

一种基于支持向量回归的 802.11 无线室内定位方法*

石柯, 陈洪生, 张仁同

(华中科技大学 计算机科学与技术学院, 湖北 武汉 430074)

通讯作者: 石柯, E-mail: keshi@mail.hust.edu.cn, http://cs.hust.edu.cn

摘要: 802.11 无线局域网技术的广泛普及, 给无线室内定位系统带来了良好的发展契机. 提出了一种基于支持向量回归的 802.11 无线室内定位方法. 该方法主要包括离线训练和在线定位两个阶段. 离线阶段的主要工作是得到精确的位置预测模型; 在线阶段的主要工作是根据移动设备的接收信号强度(received signal strength, 简称 RSS)进行在线定位. 由于存在室内环境复杂、信道拥塞、障碍物影响和节点的通信半径有限等问题, 移动设备的接收信号强度易受干扰, 复杂多变. 针对以上问题, 离线阶段对接收信号强度信息进行统计分析, 得出数据过滤规则, 对训练数据集进行过滤, 以此提高训练样本质量, 从而提高支持向量回归预测模型的质量. 在线阶段使用连续 K 次测量定位法获取信号强度信息, 保证训练样本与在线输入信息之间的一致性, 提高最终的定位精度. 通过实验对该定位方法进行了综合对比分析, 实验结果表明: 与常用概率定位法、神经网络法相比, 该方法具有更高的定位精度, 同时具有对移动设备的存储容量及其计算能力要求较低的特点.

关键词: 802.11; 无线室内定位; 支持向量机回归(SVR); 数据过滤; 连续 k 次测量

中图法分类号: TP393

中文引用格式: 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. 软件学报, 2014, 25(11): 2636-2651. <http://www.jos.org.cn/1000-9825/4545.htm>

英文引用格式: Shi K, Chen HS, Zhang RT. Indoor location method based on support vector regression in 802.11 wireless environments. Ruan Jian Xue Bao/Journal of Software, 2014, 25(11): 2636-2651 (in Chinese). <http://www.jos.org.cn/1000-9825/4545.htm>

Indoor Location Method Based on Support Vector Regression in 802.11 Wireless Environments

SHI Ke, CHEN Hong-Sheng, ZHANG Ren-Tong

(College of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China)

Corresponding author: SHI Ke, E-mail: keshi@mail.hust.edu.cn, http://cs.hust.edu.cn

Abstract: The widespread of the 802.11-based wireless LAN technology brings a good opportunity for the development of the indoor positioning system based on 802.11. In this paper, a 802.11-based indoor positioning method using support vector regression (SVR) is presented. The method consists of two periods: offline training period and online location period. The accurate position prediction model is achieved in the offline training period by SVR, and the exact position is determined in the online location period according to the received signal strength (RSS) of the mobile devices. Due to the complex indoor environment, wireless channel congestion, obstructions and limitation of node communication range, the RSS is vulnerable and changeable. To address the above issues, corresponding data filtering rules obtained through statistical analysis are applied in offline training period to improve the quality of training sample, and thus improve the quality of prediction model. In the online location period, k -times continuous measurement is utilized to obtain the high quality input of the received signal strength, which guarantees the consistency with the training samples and improves the position

* 基金项目: 国家自然科学基金(60933012); 国家“十二五”国防预研计划; 教育部高等学校博士学科点专项科研基金(20110142110062); 湖北省国际合作项目(2010BFA008)

收稿时间: 2013-01-04; 修改时间: 2013-09-09; 定稿时间: 2013-12-03

accuracy of mobile devices. Performance evaluation and comprehensive analysis are done through intensive experiments, and the results show that the presented method has a higher positioning accuracy when compared with the probability positioning method and neutral network positioning method, and its demand for the storage capacity and computing power of the mobile devices is also low at the same time.

Key words: 802.11; wireless indoor location; SVR; data filtering; k -times continuous measurement

位置信息是普适计算中重要的上下文信息.在室内环境中,通过定位提供移动设备的准确位置信息,对普适计算的各类应用具有重要的意义,可使得普适计算系统能够根据对象位置为对象提供准确、及时的服务^[1-3].具体的应用领域包括用户行为识别、智能医疗、智能建筑等.

当前,GPS^[4]占据着定位系统的主要市场,但其无线信号由于受到建筑物、墙壁等的影响,不能在室内环境中工作.为了提供室内定位服务,国内外研究者提出了基于不同技术原理的多种室内定位系统,基于 802.11 无线局域网的定位系统是其中最常见的一种系统.这是由于 802.11 无线局域网已被广泛部署,特别是在 GPS 不能工作的室内环境中;同时,802.11 接口内置了提供 RSS(received signal strength)信息的功能,基于 RSS 定位不需要改变现有系统,从而降低了成本,而且该系统能够在非视距环境下工作.

802.11 无线定位系统中最常用的定位方法是场景分析方法.该方法将整个定位过程分为两个阶段:离线阶段和在线阶段.离线阶段的主要工作是采集指纹数据,通过将移动设备依次置于已知位置的参考点采集 RSS,建立无线信号地图.在在线阶段,移动设备实时采集 RSS,根据离线阶段的无线信号地图进行定位.但 802.11 无线局域网使用 2.4 GHz 或者 5 GHz 的频段(当前,绝大多数都工作在 2.4 GHz).由于 2.4 GHz 是水的共振频率,人的身体会吸收该频率的无线信号,再加上室内环境的复杂性、信道拥塞、障碍物影响和节点的通信半径问题有限等因素的影响,在实际应用中,移动设备接收的 RSS 复杂多变,而且极不稳定,这对定位算法的精度有严重影响.另外,大多数现有定位方法要求移动设备存储无线信号地图,在大规模应用场景中,这对移动设备的存储容量和计算能力提出了较高的要求.

本文提出了一种基于支持向量回归的 802.11 无线室内定位方法,整个定位过程分为离线阶段和在线定位两个阶段.离线阶段的主要工作是基于支持向量回归方法得到精确的位置预测模型.由于室内环境复杂,移动设备接收的 RSS 信息易受干扰,复杂多变,本文根据对无线信号强度信息的统计分析提出了相应的规则,对移动设备采集的原始 RSS 数据进行过滤,以提高训练数据集的质量.然后,基于支持向量回归方法对过滤后的数据集进行监督学习,以得到较好的位置预测模型.在线阶段的主要工作是根据移动设备接收的 RSS 信息进行在线定位.为了保证位置预测模型输入信息的质量以及保持离线阶段训练数据集与在线阶段模型输入数据之间的一致性,本文提出了连续 k 次测量定位法,连续 k 次测量采集 RSS 数据信息,综合 k 条 RSS 记录得到合理的模型输入,然后基于离线阶段获取的位置预测模型进行在线定位.与其他定位方法相比,本文提出的方法通过对 RSS 信息的预处理提高了定位精度,而且不需要移动设备存储无线信号地图,对移动设备的存储和计算能力要求较低.

本文第 1 节介绍背景及相关研究工作,给出 802.11 无线定位系统常用的定位方法.第 2 节介绍基于支持向量回归的定位方法的流程及细节.第 3 节通过实验分析该方法的性能,并与其他定位方法进行对比分析,以论证该方法的有效性.最后,对全文进行总结,并讨论未来的研究方向.

