93	1.88	2.71

3 结 论

本文针对现实 WLAN 环境中由于 RSS 信号的时变性而引入的冗余和噪声问题,以及传统支持向量回归定位算法在构建位置坐标与信号强度时的单输出导致位置坐标信息之间的关联性降低的问题,提出了一种改进的支持向量回归定位算法。该算法首先对 RSS 预处理构建指纹数据库,接着在训练阶段增加训练一个校正坐标 $z = x \cdot y$,最后根据加权反 K 近邻的方法得到最优的位置坐标。实验表明本文提出的算法优于传统的支持向量回归算法,提高了定位精度,但是在实验中只考虑了静止目标的定位,对于目标移动的定位问题有待下一步解决。

参考文献

- [1] GU Y Y, LO A, NEIMEGEERS I. A survey of indoor position system for wireless personal networks[J]. IEEE Communications Surveys & Tutorials, 2009, 11(1): 13-32.
- [2] HE S, CHAN S H G. Wi-Fi fingerprint-based indoor positioning: Recent advances and comparisons[J]. IEEE Communications Surveys & Tutorials, 2016, 18(1): 466-490.
- [3] 郑学理, 付敬奇. 基于 PDR 和 RSSI 的室内定位算法研究[J]. 仪器仪表学报, 2015, 36(5):1177-1185.
ZHENG X L, FU J Q. Study on PDR and RSSI based indoor localization algorithm [J]. Chinese Journal of Scientific Instrument, 2015, 36(5):1177-1185.
- [4] HAQUE I T, ASSI C. Profiling-based indoor localization schemes [J]. IEEE Systems Journal, 2015, 9(1): 76-85.
- [5] 郑新元, 严军, 范浩, 等. 线性不稳定环境下的 WIFI 室内定位系统[J]. 电子测量技术, 2015, 38(12): 121-124.
ZHENG X Y, YAN J, FAN H, et al. WIFI indoor

localization system in linear unstable environment [J]. Electronic Measurement Technology, 2015, 38(12): 121-124.

- [6] 郑学伟. 基于权值参数的 WiFi 定位算法研究[J]. 国外电子测量技术, 2014, 33(3):35-37.
ZHENG X W. Research of WiFi position algorithm based on weight parameters [J]. Foreign Electronic Measurement Technology, 2014, 33(3):35-37.
- [7] 张勇, 徐小龙, 徐科宇. 基于加权质心法的 WLAN 室内定位系统[J]. 电子测量与仪器学报, 2015, 29(7):1036-1041.
ZHANG Y, XU X L, XU K Y. Algorithm based on weighted centroid method for WLAN indoor positioning[J]. Journal of Electronic Measurement and Instrument, 2015, 29(7):1036-1041.
- [8] ZHANG G, ZHAN X, DAN L. Research and improvement on indoor localization based on RSSI fingerprint database and K-nearest neighbor points [C]. IEEE International Conference on Communications, Circuits and Systems, 2013:68-71.
- [9] YOUSSEF M, AGRAWALA A. The Horus WLAN location determination system [C]. International Conference on Mobile Systems, Applications, and Services, 2005:205-218.
- [10] 邓志安, 徐玉滨. 基于支持向量机回归算法的 WLAN 室内定位系统[J]. 仪器仪表学报, 2009, 30(6): 578-582.
DENG ZH AN, XU Y B. A support vector regression algorithm for indoor positioning in wireless local area network [J]. Chinese Journal of Scientific Instrument, 2009, 30(6): 578-582.
- [11] WU Z L, LI C H, NG K Y, et al. Location estimation via support vector regression [J]. IEEE Transactions on Mobile Computing, 2007, 6(3):311-321.
- [12] Y ZHANG, L DONG, L LAI, L HU. Study of indoor positioning method based on combination of support vector regression and kalman filtering [J]. International Journal of Future Generation Communication and Networking, 2016, 9(3): 201-214
- [13] PAN J J, KWOK J T, YANG Q, et al. Multidimensional vector regression for accurate and low-cost location estimation in pervasive computing [J]. IEEE Transactions on Knowledge and Data Engineering, 2006, 18(9): 1181-1193.
- [14] 陈永乐, 朱红松, 孙利民. 一种抗多径和阴影的视距指纹定位算法 [J]. 计算机研究与发展, 2013,

50(3):524-531.

CHEN Y L, ZHU H S, SUN L M. A line of sight fingerprint localization algorithm resisting multipath and shadow [J]. Journal of Computer Research and Development, 2013, 50(3):524-531.

[15] FANG S H, LIN T. Principal component localization in indoor WLAN environments [J]. IEEE Transactions on Mobile Computing, 2012, 11(1): 100-110.

[16] 石柯, 陈洪生, 张仁同. 一种基于支持向量回归的802.11无线室内定位方法 [J]. 软件学报, 2014(11): 2636-2651.

SHI K, CHEN H SH, ZHANG R T. Indoor location method based on support vector regression in 802.11 wireless environments [J]. Journal of Software, 2014(11): 2636-2651.

[17] SMOLA A J, SCHOLKOPF B. A tutorial on support vector regression [J]. Statistics and Computing, 2004, 14(3): 199-222.

[18] CHANG C C, LIN C J. LIBSVM: A library for support vector machines [J]. ACM Transactions on Intelligent Systems & Technology, 2011, 2(3):1-27.

作者简介



姚英彪,2006年于浙江大学获得博士学位,现为杭州电子科技大学教授,主要研究方向为嵌入式系统的硬件/软件协同设计、并行处理、多媒体信号处理和无线传感器网络等。

E-mail:yaoyb@hdu.edu.cn

Yao Yingbiao received his Ph. D. degree from Zhejiang University in 2006. Now he is a professor in Hangzhou Dianzi University. His main research interests include hardware/software co-design of embedded systems, parallel processing, multimedia signal processing, wireless sensor networks, and etc.



毛伟勇,2014年于泉州师范学院获得学士学位,现为杭州电子科技大学研究生,主要研究方向为室内定位、机器学习。

E-mail:151080040@hdu.edu.cn

Mao Weiyong received his B. Sc. degree from Quanzhou Teachers College in 2014. Now he is a master student in Hangzhou Dianzi University. His main research interests include indoor positioning and machine learning.