

# 基于压缩感知信号重构的wifi室内定位算法研究

刘莎莉,覃锡忠,贾振红

(新疆大学信息科学与工程学院, 乌鲁木齐 830046)

**摘要:**针对位置指纹定位算法在训练阶段信号数据采集量大和定位精度不高的问题,提出一种压缩感知(CS, Compressed Sensing)与K均值改进支持向量机(SVM, Support Vector Machine)相结合的定位算法模型(CS-KSVM)。CS算法在训练阶段利用已采集到的部分参考点wifi信号强度数据对整个指纹信号库进行重构以降低信号采集工作量,再用K均值改进SVM算法来实现测试点的准确分类。实验仿真结果表明,CS-KSVM算法在相同采样点条件下的定位精度明显要高于传统定位算法,同时在相同定位精度条件下大大减少了定位需要的采样点数。CS-KSVM算法在3米之内的定位准确度可以达到93.2%。

**关键词:**位置指纹;压缩感知;支持向量机;wifi;室内定位

**中图分类号:**TP391.9

**文献标识码:**A

**DOI编码:**10.14016/j.cnki.jgzz.2014.09.082

## Research on WiFi indoor positioning system based on compressed sensing signal reconstruction

LIU Sha-li, QIN Xi-zhong, JIA Zhen-hong

(School of Information Science and Engineering, Xinjiang University, Urumqi 830046)

**Abstract:** In view of the problem that in the training phase location fingerprinting localization algorithm has large amount of signal data acquisition number and its positioning accuracy is not high, put forward a kind of localization algorithm model(CS-KSVM) based on compressed sensing (CS, Compressed Sensing) and the improved support vector machine with K- means clustering (SVM, Support Vector Machine). The CS algorithm in training stage refactors the whole fingerprint library using the collected some reference point wifi signals strength data in order to reduce the workload of signal acquisition, and then uses k-means to improve the SVM algorithm for the sake of achieving accurate classification test point. The experimental simulation results show that the positioning accuracy of CS-KSVM algorithm is significantly higher than that of traditional localization algorithm under the condition of the same sampling point, at the same time the sampling points of positioning are greatly reduced under the condition of same positioning accuracy. The positioning accuracy of CS-KSVM algorithm within 3 meters can reach 93.2%.

**Key words:** Position fingerprint; Compressed sensing; Support vector machine; WiFi; Indoor positioning

随着无线通信技术的发展和智能移动终端的普及,使得基于位置的服务(LBS, Location-based Service)也得到飞速发展。目前,室内定位服务在商业活动、科学研究乃至紧急救助等行业都有着重要的应用,在救援、救灾应急指挥调度等特殊应用条件下,高精度室内定位尤为重要。由于室内定位环境比室外复杂,在室内场景下对人员进行定位时,信号传播很容易受到室内障碍物的阻挡而导致定位误差大大增加,使得传统GPS定位技术无法直接应用到室内定位中,因此如何提高无线网络室内定位精度的技术成为研究热点。

现如今,wifi室内定位凭借其低成本、高灵活性、

热点分布广的优势也得到了众多研究室内定位人士的青睐,大多数利用wifi信号进行室内定位的系统采用的都是以接收信号强度(RSS, Received signal Strength)作为多径特征来实现位置指纹识别算法实现定位<sup>[1]</sup>,但此算法存在以下两个问题:1)位置指纹识别算法在训练阶段需要采集大量的样本点,影响了大场景的应用;2)指纹数据库的完整性以及定位阶段匹配算法的选择也会定位精度产生一定程度上的影响<sup>[2]</sup>。迄今为止,国内外已经研究出较多的定位算法,如卡尔曼滤波、WKNN、概率定位、支持向量机<sup>[3]</sup>等,然而这些方法虽然在一定程度上提高了指纹定位精度,但是都需要测试大量的数据,测试时间较长<sup>[4]</sup>,因此如何降低信号采集工作量并进一步提高室内定位精度就成为我们亟需解决的问题。