1 背景及相关工作

1.1 802.11 室内定位系统常用的定位方法

802.11 室内定位系统最常用的定位方法是场景分析方法^[5].该方法首先离线建立无线信号地图,利用统计学习理论得到定位模型,然后,移动设备实时采集 RSS 信息根据离线阶段的无线信号地图进行在线定位.目前常用的定位模型推导方法有 K 最近邻法(KNN)^[6,7]、最小 M 顶点多边形方法^[8]、概率法^[9]、神经网络法^[10]和支持向量机法^[11-15]等.

- ParamvirBahl^[6]提出了 KNN 方法进行定位.该方法首先通过移动节点在已知位置接收无线信号建立

“指纹”数据,然后在在线定位过程中基于最小二乘法搜索 k 个最接近的位置,然后对 K 个位置求位置平均,将获得的结果作为目标的位置.为了提高 KNN 定位方法的精度,Bhasker 等人^[7]在定位过程中引入了用户对定位结果的在线反馈.

- 最小 M 顶点多边形方法^[8]的离线训练阶段与 KNN 方法相似:在在线定位阶段,该方法根据实时检测的 M 维信号强度向量中的每个信号强度值,在其中寻找对应的信号强度值最接近的位置单元作为可选顶点,在所有 M 顶点多边形中,选择周长最小的多边形进行位置平均,以此作为位置预测结果.
- 概率法^[9]把整个定位场景划分为一系列的离散位置区域,通过训练得到各位置区域的 RSS 分布信息(通常以直方图或高斯模型的分布参数形式存储).当在线定位时,移动设备实时采集 RSS 信息向量,计算移动设备出现在各个位置的概率.但这种方法只能得到离散的位置信息,移动设备连续的位置信息只能通过插值处理得到.
- 文献[10]中提出了基于人工神经网络的室内定位系统.整个定位过程也包括离线和在线两个阶段:在离线训练阶段,通过(位置单元坐标,RSS)对神经网络模型进行训练;在在线定位阶段,将在线检测到的 RSS 信息作为输入,由神经网络模型输出在线定位结果.与概率模型相比,该方法的定位精度较高,实验结果显示,平均定位误差从 2.54 m 提高到 1.43 m.

综上所述,在文献[6-9]中,移动设备需要存储整个无线信号地图,这对容量有限的移动设备来说是较大的开销;同时,由于在线定位阶段也需要搜索整个地图,定位延时较大,而且定位精度不高.文献[10]提出的方法由于计算开销大,具体定位过程不是由移动设备完成,而是由服务器完成的,这种集中式方式的定位带来了传输时延,并可能使服务器成为瓶颈,并且用户隐私的保护也是一个难点.

1.2 基于支持向量机的定位方法

在数据分类归并、回归估计以及函数逼近方面,基于核函数的学习方法取得了巨大成就^[11,12].支持向量机(support vector machine,简称 SVM)就是属于此类方法的一种新型机器学习方法,具有拟合精度高、选择参数少、推广能力强、全局最优等特点.一般认为,当数据扩展到高维特征空间中时,数据具有更好的可分性,从而适合使用支持向量分类、支持向量回归以及其他核函数学习方法^[16].支持向量机可以分为支持向量分类(SVC)和支持向量回归(SVR)两类.其中,支持向量回归用来处理函数回归问题,已经成功地应用于系统识别、非线性系统的预测等方面,并且取得了较好的效果.

Battiti 等人^[13]对比分析了基于统计学习理论的各种定位方法在场景分析中的应用,结果表明,基于支持向量机的方法具有最好的结果,并推荐使用支持向量分类来进行定位.Nguyen 等人^[14]提出了应用基于支持向量机的方法,在 ad hoc 网络中进行定位.该方法假设所有节点都能收到对方的信号,同时,网络中有若干位置已知的锚节点,并将这些已知位置的节点分为 n 个类别.系统采集所有锚节点的数据作为支持向量机的训练数据集,得到相应的分类模型,最后,基于该模型进行在线定位.Tran 等人^[15]研究了基于支持向量机的无线传感器网络定位方法,分析了定位误差的上界,并提出了修正的最大似然法优化方案以提高定位精度.Yin 等人^[16]也做了类似于文献[15]的工作,提出了一个层级支持向量机(H-SVM)方案来进行大规模传感器网络的定位,精确分析了该定位方法的平均误差及方差以及定位误差的概率分布.

Wu 等人^[17]基于支持向量机回归原理进行移动通信系统的定位研究,通过设计新的核函数应对移动通信系统中信号的漏检问题,并且通过一个综合实验对比了不同定位方法的定位精度,结果表明,在地形复杂的定位场景中,支持向量机回归方法的性能最好.

综上所述,文献[13-17]基于支持向量机的分类原理研究了 ad hoc 与传感器网络中的定位问题,这与室内定位应用场景不同;同时,由于只能将位置对应于离散的类别,所以类别的离散程度与定位精度密切相关,在文献[15]中,取预测类和相邻下一个类别的中点作为定位位置.文献[17]采用基于支持向量机回归原理进行定位研究,并提出设计新的核函数来处理移动通信信号的不稳定性,提高定位精度.但其针对的是蜂窝通信系统,而且没有提出对训练数据集以及在线输入进行预处理.

2 基于 SVR 的定位方法

2.1 概述

基于 SVR 的定位方法的核心是通过支持向量回归原理找出 802.11 无线局域网 RSS 信息与其地理位置的关系.其总体框架如图 1 所示,移动设备通过与 802.11 无线接入点的通信获取自身的位置信息.

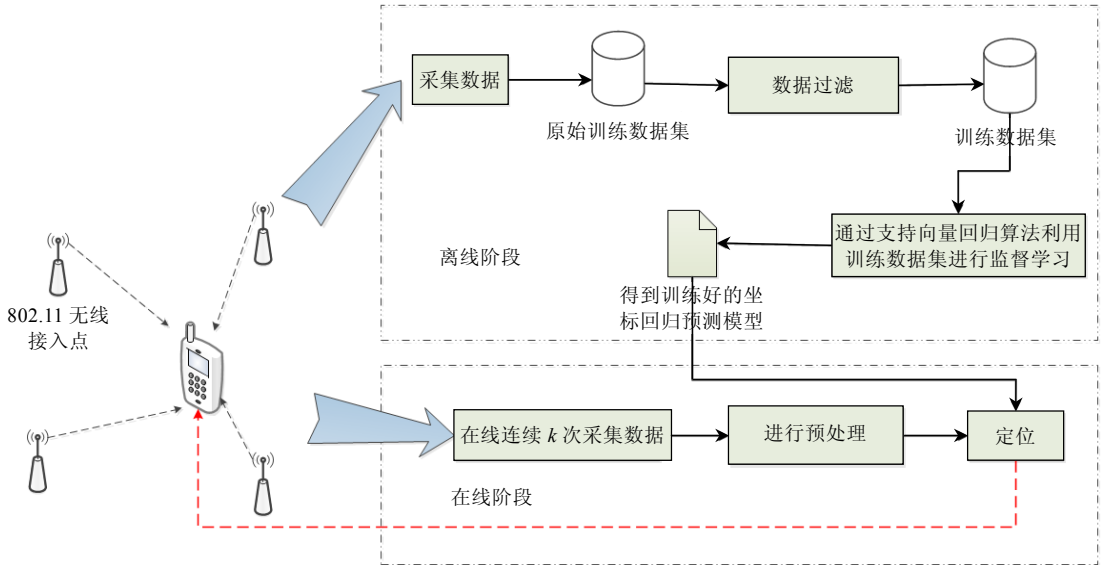


Fig.1 Framework of SVR-based location method

图 1 基于 SVR 的定位方法的总体框架

整个定位过程包括离线和在线两个阶段:

(1) 离线阶段

离线阶段的主要工作包括:

- 采集数据

通过将移动设备依次置于预先部署好的位置已知的参考点,采集各 802.11 无线接入点的 RSS 信息,得到原始训练数据集.

