

基于核直接判别分析和支持向量回归的 WLAN 室内定位算法

徐玉滨 邓志安* 马琳

(哈尔滨工业大学电子与信息工程学院 哈尔滨 150001)

摘 要: 该文针对 RSS 信号的时变性降低 WLAN 室内定位精度的问题, 提出了一种新的基于核直接判别分析和支持向量回归的定位算法。该算法利用核直接判别分析对原始 RSS 信号进行定位信息重组, 去除冗余定位特征和噪声, 提取最具判别力的定位特征, 然后采用支持向量回归算法建立定位特征与物理位置的映射关系。实验结果表明, 提出算法的定位精度明显高于传统定位算法, 且大大降低了离线阶段数据采集的工作量。

关键词: 无线局域网; 室内定位; 核直接判别分析; 支持向量回归

中图分类号: TP393.17

文献标识码: A

文章编号: 1009-5896(2011)-04-0896-06

DOI: 10.3724/SP.J.1146.2010.00813

WLAN Indoor Positioning Algorithm Based on KDDA and SVR

Xu Yu-bin Deng Zhi-an Ma Lin

(School of Electronics and Information Engineering, Harbin Institute of Technology, Harbin 150001, China)

Abstract: The time-varying Received Signal Strength (RSS) drastically degrades the indoor positioning accuracy in Wireless Local Area Network (WLAN). A new positioning algorithm based on Kernel Direct Discriminant Analysis (KDDA) and Support Vector Regression (SVR) is proposed to resolve the problem in this paper. The proposed algorithm employs KDDA to reconstruct the localization information contained in the RSS signal. The most discriminative localization features are then extracted while the redundant localization features and noise are discarded by KDDA. The extracted localization features are taken as inputs to SVR learning machine and the mapping between localization features and physical locations is established. The experimental results show that the proposed algorithm obtains significant accuracy improvement while requiring a much smaller set of RSS training data than previous methods.

Key words: Wireless Local Area Network (WLAN); Indoor positioning; Kernel Direct Discriminant Analysis (KDDA); Support Vector Regression (SVR)

1 引言

近年来, 人们对室内位置信息服务如个性化信息传递、医疗服务、目标发现等的需求与日俱增^[1]。由于传统的 GPS 和蜂窝网定位技术无法满足室内定位的要求, 基于接收信号强度(RSS)的无线局域网(WLAN)室内定位系统受到越来越广泛的关注^[2]。该定位系统仅由现有的广泛存在的 WLAN 架构和移动终端即可实现, 不需要任何额外的专用设备, 而且其较高的定位精度能满足复杂室内环境下大多数位置信息服务的要求。

目前最普遍的 WLAN 室内定位系统是基于位置指纹定位架构的^[3]。该方法分为离线和在线两个阶段。离线阶段, 在选定的参考点位置上采集来自各个接入点(Access Point, AP)的 RSS 值, 以建立位

置指纹数据库即所谓的无线电地图; 在线阶段, 将实时采集的 RSS 信号与无线电地图比较, 匹配出用户位置。位置指纹定位法可以被看作是典型的模式识别过程, 目前已有部分模式识别算法^[3-5]应用于 WLAN 室内定位。

在实际 WLAN 室内定位环境中, 由于多径、阴影效应、人员走动等影响^[6], 固定位置上来自各个 AP 的 RSS 值往往表现出复杂的时变统计特性, 降低定位精度和增加系统复杂度。一方面, 直接采用各个 AP 的 RSS 值为输入的定位特征分量, RSS 信号的时变性将严重影响定位精度。首先是因为方差较大, 位置分辨能力较弱的 AP 的 RSS 信号将引入大量的噪声; 其次不同 AP 的定位特征分量的相关性造成输入特征空间的冗余。另一方面, 为获得 RSS 信号的时变统计特性, 离线阶段需要耗费大量的人工时间在每个参考点上采集多个 RSS 样本, 不利于定位系统的大规模部署。针对输入定位特征, 文献[7,8]中分别采用不同的方法选择最具判别力的 AP