本文提出了一种融合压缩感知与K均值改进支持向量机相结合的算法<sup>[5]</sup>(CS-KSVM算法),该算法只需采集少量样本点,就可以重构出高精度指纹信号库,定位阶段采用对信号进行K均值聚类,然后使用支持向量机实现分类定位。该方法解决了指纹信号库建立耗时耗力以及指纹信号库不完整导致定位精度下降的问题,实验结果证明了这种方法比已有的定位算法具有更高的定位精度,并且大大减少训练阶段信号数据采集工作量。

## 1 压缩感知算法

压缩感知算法指出,对于可压缩的或者在某个变换域内稀疏的信号,能够从远比奈奎斯特采样率低的采样数据或测量数据中以高概率得到无损恢复<sup>[6]</sup>。本文定位过程中采用的是wifi接入点(AP, Access point)的接收信号强度与位置坐标之间的关系进行定位,由于来自同一AP的RSS在二维空间中的变化具有一定的平滑性,故RSS相对应频域内的傅里叶系数存在一定的稀疏性。本文正是利用这一点,从而达到采用压缩感知算法来重构位置指纹信号库。将此技术用于不完整采样下指纹信号库的重构,不仅大大减少了训练阶段指纹库的采集工作,也给指纹库的更新带来新的解决途径。

本文中采用压缩感知重构算法的主要理论<sup>[7]</sup>如下:

令指纹信号  $x = x_c + x_g \in Q \oplus Q^\perp$ , 其中  $Q$  是  $\mathbb{R}^n$  的子空间。假设  $x_c$  是已测量得到的位置指纹信息,  $x_g$  是未测量的位置指纹信息,现在的目标是要重构出  $x$ , 也就是指纹信号。我们不知道指纹信号  $x$ , 但是知道它在变换基  $\Psi$  下是稀疏的, 分别令  $P_Q$  和  $P_{Q^\perp}$  为在  $Q$  和  $Q^\perp$  上的正交投影, 现在只需要解欠定方程

$$P_Q \Psi_c = P_Q x \quad (1)$$

求解(1)式的最稀疏解  $C$ 。

从压缩感知的角度来说,就是要解下面的方程

$$\min_c \|C\|_1, s.t. P_Q \Psi_c = P_Q x \quad (2)$$

知道了稀疏矩阵  $C$ , 则指纹信号可以通过  $x = \Psi C$  来重构。

根据定理<sup>[8]</sup>可知,最稀疏解有唯一性,可以得到如果满足  $\|\Psi^T x\|_0 < \frac{1}{2}(1 + \mu(P_Q \Psi)^{-1})$ , 那么就能求出唯一解:

$$\Psi^T x = \arg \min_c \|C\|_1, s.t. P_Q \Psi_c = P_Q x \quad (3)$$

基于位置指纹方法进行定位的缺点之一就是训练数据阶段十分耗时,CS-KSVM算法可以通过只测量一部分点,进而利用压缩感知信号重构的方法来完成

善位置指纹信号库,这样不仅缩短了训练数据时间,同时定位精度也高于插值方法<sup>[9]</sup>。

## 2 K均值聚类改进支持向量机

K均值聚类算法的基本思想是,假设含有  $M$  个数据的集合  $Y \sim (y_1, y_2, \dots, y_n)$ , 将这个数据集合划分为  $K$  个聚类中心集合  $X \sim (x_1, x_2, \dots, x_k)$  的问题,假设第  $k$  类的样本数目为  $N_k$ , 则  $N = \sum_{k=1}^K N_k$ , 每类  $X_k$  的均值为  $(m_1, m_2, \dots,$

$m_k)$ , 则  $m_k = \frac{1}{N_k} \sum_{i=1}^{N_k} y_i, k=1, 2, \dots, K$ 。K均值聚类的算法