- 数据过滤

由于室内环境复杂,再加上信道拥塞等因素的影响,移动设备接收的 RSS 信息易受干扰,复杂多变,因此,移动设备采集到的原始数据集质量较差.在数据过滤阶段,采用数据过滤规则对采集的原始数据集进行过滤处理,以提高训练样本的质量.

- 通过支持向量回归算法利用训练数据进行监督学习

利用支持向量回归算法训练过滤后的数据集,以得到相应的位置预测模型.

(2) 在线阶段

- 连续 k 次采集 RSS 信息

由于 RSS 信息易受干扰,复杂多变,即使在同一位置,移动设备每次测量得到 RSS 信息也不一致.根据离线阶段对移动设备接收 RSS 信息的统计分析,移动设备接收到的 RSS 信息的质量与其所在位置、环境以及各 802.11 无线接入点(WAP)的位置密切相关.为了得到高质量的模型输入,在线阶段采用连续 k 次测量法,即,移动设备在同一位置连续测量 k 次得到位置预测模型的原始 RSS 输入信息.

- 原始 RSS 输入信息预处理

对原始 RSS 信息进行预处理,综合 k 条 RSS 记录信息得到模型的最最终输入.

- 定位

将预处理后的 RSS 信息作为离线阶段训练得到的位置预测模型的输入,确定移动设备的位置.

通过对训练数据进行过滤,并在采集数据时连续多次测量,能够有效地降低各种干扰造成的 RSS 信息瞬时变化对定位的影响.然而,当周围环境或设备发生永久性变化时,如室内布局的改变、无线接入点位置的改变,会引起 RSS 信息的永久性变化,需要重新采集训练数据,重新进行离线训练,得到新的定位模型.离线训练的触发可以基于对定位结果的误差分析,即定期在参考点采用定位模型进行定位,分析其误差,当误差在一段时间内都超出给定阈值时,就重新进行训练.

2.2 基于支持向量回归的位置预测模型

由图 1 可知,位置预测模型描述移动设备所处的物理位置与其从各 802.11 无线接入点接收到的 RSS 信息之间的关系.假设定位系统共有 n 个 802.11 无线接入点,移动设备在物理位置 s 每次测量接收到的 RSS 信息为 $x_i = \{mac_1(s), mac_2(s), \dots, mac_n(s)\}$, $mac_i(s)$ 为移动设备在 s 参考点位置接收到第 i 个无线接入点的信号强度值. L_i 为移动设备在参考点 s 的坐标,则对于给定的数据集 $D = \{(x_i, L_i)\}_{i=1}^N$ ($x_i \in R', L_i \in R$), 函数回归理论的目标是寻找映射 $f: R' \rightarrow R$, 使得 $f(x_i) \approx L_i$. 由于该映射关系属于非线性映射关系,根据支持向量机方法^[12,17],对于非线性问题,可以首先通过非线性函数 $\Phi(x)$ 将输入空间 R' 映射到高维的特征空间 R'' , 然后在高维特征空间中利用线性函数对数据集 $\{(\Phi(x_i), L_i)\}_{i=1}^N$ 进行拟合:

$$L = f(x) = W^T \Phi(x) + b \quad (1)$$

其中, $W \in R''$ 为权重向量, b 为偏置项.

根据统计学习理论,可以通过使如下目标函数(即训练误差函数)取最小值计算得到 W, b :

$$R(w) = \frac{1}{2} w^T w + \lambda \sum_{i=1}^N |L_i - (W^T \Phi(x) + b)|_{\varepsilon} \quad (2)$$

其中,

1. ε, λ 为经验参数;
2. $|L_i - (W^T \Phi(x) + b)|_{\varepsilon}$ 称为 ε 不敏感损失函数,其取值如下:

$$|L_i - (W^T \Phi(x) + b)|_{\varepsilon} = \begin{cases} 0, & |L_i - (W^T \Phi(x) + b)| < \varepsilon \\ |L_i - (W^T \Phi(x) + b)| - \varepsilon, & |L_i - (W^T \Phi(x) + b)| \geq \varepsilon \end{cases} \quad (3)$$

即,当预测值的误差小于 ε 时,损失函数的取值为 0; 否则,对其进行线性惩罚.

引入两个松弛变量 ξ_i, ξ'_i , 等价于下列优化问题:

使得:

$$R(w) = \frac{1}{2} w^T w + C \left(\sum_{i=1}^N (\xi_i + \xi'_i) \right) \quad (4)$$

取最小值,并满足:

$$\begin{cases} L_i - W^T \Phi(x_i) \leq \varepsilon + \xi_i \\ W^T \Phi(x_i) - L_i \leq \varepsilon + \xi'_i \\ \xi_i \geq 0 \\ \xi'_i \geq 0 \\ i = 1, \dots, N \end{cases}$$

参数 ε 表明系统对回归函数在训练数据集上的误差要求, ε 越小,回归函数在训练数据集的误差越小,所得到的回归函数估计精度越高,但支持向量数越多. 参数 C 是对回归估计函数的误差大于 ε 的训练数据集上数据元组的惩罚. C 越大,表示对这些数据元组的惩罚越大.

上述优化问题可定义为如下拉格朗日函数:

$$L(w, \xi_i, \xi'_i, \alpha, \alpha', t, t') = \frac{1}{2} w^T w + C \left(\sum_{i=1}^N (\xi_i + \xi'_i) \right) - \sum_{i=1}^N (t_i \xi_i + t'_i \xi'_i) - \sum_{i=1}^N \alpha_i (W^T \Phi(x_i) - L_i + \varepsilon + \xi_i) - \sum_{i=1}^N \alpha'_i (L_i - W^T \Phi(x_i) + \varepsilon + \xi'_i) \tag{5}$$

其中, α, α', t, t' 为拉格朗日乘数, 当其取得极值时, 必满足:

$$\frac{\partial L(w, \xi_i, \xi'_i, \alpha, \alpha', t, t')}{\partial w} = w - \sum_{i=1}^N (\alpha_i - \alpha'_i) \Phi(x_i) = 0 \tag{6}$$

$$\frac{\partial L(w, \xi_i, \xi'_i, \alpha, \alpha', t, t')}{\partial \xi_i} = t_i - (C - \alpha_i) = 0 \tag{7}$$

$$\frac{\partial L(w, \xi_i, \xi'_i, \alpha, \alpha', t, t')}{\partial \xi'_i} = t'_i - (C - \alpha'_i) = 0 \tag{8}$$