2010-08-05 收到, 2010-11-01 改回

国家 863 计划项目(2008AA12Z305)资助课题

*通信作者: 邓志安 dengzhianan@163.com

对应的 RSS 信号作为输入定位特征。由于未选中 AP 所对应的 RSS 定位信息都将被抛弃，将降低定位精度，尤其是当可利用的 AP 个数较少的时候。文献 [9] 利用主成份分析 (Principal Component Analysis, PCA) 进行定位特征提取，去除定位特征分量之间的相关性。但是 PCA 的线性特性限制了非线性定位特征的提取，而且 PCA 无法有效提取有判别力的定位特征。

本文提出基于核直接判别分析 (KDDA) 和支持向量回归 (SVR) 的 WLAN 室内定位算法 (KDDA-SVR)。KDDA 可以有效提取非线性定位特征。而且，与其它特征提取方法 [10] 不同的是，KDDA 按照判别力大小重组 RSS 定位信息，从中提取最具判别力的定位特征，同时将判别力小的冗余定位特征当作噪声去除。判别力即定位特征区分不同物理位置的能力，是影响 WLAN 指纹定位精度的关键因素。基于 SVR 在解决有限样本的非线性回归问题的特有优势 [11]，结合 SVR 学习和建立定位特征与物理位置的映射关系。实验结果表明，本文算法在明显提高定位精度的同时，大大减少了参考点数据采集的工作量。

2 基本理论

2.1 KDDA 变换基本原理

KDDA [10] 算法是 Fisher 线性判别分析 (Linear Discriminant Analysis, LDA) 的核推广形式。LDA 作为一种常用的特征提取方法，被广泛应用于多类分类问题中，以寻找能最优化区分不同类别的特征分量。核方法为许多标准的线性变换提供了有效的非线性拓展 [11]，因此 KDDA 可以有效提取 RSS 信号的非线性定位特征。

为了挖掘非线性定位特征，KDDA 将原始 RSS 信号 \mathbf{r} 映射至高维非线性空间： $\phi: \mathbf{r} \in \mathbf{R}^d \rightarrow \phi(\mathbf{r}) \in \mathbf{F}$ ，其中 \mathbf{R}^d 表示由 d 个 AP 组成的 RSS 信号空间。然后 KDDA 以 Fisher 判别式为优化准则，在核空间 \mathbf{F} 中寻找最具判别力的基向量 $\psi' = [\psi'_1, \dots, \psi'_M]$ ：

$$\psi' = \arg \max_{\psi} \left(\frac{\text{tr}(\psi^T \mathbf{S}_{\text{BTW}} \psi)}{\text{tr}(\psi^T \mathbf{S}_{\text{WTH}} \psi)} \right) \quad (1)$$

式 (1) 中 \mathbf{S}_{BTW} 和 \mathbf{S}_{WTH} 分别为类间和类内离散度矩阵：

$$\mathbf{S}_{\text{BTW}} = \frac{1}{C} \sum_{i=1}^C (\bar{\phi}_i - \bar{\phi})(\bar{\phi}_i - \bar{\phi})^T \quad (2)$$

$$\mathbf{S}_{\text{WTH}} = \frac{1}{L} \sum_{i=1}^C \sum_{t=1}^n (\phi(\mathbf{r}_i(t)) - \bar{\phi}_i)(\phi(\mathbf{r}_i(t)) - \bar{\phi}_i)^T \quad (3)$$

其中 $\bar{\phi}_i = 1/n \cdot \sum_{t=1}^n \phi(\mathbf{r}_i(t))$ ， $\bar{\phi} = 1/C \cdot \sum_{i=1}^C \bar{\phi}_i$ 分别为参考点 L_i 上对应 C_i 类的均值中心和所有类别的

均值中心，设 $\mathbf{r}_i(t)$ 和 n 分别为参考点 L_i 上的第 t 个 RSS 向量样本和样本数目， C 为整个定位环境中总的参考点数目， $L = n \cdot C$ 为 RSS 样本总数。