是基于误差平方和准则的,即K均值聚类算法的最小目标函数为<sup>[10]</sup>

$$E = \sum_{k=1}^K \sum_{i=1}^{N_k} \|y_i - m_k\|^2 \quad (4)$$

用K均值聚类改进支持向量机的具体实现方法是:先采用K均值聚类对指纹信号库中的接收信号强度值进行一个区域划分,然后记录下各个聚类中心的RSSI值以及各位置坐标,将待定位的移动设备接收到的RSSI值与各个聚类中心的RSSI值进行匹配,找出空间距离上最近的那个类,对该类中心所在的那个区域内所有指纹点进行数据库网格分类匹配定位,最终实现位置估计。

支持向量机<sup>[11]</sup>的概念是Corinna Cortes和Vapnik等于1995年首先提出的,此算法最终归结为求解一个凸二次规划问题,因此,能够保证获得全局最优解。而传统的基于采样匹配算法在处理此类非线性问题时有很多不足,并且在定位时采集到的信号存在变化幅度大以及采集时间短的问题,如果直接与指纹数据库中采样点的信号强度直接进行匹配运算时容易造成误差,导致定位精度不高,所以采用K均值聚类改进支持向量机分类的方法来代替传统算法进行匹配定位,达到更高的定位精度,可以很好的完成本文精确定位的目的。

## 3 CS-KSVM算法流程

本文的定位算法分为两个阶段:1)训练阶段,首先采集少量的位置指纹数据,再通过压缩感知技术中正交匹配追踪算法对未采集点进行重构,从而达到减少测试数据的目的。由于位置指纹定位主要是根据移动设备接收到的信号强度与数据库中的信号强度做比较来确定位置坐标的,因此,对指纹采集点数据重构的准确度直接影响到定位的精度。2)定位阶段,通过移动设备实时接收到的信号强度与重构后的信号库做比对,采用K均值聚类改进支持向量机的方法

更容易对移动设备当前的位置做出精确的定位,并输出移动终端的位置坐标。

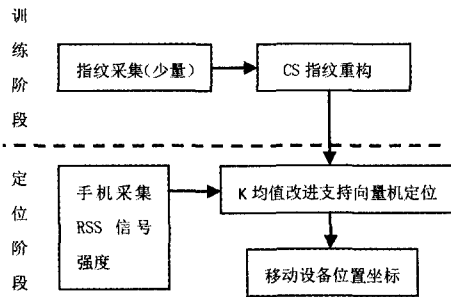


图1 CS-KSVM算法流程

该重构算法只适用于训练阶段对指纹信号库的完整建立,信号的重构主要是从长度为M的测量数据样本中重构出长度为N的完整原始信号的过程( $M < N$ ),本文M取20,N取80。压缩感知技术中有很多种重建信号的算法,基于正交匹配追踪算法<sup>[12]</sup>的指纹信号库重构算法具体流程如下<sup>[7]</sup>:

输入:已采集少量指纹信号数据库W,稀疏度为P,给定误差限定 $\epsilon$

(1)任取位置点,设其信号强度y,判断该位置点是否为未采集点,如果是则予以剔除,并令 $y=0$ ;否则予以保留;

(2)根据(1)中剔除的未采集点元素的位置,设计阶梯型观测矩阵;

(3)使用正交匹配追踪算法重构;

(4)初始化:  $r_0 = y$ , 索引集  $A_0 = \emptyset$ , 迭代次数  $t=1$ ;

(5)找到索引集  $G_t$ , 使得

$$G_t = \arg \max_{i=1,2,\dots,n} \left| \langle r_{t-1}, \Phi_i \rangle \right|; r_{t-1} \text{ 为信号剩余分量}, \Phi$$

是  $M \times N$  维的测量矩阵。

(6)令  $A_t = A_{t-1} \cup \{G_t\}$ ;

(7)通过求解最小二乘问题得到指纹信号的近似解:  $\hat{s} = \arg \min_s \|y - \Phi_s\|_2$ ;

(8)计算:  $r_t = y - \Phi \hat{s}$ ;