将公式(6)~公式(8)代入公式(5), 同时根据 SVR 理论^[12]引入核函数 $K(x_i, x_j)$, 并利用 Wolfe 对偶技巧^[12]将上述问题转化为下述对应的对偶优化问题:

使得:

$$Q(\alpha, \alpha') = \sum_{i=1}^N L_i (\alpha_i - \alpha'_i) - \varepsilon \sum_{i=1}^N (\alpha_i - \alpha'_i) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N (\alpha_i - \alpha'_i) (\alpha_j - \alpha'_j) K(x_i, x_j) \tag{9}$$

取最大值, 并满足:

$$\begin{cases} \sum_{i=1}^N (\alpha_i - \alpha'_i) = 0 \\ 0 \leq \alpha_i \leq C \\ 0 \leq \alpha'_i \leq C \\ i = 1, \dots, N \end{cases} \tag{10}$$

从而, 对应的回归公式(1)可改写为下式:

$$L = \sum_{i=1}^N (\alpha_i - \alpha'_i) K(x, x_i) + b \tag{11}$$

根据支持向量回归函数的性质^[12]: $\alpha_i \times \alpha'_i = 0, \forall i \in \{1, 2, \dots, N\}$, 只有少数的 $\alpha_i - \alpha'_i$ 不等于 0, 这些参数对应的向量为支持向量, 回归函数 $f(x)$ 完全由其决定.

经常使用的核函数有:

- 线性核:

$$K(x_i, x_j) = x_i^T \cdot x_j \tag{12}$$

- 多项式核:

$$K(x, x_j) = (\gamma(x_i^T \cdot x_j) + c)^d \tag{13}$$

其中, γ, c, d 为参数.

- 高斯核(RBF 核):

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \tag{14}$$

- Sigmoid 核:

$$K(x_i, x_j) = \tan h(\gamma(x_i \cdot x_j) + \beta) \tag{15}$$

其中, γ, β 为参数.

最终得到的位置预测模型的结构如图 2 所示.

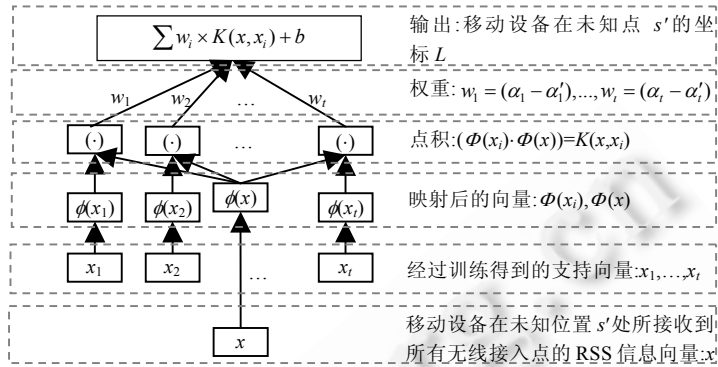


Fig.2 Location predicting model trained by SVR
图 2 基于 SVR 得到的位置预测模型

2.3 离线阶段数据过滤

在支持向量回归方法中,训练样本的质量决定了最终模型的质量.训练样本数据的质量越好,得到的模型也就越好.与此同时,模型输入的信息越精确,模型获得的结果也就越精确.在无线室内环境中,无线信号的传输受到多种因素的影响,移动设备接收到的 RSS 信息复杂多变,在某一固定点,每次测量所采集的 RSS 信息也会不断变化,从而导致训练样本以及在线输入的 RSS 信息质量较低,影响最终的定位结果.因此,在离线阶段,训练数据需进行过滤,以获得高质量的样本;在在线阶段,RSS 信息也要进行预处理,以获取更精确的输入.因此,本节将讨论离线阶段数据过滤,下一节讨论在线数据预处理.

通过对采集到的数据进行分析,移动设备在固定点多次测量接收到各无线接入点的 RSS 信息概率分布图如图 3 所示.我们发现,RSS 的变化规律与文献[18,19]描述的规律一致.即,移动设备在某个固定点所接收到的某个 802.11 无线接入点的 RSS 信息服从对数正态分布.

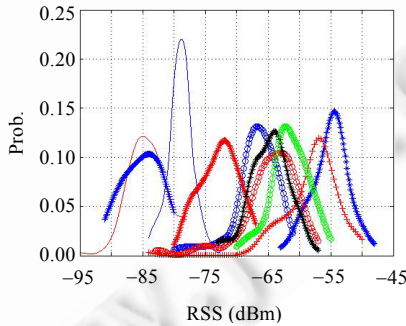


Fig.3 Distribution of the RSS information received by mobile device communicating with stationary AP
图 3 移动设备在固定位置与固定无线接入点通信时所接收 RSS 信息的概率分布

令

$$x = -\text{mac}_i(s) > 0 \tag{16}$$

则 x 的概率密度函数为

$$f(x) = \frac{1}{\sqrt{2\pi} \times \sigma \times x} \times \exp \left(-\frac{(\ln x - \mu)^2}{2\sigma^2} \right) \tag{17}$$

其中, $\text{mac}_i(s)$ 为移动设备在参考点 s 所接收到第 i 个无线接入点 RSS, μ 为 x 对数均值, σ 为 x 对数标准差.于是, x 的期望值为

$$E(x) = e^{\mu + \frac{\sigma^2}{2}} \tag{18}$$

根据上述概率分布,我们可以设置一个正的阈值 θ ,当 $0 < E(x_i) - \theta < -mac_i(s) < E(x_i) + \theta$ 时,才认为该值是一个有效值.当 RSS 值超出了上述范围时,就认为该信号由于干扰等原因出现了非正常的信号值过高或过低,属于异常信号.对于每一个测量值 $x = -mac_i(s)$,其有效的概率为

$$P_V(E(x_i) - \theta < x < E(x_i) + \theta) = \int_{E(x_i) - \theta}^{E(x_i) + \theta} \frac{1}{x \times \sigma \times \sqrt{2\pi}} \times \exp\left(-\frac{(\ln x - \mu_i)^2}{2\sigma^2}\right) dx \tag{19}$$

除了信号值异常之外,在定位时,无线网络也在传输数据.由于网络拥塞等原因,还会造成移动设备在某参考点的某次测量中可能没有得到第 i 个无线接入点的 RSS 信息,从而造成对应原始数据记录中的 $mac_i(s) = 0$.因此,还需要区分到底是网络拥塞造成的丢包使 $mac_i(s) = 0$,还是对应的位置关系使信号强度值本身就很弱.因此,我们基于多次测量来区分这两种情况.

假设移动设备每次测量之间相互独立,对于每一个参考点,分别统计移动设备在该位置每次测量无法接收到第 i 个无线接入点 RSS 信息的概率为 P_{M_i} :

$$P_{M_i} = \frac{m - j}{m} \tag{20}$$

其中, m 为在移动设备在参考点总共测量次数, j 为移动设备在该参考点 m 次测量中接收到第 i 个无线接入点 RSS 信息的总次数.