2.2 SVR 统计学习理论

SVR 算法以统计学习的 VC (Vapnik and Chervonenkis) 维 [11] 理论为基础。学习机器的实际风险 $R(w)$ 的上界由经验风险值 $R_{em}(w)$ 和置信范围值 $\phi(n/h)$ 两部分决定：

$$R(w) \leq R_{em}(w) + \phi(n/h) \quad (4)$$

其中 w 表示学习函数的参数， $\phi(n/h)$ 是单调递减函数， h 表示学习函数的 VC 维， n 为训练样本数。传统学习机器 (如神经网络) 只强调了训练样本的经验风险最小化，没有最小化置信范围值，其泛化能力较差，很可能出现对训练样本误差很小，而实际测试样本误差很大的情况，即所谓的过拟合问题。SVR 以经验风险值为优化问题的约束条件，置信范围值最小化作为优化目标，从而使实际风险的上界最小化。特别是对于训练样本有限的情况下，SVR 能有效防止过拟合的情况，其泛化能力明显优于传统的学习机器。

3 KDDA-SVR 定位算法

如图 1 所示，给出 KDDA-SVR 定位算法流程。离线阶段，在定位区域选定参考点，每个参考点采集一定数量的 RSS 样本，以构建无线电地图。通过 KDDA 变换，从原始 RSS 样本中提取定位特征，构成定位特征数据库。以定位特征和相应的参考点位置为输入训练样本对，完成 SVR 学习训练过程，得出定位特征与物理位置的映射关系。在线阶段，实时采集 RSS 信号，对 RSS 信号进行 KDDA 变换，然后输入已构建好的支持向量回归函数，估计出用户的实际位置。

3.1 KDDA 定位特征提取

输入特征的判别力是影响学习机器性能的关键因素 [12]。由于 RSS 信号的时变性，直接采用各个

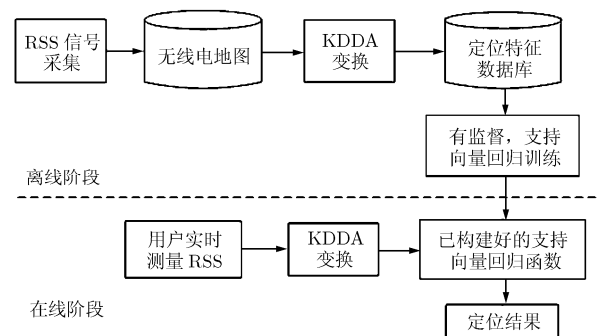


图 1 KDDA-SVR 定位算法流程

AP 的 RSS 值为输入定位特征, 将引入方差较大, 位置分辨力较弱的冗余特征和噪声。本文采用 KDDA 进行定位特征提取, 以减小 RSS 信号时变性的影响, 增强输入定位特征的判别力。

KDDA 变换以式(1)为优化目标在核空间 \mathbf{F} 中寻找最具判别力的基向量 ψ' , 然后在基向量 ψ' 的投影为输入定位特征。式(1)中的分子度量定位特征对不同参考点位置的区分能力, 即类间离散度; 而式(1)中的分母度量定位特征的平均方差, 即类内离散度。如图 2 所示, 满足式(1)的最优基向量 ψ' 所提取的特征使得类间离散度最大化, 类内离散度最小化。无线电地图中的每个参考点和对应的定位特征可以被看作是一个类别。KDDA 变换所提取的定位特征能最大化区分不同类别, 具有最优判别定位能力。