(9) $t=t+1$ ,如果  $t < P$  或者  $\|r_t\|_2 \geq \epsilon$ , 返回2), 依次迭代。

输出:索引集A,重建指纹信号,余量 $r_t$ 。

根据以上算法完成指纹信号库的重构,得到一个完整的指纹信号库,由于在定位阶段将输入的指纹信号库中的信号强度值做了聚类,就可以得到各个类的中心,支持向量机分类匹配过程中就直接将移动设备接收到的RSSI值与各个类中心进行匹配,从而得到与移动设备在空间距离上最近的那个类,最后根据支持向量机分类模型对待定位移动设备位置进行分类,分类后对所得到的x和y坐标再进行组合后就可以得

到待定位移动设备的定位坐标。在此引入K均值聚类来改进支持向量机是为了在定位阶段快速缩小范围,提高定位的效率。

## 4 仿真与结果分析

本实验采集样本环境设定在某教学楼,室内环境如图2所示,移动终端采用小米2s手机,软件为wifi信号检测,平均2s/sample,每个位置点采集120个样本值,将每个点采集的样本值进行去异常值和取均值处理<sup>[13]</sup>。

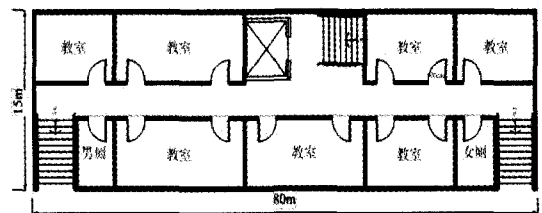


图2 实验环境平面图

为了检验本文算法的定位效果,选取WKNN、SVM和插值三种传统定位方法的效果作为对比<sup>[14]</sup>,使用MATLAB进行仿真<sup>[15]</sup>,结果如图3和图4所示。

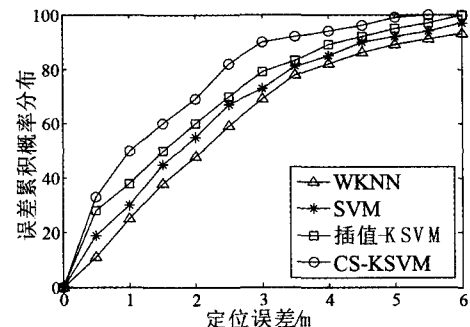


图3 各种算法的误差累积概率分布

图3所示为选取20个点作为参考点时,各种算法的误差累计概率分布图,描述了四种算法在测试点处的平均定位精度。

表1 四种算法误差累积概率分布比较

	WKNN	SVM	插值-KSVM	CS-KSVM
小于1m	26.1%	31.2%	38.4%	49.5%
小于2m	48.9%	50.4%	60.1%	70.2%
小于3m	64.3%	68.5%	79.6%	93.2%

由表1可以看出,本文提出的CS-KSVM算法的定位精度明显要高于别的定位算法。由于WKNN不能够充分提取出WiFi信号中的统计信息,所以精度最低。KSVM方法能够利用SVM的优秀的泛化能力对聚类之后的数据进行更精确的分类,其精度比SVM高。CS算法重构数据库的能力要优于插值方法,在线阶段的定位能够利用更多的信号信息,因此定位精度最高。

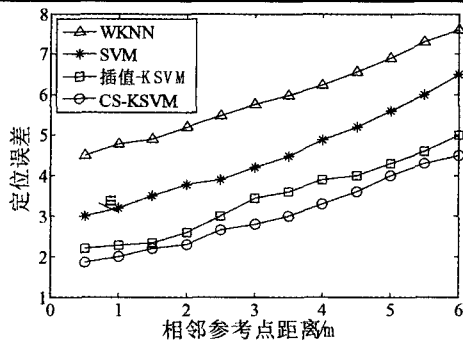


图4 改变参考点间距离时的定位误差

图4所示为不同相邻参考点距离下进行采样后,四种定位算法的平均定位误差的比较。

表2 不同相邻参考点距离采样四种算法误差比较

	WKNN	SVM	插值-KSVM	CS-KSVM
1m	4.82m	3.25m	2.31m	1.98m
2m	5.83m	4.18m	3.48m	2.63m
3m	7.62m	6.52m	4.66m	4.12m