若 P_{M_i} 很高,则可以认为该位置距第 i 个无线接入点较远,接收到的信号强度本身就很低,或者该位置与第 i 个无线接入点的通信受到了干扰(如出现了障碍等),信号强度信息不稳定,以至于多次测量也无法获取有效的信号强度信息;若 P_{M_i} 很低,该位置应该处于第 i 个无线接入点的通信范围之内,能够接收到正常的信号,此时,若出现 $mac_i(s) = 0$,则可以认为由于网络拥塞造成的丢包导致无法检测到信号强度信息.为了提高预测模型的精度,这些数据记录应该从训练数据样本中去除.因此,我们定义有效测量阈值 ζ :当 $P_{M_i} \leq \zeta$ 时,我们认为移动设备在该参考点能够和第 i 个接入点进行有效的通信,获取信号强度信息,多次测量获得的数据大多数都是有效的;当 $P_{M_i} > \zeta$ 时,我们认为移动设备在该参考点和第 i 个接入点的通信受到较大的干扰,无法获取信号强度信息,多次测量获得的数据大多数都是无效的.基于上述分析,我们定义了相应的离线阶段数据过滤规则:

- (1) 对于移动设备在参考点采集的相对各个无线接入点的信号强度记录,如果该记录对应的 $P_{M_i} \leq \zeta$, ζ 为有效测量阈值,而该记录的 $mac_i(s) = 0$,则认为该记录无效,做过滤处理;如果记录对应的 $P_{M_i} > \zeta$,而该记录的 $mac_i(s) \neq 0$,则认为该条记录无效,同样做过滤处理.
- (2) 对于移动设备在参考点采集的每一条记录,如果该记录的 $P_{M_i} \leq \zeta$,同时, $-mac_i(s) < E(x_i) - \theta$ 或 $-mac_i(s) > E(x_i) + \theta$,则认为该记录无效.

规则(1)中的第 1 种情况描述了应该过滤掉网络拥塞造成的测量数据包丢失所导致的正常信号漏检,第 2 种情况描述了应该过滤掉移动设备和无线接入点之间无法正常通信时得到的少量不稳定的信号强度数据.规则(2)描述了过滤掉移动设备和无线接入点之间无法正常通信时受到干扰时产生的异常信号.

2.4 在线阶段数据预处理

离线阶段数据过滤去除了训练样本中不稳定、非正常的 RSS 数据,得到了高精度的预测模型.在线阶段使用该模型进行位置预测时,必须也输入高精度的 RSS 信息,避免不稳定、非正常的 RSS 数据,才能得到高精度的位置信息.因此,在线阶段采用 k 次测量法获取 RSS 信息. k 次测量法首先必须保证通过 k 次测量移动设备在在线阶段一定能够获取有效的 RSS 信息;然后,通过数据预处理,去除 k 条 RSS 信息记录中可能存在的不稳定、受干扰的数据,获取有效的真正反映位置关系的数据,以提高位置预测模型 RSS 输入信息的质量.

为了保证在在线阶段一定能够获取有效的 RSS 信息,首先必须合理设置 k 的大小.设移动设备在未知点的每次测量中能够测量到第 i 个无线接入点有效的 RSS 值为 $mac_i(s')$,令

$$x = -mac_i(s') > 0 \tag{21}$$

根据上面的分析,该值服从对数正态分布,若 $E(x_i)$ 为 x 的期望值,则该值在指定范围 $(E(x_i)-\theta, E(x_i)+\theta)$ 的概率为 $P_i(E(x_i)-\theta < x < E(x_i)+\theta)$, 其中, $E(x_i)-\theta > 0$. 考虑无法正常接收的情况,设其每次测量 RSS 值都在 $(E(x_i)-\theta, E(x_i)+\theta)$ 范围内的概率为 $P(E(x_i)-\theta < x < E(x_i)+\theta)$, 其中, $E(x_i)-\theta > 0$, 则有:

$$P(E(x_i)-\theta < x < E(x_i)+\theta) = (1 - P_{M_i}) \times P_i(E(x_i)-\theta < x < E(x_i)+\theta) \quad (22)$$

P_{M_i} 为移动设备在该位置每次测量无法接收到第 i 个无线接入点 RSS 信息的概率.

为了确保移动设备在线阶段能够在 k 次测量中至少有 1 次能够接收到第 i 个无线接入点的有效 RSS 信息, 则有:

$$1 - (1 - P(E(x_i)-\theta < x < E(x_i)+\theta))^k = \tau \quad (23)$$

其中, τ 为趋近于 1 的常数.

于是有:

$$k = \log_{(1 - P(E(x_i)-\theta < x < E(x_i)+\theta))} (1 - \tau) \quad (24)$$

因此,真实系统的 k 值必须大于等于公式(24)得出的值,才能保证获取到有效数据. k 值越大,有效测量次数越多,得到的有效数据越多.但 k 值越大,同时也意味着移动设备测量的时间就越长,这会使位置信息获取的时间变长,影响用户的体验,限制应用的范围(如用户移动的速度).公式(24)给出了 k 值的下限,具体 k 的取值与无线环境、用户的需求相关.一般来说,无线传输环境简单,无线网络的负载较轻,出现丢包、信号异常的概率较小, k 值可以较小;用户对位置信息精确度的要求较低, k 值也可以较小.反之, k 值应该较大.从公式(24)可以看出: k 值的下限的计算与移动设备在该位置所能接收到的无线接入节点 RSS 信息的概率、参数 θ 和 ζ 直接相关,移动设备在该位置所能接收到的无线接入节点 RSS 信息的概率由无线传输和网络环境所决定, θ 反映了用户对信号变化的容忍范围,参数 ζ 反映了保证的可靠性.具体的最佳取值,有待真实系统的校正确定.

确定 k 值后,移动设备进行 k 次测量,然后对获取的数据进行预处理.预处理的过程如下:

令

$$x_i = \text{mac}_i(s') > 0 \quad (25)$$

对于每一个无线接入点,首先统计其在这 k 次测量中的 P_{M_i} , 如果 $P_{M_i} \geq \zeta$, 则认为本次测量的 $\text{mac}_i(s') = 0$; 接下来分别统计 x_i 对数的均值 μ_i 和标准差 σ_i , 并根据公式(18)计算 x 的期望值 $E(x_i)$, 如果 $|\text{mac}_i(s') - E(x_i)| < \theta$, 则认为该值有效,记为 $\text{mac}'_{ij}(s')$, 否则认为无效.最后, $\text{mac}_i(s')$ 的取值为所有有效值的均值,即

$$\text{mac}_i(s') = \begin{cases} 0, & P_{M_i} \geq \zeta \\ \frac{\sum_{j=1}^t \text{mac}'_{ij}(s')}{t}, & P_{M_i} < \zeta \end{cases} \quad (26)$$

其中, ζ 为测量阈值, $\text{mac}'_{ij}(s')$ 为对应第 i 个无线接入点在 k 次测量中的第 j 个有效值, t 为对应第 i 个无线接入点在 k 次测量中其有效值的总数.

通过以上预处理,选择合适的 k 值,能够获取真实、有效的 RSS 信息,与离线训练集的数据一致,保证模型输入信息的精度,从而在很大程度上提高了回归模型定位精度.

3 性能评价分析

3.1 仿真环境

我们通过仿真的形式对本文提出的定位方法进行了详细的性能评价分析.仿真采用的数据来源于真实系统,即,King 和 Kopf 等人研究的基于 802.11 的无线定位系统的定位数据.该定位系统的部署如图 4 所示,其中,深色点代表参考点的位置,房间中的浅灰色点代表无线接入点的部署位置,浅色位置为在线测试位置.整个系统覆盖面积为 221 m², 整个区域内部署了 25 个无线接入点,即, $n=25$. 在每一个参考点,重复 110 次采集各个无线接入点的信号强度,即, $m=110$, 每次测量耗时 250 ms.



Fig.4 Deployment of the experimental system^[20]

图 4 实验系统部署^[20]

我们使用 libsvm-3.1 来实现 SVR 算法,参数 C 和 ε 分别取经验值 1 000 和 0.01.同时,为了度量定位误差,设模型对某个点预测得到的位置坐标为 $\{x_p, y_p\}$,而该点的实际坐标为 $\{x_r, y_r\}$,则有定位误差距离 $Error$ 为

$$Error = \sqrt{(x_p - x_r)^2 + (y_p - y_r)^2} \tag{27}$$

在实验中,默认对训练样本根据相应规则进行了过滤处理,且在线定位阶段进行连续 k 次测量,并通过预处理得到模型的输入.同时,在实验中通过设计相应的实验验证了不做训练样本过滤以及在线定位阶段预处理时对模型的定位精度的影响.对实验中参数的说明见表 1.

Table 1 Experimental parameters description

表 1 实验参数说明

参数名称	默认值	参数说明
训练样本大小	5 000	实验中所有训练样本大小的默认为 5 000,同时,在实验中验证了样本大小对定位模型的影响
无线接入点数量	17	实验中,无线接入点的数量默认为 17,同时,在实验中通过改变无线接入点数量验证了其对应定位结果的影响
核函数	高斯核函数	实验中默认为高斯核函数,同时,在实验中通过使用不同的核函数验证了其对应定位结果的影响
重复测量次数 k	8	实验中,在线定位阶段多次测量次数 k 默认值为 8,同时,在实验中通过改变测量次数验证了其对应定位结果的影响
ζ	0.9	实验中的默认值为 0.9, θ 的默认值为 4,同时,在实验中通过改变其取值验证了其对应测量次数 k 的影响
θ	4	实验中 θ 的默认值为 4,同时,在实验中通过改变其取值验证了其对应测量次数 k 的影响
测试样本大小	1 000	-

3.2 数据质量对模型的影响

离线阶段的数据过滤和在线阶段的数据预处理,分别提高了训练数据的质量和模型输入数据的质量.为了分析相应数据质量对基于支持向量回归的定位模型精度的影响,分别进行了以下 4 种实验,对不同场景中定位模型的精度进行对比与分析:

- 在实验 1 中,离线阶段不对训练样本进行过滤处理,在线阶段移动设备只在未知点测量 1 次各无线接入点的 RSS 信息,将其作为模型的输入.
- 在实验 2 中,离线阶段不对训练样本进行过滤处理,在线阶段移动设备在未知点重复测量 8 次,然后综合所采集的 8 条 RSS 信息作为模型的输入.
- 在实验 3 中,离线阶段对训练样本依据本文提出的相应规则进行数据过滤,在线阶段移动设备只在未知点测量 1 次各无线接入点的 RSS 信息,将其作为模型的输入.
- 在实验 4 中,离线阶段对训练样本依据本文提出的相应规则进行数据过滤处理,在线阶段移动设备在

未知点重复测量 8 次,然后综合所采集的 8 条 RSS 信息作为模型的输入.定位模型的误差累积概率分布如图 5 所示,定位误差见表 2.

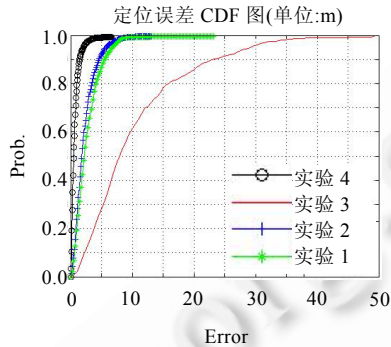


Fig.5 Impact of data quality on the location model

图 5 数据质量对模型定位的影响

Table 2 Error of location results in different experiments (m)

表 2 各实验定位误差结果分析 (m)

	实验 1	实验 2	实验 3	实验 4
50%测试样本	1.78	2.23	7.78	0.48
80%测试样本	3.41	4.02	15.60	1.00
90%测试样本	4.51	5.40	22.40	1.40
平均定位误差	2.25	2.7	10.37	0.68

由图 5 和表 2 可知:

- 实验 4 的定位精度最好,其测试样本的平均定位误差距离为 0.68 m,即,同时应用离线数据过滤和在线数据预处理的定位效果最好;
- 实验 1 次之,其测试样本的平均定位误差距离为 2.25 m;
- 再次是实验 2,其测试样本平均定位误差距离为 2.70 m;
- 实验 3 最差,其测试样本的平均定位误差距离为 10.37 m.

因此,离线阶段的数据过滤和在线阶段的数据预处理能够有效地提高模型的定位精度,但使用高质量的训练样本得出的模型也需要高质量的输入数据.

3.3 与其他定位方法的对比

为了验证本文提出的定位方法的有效性,我们与基于概率模型的方法^[18]、无过滤和预处理的 SVR 方法、ANN 方法^[10]进行了对比实验和分析.最终的定位误差累积概率分布函数图如图 6 所示,定位误差见表 3.

由图 6 和表 3 可知:带预处理的 SVR 定位方法(即我们提出的方法,包括过滤和预处理)定位精度最高,优于无预处理的 SVR 方法、ANN 方法、带预处理的 ANN 方法和概率模型定位方法;带预处理的 ANN(同样在离线阶段对样本进行数据过滤,在在线阶段进行 k 次测量,然后经过相同的预处理方法进行定位)的表现接近有预处理的 SVR 定位方法.

- 在带预处理的 ANN 定位方法中,测试样本的平均定位误差距离为 0.886 m;
- 在带预处理的 SVR 定位方法中,测试样本平均定位误差距离为 0.683 m;
- 在无预处理的 SVR 定位方法中,测试样本的平均定位误差距离为 2.24 m;
- 在无预处理的 ANN 定位方法中,测试样本的平均定位误差距离为 2.24 m;
- 在经典概率模型定位方法中,测试样本的平均定位误差距离为 2.857 m.

因此,有效的数据过滤和预处理能够大幅度提高基于机器学习的定位方法的精度,而 SVR 方法的表现要优

于 ANN 方法.

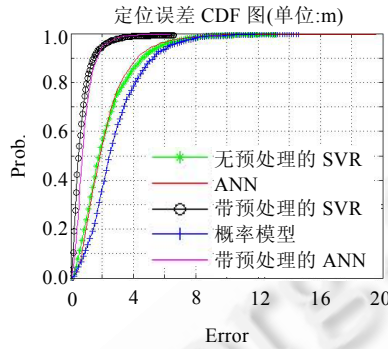


Fig.6 Comparison with the other location methods
图 6 与其他定位方法的对比

Table 3 Error of location results for different location methods (m)
表 3 各定位方法定位误差结果分析 (m)

	带预处理的 ANN	带预处理的 SVR	无预处理的 SVR	ANN	概率模型
50%测试样本	0.74	0.48	1.78	1.87	2.47
80%测试样本	1.25	1.00	3.40	3.17	4.10
90%测试样本	1.66	1.40	4.53	4.24	5.21
平均定位误差	0.886	0.683	2.24	2.24	2.857

3.4 训练集大小对定位方法的影响

通过不同大小的训练集训练得到定位模型的定位精度不同,图 7 描述了不同大小的训练数据集对模型定位误差的影响.从图中可知:

- 当训练数据集规模小于 2 000 时,训练所得到的模型平均定位误差距离随着训练数据集大小的增加而迅速地缩短;
- 当训练数据集达到 2 000 而小于 4 000 时,随着训练数据集的增加,训练所得到的模型平均定位误差距离缩短的速度明显变慢;
- 当训练样本大于 4 000 时,所得到的模型平均定位误差距离随着样本的增加而基本不变.