点距离相同条件下,可看出本文提出的算法平均定位精度都明显高于其余三种算法,由此可得,改变相邻参考点间距离进行采样时的定位误差,随着相邻距离的增加,即参考点个数的减少,每种算法的误差都随之增大,说明在训练阶段采集越多的样本数据,最后的定位误差会在一定程度上减小,但也不是采集的样本数据越多就一定越好。同时WKNN和SVM方法只能从训练阶段提供的有限数据中提取信息,因此它们在参考点个数较少时候的定位误差会变得很大。而插值法和CS方法具有重构数据的能力,减少参考点对它们的影响并不大,定位结果也说明了CS算法比插值法具有更好的重构数据能力,能够给定位阶段的定位提供更多参考消息,因此定位误差最小,并且大大减少了训练阶段数据量的采集。

## 5 结束语

本文针对室内定位服务系统进行分析,针对传统算法在训练阶段指纹信号数据采集量大和定位精度不高的问题,引入将压缩

感知算法与K均值改进的支持向量机相结合的CS-KSVM算法,该算法在相同采样点条件下有效的提高了定位精度,并且大大减少在相同定位精度条件下需要的采样点数。实验结果验证了本文算法具有少工作量、快速精确定位的特点,为大场景定位应用提供了可行性,具有更高的实用价值。

未来的工作主要包括以下两个方面:一是引入时间复杂度的概念,提高定位的效率;二是引入动态定位,增强人员定位的实用性以及人性化。

## 参考文献

- [1] 卢恒惠,刘兴川,张超等.基于三角形与位置指纹识别算法的WiFi定位比较[J].移动通信,2010,34(10):72-76.
- [2] Vuckovic, M.; Petrovic, I.; Vidovic, D. et al. Space grid resolution impact on accuracy of the indoor localization fingerprinting[C], Telecommunications Forum(TELFOR), 2011: 321-324.
- [3] 秦泗明.基于位置指纹的WiFi室内定位技术研究[D].西安:电子科技大学,2010年.
- [4] 夏英,王磊,刘兆宏.基于无线局域网接收信号强度分析的混合室内定位方法[J].重庆邮电大学学报:自然科学版,2012,24(2):217-221.
- [5] 朱宇佳,邓中亮,刘文龙等.基于支持向量机多分类的室内定位系统[J].计算机科学. 2012,39(4):32-35
- [6] Yukai Yao, Yang Liu, Yongqing Yu, et al. K-SVM: An effective SVM algorithm based on K-means clustering[J]. Journal of computers, 2013, 8(10):2632-2639.
- [7] David L. Donoho, Member. Compressed sensing[C]. IEEE, Transactions on Information Theory, April 2006.4(52): 1289-1306.
- [8] 屈冉.压缩感知算法及其应用研究[D].南京:南京邮电大学,2013.
- [9] Arian Maleki, David L. Donoho. Optimally tuned iterative reconstruction algorithms for compressed sensing[J]. IEEE Journal of selected topics in signal processing, 2010, 4(2): 330-341.
- [10] Ryosuke Kubota, Shigeaki Tagashira, Yutaka Arakawa, et al. Efficient survey database construction using location fingerprinting interpolation[C]. 2013 IEEE 27th International Conference on Advanced Information Networking and Applications. 2013:469-476.
- [11] Bulut Altintas, Tacha Serif. Improving rss-based indoor positioning algorithm via k-means clustering[J]. European Wireless, 2011, 681-685.
- [12] 丁世飞,齐丙娟,谭红艳.支持向量机理论与算法研究综述[J].电子科技大学学报(自然科学版).2011,40(1):2-10.
- [13] J.Tropp, A.Gilbert. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE, Transactions on Information Theory, 2007, 53(12): 4655-4665.
- [14] 杨帆,赵东东.基于Android平台的WiFi定位[J].电子测量技术,2012,35(9):116-119,124.
- [15] Bagosi, T. Baruch, Z. Indoor localization by wifi[J]. Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on Digital Object. 2011:449-452.
- [16] 张德丰. MATLAB 通信工程仿真[M].北京:机械工业出版社,2010.