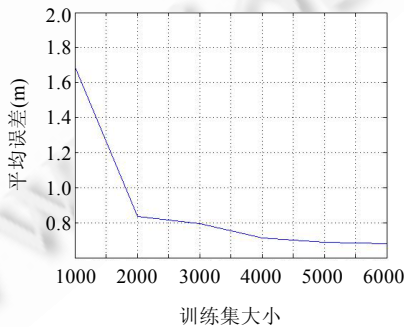


Fig.7 Impact of training data size on the location model
图 7 不同训练集的大小对定位模型的影响

3.5 不同核函数对定位方法的影响

根据支持向量机回归理论,满足 Mercer 定理的核函数能够把高维特征空间中的点积运算转化为低维输入

空间的核函数运算^[16],从而解决维数灾难问题.不同核函数训练得到的模型会有不同的表现.我们通过实验分别采用线性核、多项式核、RBF 核以及 Sigmoid 核对样本进行训练,并通过测试集分别测试相应模型的定位精度.

实验结果如图 8 所示,在所有核函数中,采用高斯核函数训练得到的模型的精度最高:

- 当使用线性核函数时,测试样本的平均定位误差距离为 2.44 m;
- 当使用 Sigmoid 核函数时(其中, $\gamma=1/17,\beta=0$),测试样本的平均定位误差距离为 2.46 m;
- 当使用 RBF 核函数时,测试样本平均定位误差距离为 0.683 m;
- 当使用多项式核函数时($\gamma=1/17,c=0,d=3$),测试样本的平均定位误差距离为 1.15 m.

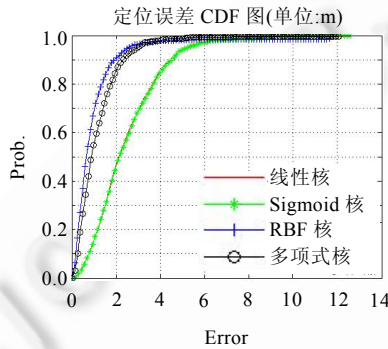


Fig.8 Impact of kernel function on the location model

图 8 不同核函数对定位模型的影响

3.6 锚节点数量对定位方法的影响

一般来讲,在一定的空间区域内,锚节点的数目越多,则定位误差越小.为了分析锚节点数量对定位方法的影响,我们通过改变给定实验配置中的锚节点数量进行了对比分析,并与基于概率模型的方法、无过滤和预处理的 SVR 方法、ANN 方法进行了比较.在实验中,每次从所有锚节点中随机选择相应数量的锚节点,其数量从 1 逐一增加到 20.实验结果如图 9 所示,定位模型的定位精度随着锚节点的增多而提高.当锚节点数目基本相同时,精度差别不大,此时,若锚节点的位置相邻,它们之间可能会有信号干扰,从而使采集到的数据不够准确,但经过大量实验取平均值后,几乎不存在波动情况.当锚节点达到一定数量时,即可得到较高的定位精度,并且定位精度差别不是很大.

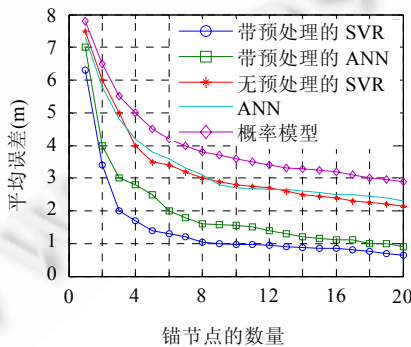


Fig.9 Impact of the number of WAPs on the location model

图 9 锚节点数量对定位模型的影响

如图 9 所示,本文提出的定位方法在锚节点数量增加到 8 个以后时,当锚节点的数量再增加时,模型的定位精度并没有明显的提高.此外,无论锚节点的数目如何变化,本文方法的定位精度都优于其他方法.

3.7 重复测量次数 k 对定位方法的影响

如第 2.4 节所述,参数 k 对系统的定位精度有着重要影响,与相应的工作环境有很大的联系,同时又与参数 θ, ζ 密切相关.具体的设置值要根据无线传输环境、用户对延时以及定位精度的要求而定,在满足用户需求的同时保证较好的定位精度.为了验证 k 值对模型定位的影响,我们将 k 值从 1 逐一增加到 12,并对不同 ζ, θ 参数值进行了对比分析.实验结果如图 10 所示,从图中可知,总体来讲,模型的定位精度随着 k 值的增大而提高.同时,参数 ζ 和 θ 的取值标识着对训练数据样本进行过滤的标准,该标准越严格,则定位模型定位精度达到收敛时所需要的测量次数 k 就越多,定位精度也就越高;该标准越低,则相反,定位模型定位精度达到收敛时所需要的测量次数 k 就越少,定位精度也就越低.

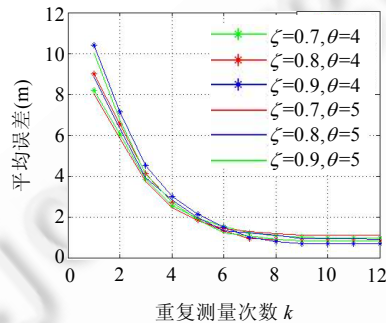


Fig.10 Impact of the number of measurements on the location model with different system parameters

图 10 不同系统参数下,重复测量次数对定位模型的影响

3.8 对移动设备的性能需求

由于移动设备的存储空间和计算能力有限,定位方法如何减少对移动设备数据存储量和计算能力的需求也相当重要.神经网络定位方法^[10]需要进行复杂的计算,移动设备无法满足其计算需求,因此,定位计算通常在服务器上完成,当移动设备需要位置信息时,需要向服务器提出请求,服务器计算完成后返回位置信息给移动设备.这种模式增加了网络负载,且定位精度可能会受网络延迟的影响,而且不利于保护用户隐私.

概率性的定位方法需要存储信号地图.文献[20]所描述的系统的无线信号地图的容量为 6 110 KBytes,尽管其提出了基于 802.11 的定位系统的指纹选择算法,减小了移动设备存储的数据量,其实验结果也表明,移动设备终端平均只需存储 12 Kbytes 指纹数据即可,但移动设备的指纹数据只存储了移动设备附近的局部信息,随着移动设备位置的变化会发生变化,需要不断更新,从服务器获取新的指纹数据,这极大地加重了移动设备的通信负载.

本文提出的方法其计算的主要工作在离线阶段,在移动设备只需存储离线阶段训练好的位置预测模型,执行常数时间复杂度的计算,就能得到移动设备的位置信息,因而对移动设备的计算和存储能力要求不高.在我们的实验中,若训练集大小为 5 000,核函数为高斯核函数,锚节点的数量为 20,则训练所得到的位置预测回归模型的大小为 81 KBytes,这远远小于整个无线信号地图的容量,而且无需随着移动设备的移动和服务器通信进行更新.

4 结论以及未来工作

在室内环境中,通过定位提供移动设备的准确位置信息,对普适计算的各类应用具有重要的意义,可使得普适计算应用系统根据对象位置为对象提供准确、及时的服务.本文针对室内环境复杂、信号容易存在漏检等问题,提出了基于支持向量回归的多次测量定位法,通过离线阶段的数据过滤和在线阶段的数据预处理提高数据的质量,并通过实验对该定位方法进行了综合对比分析.实验结果表明:相对于经典的概率定位法以及神经网络法,该方法定位精度更高;同时,对移动设备的存储容量以及计算能力要求低.

从实验中可知:由于各锚节点对模型的重要程度不一样,有时候锚节点数量的增多并没有提高定位模型的定位精度.因而,当把合适数量的锚节点部署在合适的地点时,对定位模型乃至整个系统都具有重要意义.未来的研究重点将集中于分析锚节点与定位模型定位精度之间的关系.

References:

- [1] Wang YC, Zhao JH, Fukushima T. LOCK: A highly accurate, easy-to-use location-based access control system. In: Choudhury T, *et al.*, eds. Proc. of the LoCA 2009. LNCS 5561, Berlin, Heidelberg: Springer-Verlag, 2009. 254–270. [doi: 10.1007/978-3-642-01721-6_16]
- [2] Zheng VW, Cao B, Zheng Y, Xie X, Yang Q. Collaborative filtering meets mobile recommendation: A user-centered approach. In: Proc. of the 24th AAAI Conf. on Artificial Intelligence (AAAI 2010). AAAI, 2010. 236–241. <http://www.aaai.org/ocs/index.php/AAAI/AAAI10/paper/viewFile/1615/1964>
- [3] Kizhakkepurayil S, Oh JY, Lee Y. Mobile application for healthcare system—Location based. In: Sobhet T, *et al.*, eds. Innovations and Advances in Computer Sciences and Engineering. Springer-Verlag, 2010. 297–302. [doi: 10.1007/978-90-481-3658-2_51]
- [4] Kaplan E, Hegarty C, eds. Understanding GPS: Principles and Applications. 2nd ed., Artech House, Incorporated, 2005. 1–19.
- [5] Liu H, Darabi H, Banerjee P, Liu J. Survey of wireless indoor positioning techniques and systems. IEEE Trans. on Systems, Man, and Cybernetics—Part C: Applications and Reviews, 2007,37(6):1067–1080. [doi: 10.1109/TSMCC.2007.905750]
- [6] Bahl P, Padmanabhan VN. RADAR: An in-building RF-based user location and tracking system. In: Proc. of the INFOCOM 2000. IEEE, 2000. 775–784. [doi: 10.1109/INFCOM.2000.832252]
- [7] Bhasker ES, Brown SW, Griswold WG. Employing user feedback for fast, accurate, low-maintenance geolocation. In: Proc. of the 2nd IEEE Annual Conf. on Pervasive Computing and Communications (PERCOM 2004). IEEE Computer Society, 2004. 111–120. [doi: 10.1109/PERCOM.2004.1276850]
- [8] Prasithsangaree P, Krishnamurthi P, Chrysanthis PK. On indoor position location with wireless LANs. In: Proc. of the IEEE Int'l Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC 2002). IEEE, 2002. 720–724. [doi: 10.1109/PIMRC.2002.1047316]
- [9] Kontkanen P, Myllymaki P, Roos T, Tirri H, Valtonen K, Wettig H. Topics in probabilistic location estimation in wireless networks. In: Proc. of the IEEE Int'l Symp. on Personal, Indoor and Mobile Radio Communications (PIMRC 2004). IEEE, 2004. 1052–1056. [doi: 10.1109/PIMRC.2004.1373859]
- [10] Mehmood H, Tripathi NK, Tipdecho T. Indoor positioning system using artificial neural network. Journal of Computer Science, 2010,6(10):1206–1212.
- [11] Smola A. Regression estimation with support vector learning machines. Technical Report, Technical University of Munich, 1996. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.10.3628&rep=rep1&type=pdf>
- [12] Smola AJ, Schölkopf B. A tutorial on support vector regression. Statistics and Computing, 2004,14(3):199–222. <http://eprints.pascal-network.org/archive/00000856/01/fulltext.pdf> [doi: 10.1023/B:STCO.0000035301.49549.88]
- [13] Battiti R, Brunato M, Villani A. Statistical learning theory for location fingerprinting in wireless LANs. Technical Report, DIT-02-0086, Università di Trento, 2002. <http://eprints.biblio.unitn.it/238/1/86.pdf>
- [14] Nguyen XL, Jordanyz MI, Sinopoli B. A kernel-based learning approach to ad hoc sensor network localization. ACM Trans. on Sensor Networks, 2005,1(1):134–152. [doi: 10.1145/1077391.1077397]
- [15] Tran DA, Nguyen T. Localization in wireless sensor networks based on support vector machines. IEEE Trans. on Parallel and Distributed Systems, 2008,19(7):981–994. [doi: 10.1109/TPDS.2007.70800]
- [16] Feng VS, Wang TC, Chang SY, Ma HP. Location estimation in indoor wireless networks by hierarchical support vector machines with fast learning algorithm. In: Proc. of the 2010 Int'l Conf. on System Science and Engineering (ICSSE 2010). IEEE, 2010. 321–326. [doi: 10.1109/ICSSE.2010.5551780]
- [17] Wu ZL, Li CH, Ng JKY, Leung KRPH. Location estimation via support vector regression. IEEE Trans. on Mobile Computing, 2007,6(3):311–321. [doi: 10.1109/TMC.2007.42]
- [18] Sklar B. Rayleigh fading channels in mobile digital communication systems, Part I: Characterization. IEEE Communications Magazine, 1997,35(7):90–100. [doi: 10.1109/35.601747]

- [19] Kaemarungsi K, Krishnamurthy P. Analysis of WLAN's received signal strength indication for indoor location fingerprinting, *Pervasive and Mobile Computing*, 2012,8(2):292-316. [doi: 10.1016/j.pmcj.2011.09.003]
- [20] King T, Haenselmann T, Effelsberg W. On-Demand fingerprint selection for 802.11-based positioning systems. In: *Proc. of World of Wireless, Mobile and Multimedia Networks (WoWMoM 2008)*. IEEE, 2008. 1-8. [doi: 10.1109/WOWMOM.2008.45948 39]



石柯(1973—),男,河南社旗人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为无线自组织与传感器网络,嵌入式系统,智能控制,并行分布式计算.

E-mail: keshi@mail.hust.edu.cn



陈洪生(1981—),男,博士生,讲师,主要研究领域为无线传感器网络.

E-mail: chenhs1981@163.com



张仁同(1986—),男,硕士,主要研究领域为无线传感器网络.

E-mail: 844618821@qq.com

www.jos.org.cn