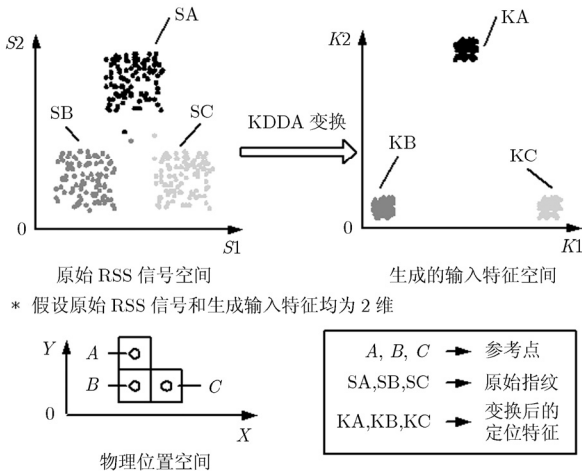


图 2 KDDA 变换对 RSS 信号进行定位特征提取

KDDA 变换的具体实现是通过 \mathbf{S}_{BTW} 的非零空间和 \mathbf{S}_{WTH} 的零空间交集来寻找最优化的基向量 ψ' [10]。首先, 计算得出 \mathbf{S}_{BTW} 的按降序排列的 C 个特征值 $[\lambda_1 \cdots \lambda_C]$ 和对应特征向量 $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_C]$, 保留其中最具判别力的 m 个特征向量: $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_m]$, 使得

$$\mathbf{V}^T \mathbf{S}_{\text{BTW}} \mathbf{V} = \mathbf{\Lambda}_b \quad (5)$$

其中 $\mathbf{\Lambda}_b = \text{diag}[\lambda_1 \cdots \lambda_m]$ 是一个 $m \times m$ 对角阵。然后求 \mathbf{S}_{WTH} 的零空间与 \mathbf{S}_{BTW} 的非零空间交集。设 $\mathbf{U} = \mathbf{V} \mathbf{\Lambda}_b^{-1/2}$, 使得 $\mathbf{U}^T \mathbf{S}_{\text{BTW}} \mathbf{U} = \mathbf{I}$, 保留矩阵 $\mathbf{U}^T \mathbf{S}_{\text{WTH}} \mathbf{U}$ 的 $M (\leq m)$ 个最小特征值 $[\lambda'_1 \cdots \lambda'_M]$ 对应的特征向量 $\mathbf{P} = [\mathbf{p}_1 \cdots \mathbf{p}_M]$ 。令矩阵 $\mathbf{Q} = \mathbf{U} \mathbf{P}$ 得

$$\mathbf{Q}^T \mathbf{S}_{\text{WTH}} \mathbf{Q} = \mathbf{\Lambda}_w \quad (6)$$

其中 $\mathbf{\Lambda}_w = \text{diag}[\lambda'_1 \cdots \lambda'_M]$, 是一个 $M \times M$ 对角阵。 M 个最优判别基向量为 $\psi' = \mathbf{Q} \mathbf{\Lambda}_w^{-1/2} = [\psi'_1 \cdots \psi'_M]$ 。对于任意的输入 RSS 向量 \mathbf{r} 的 KDDA 变换, 即它

在最优判别基向量上的投影为

$$\mathbf{x} = (\psi')^T \cdot \phi(\mathbf{r}) = \mathbf{\Theta} \cdot \gamma(\phi(\mathbf{r})) \quad (7)$$

其中 \mathbf{x} 为生成的 M 维定位特征, $\mathbf{\Theta}$ 为 $M \times L$ 变换矩阵 [10]。 $\gamma(\phi(\mathbf{r}))$ 为

$$\gamma(\phi(\mathbf{r})) = [k(\mathbf{r}_1(1), \mathbf{r}) \ k(\mathbf{r}_1(2), \mathbf{r}) \cdots k(\mathbf{r}_C(n), \mathbf{r})]^T \quad (8)$$

其中 $k(\cdot, \cdot)$ 为 KDDA 变换的核函数。

3.2 定位特征的维数选取

定位特征的维数即 KDDA 变换的参数 M , 且参数 $M \leq m$ 。参数 m 和 M 的大小直接影响 KDDA 变换对 RSS 定位信息的重组和定位特征的提取。 m 表示类间离散度矩阵 \mathbf{S}_{BTW} 保留 m 个最大特征值对应的特征向量。RSS 信号的类间离散度信息按每个特征向量方向 \mathbf{v}_i 重组, 大小由对应特征值 λ_i 度量。特征向量对应特征值越大, 该方向对应类间离散度越大, 也就意味着所提取定位特征使得参考点之间的分辨能力越强。参数 m 对应的总类间离散度信息量 ρ 可以表示为

$$\rho(m) = \sum_{i=1}^m \lambda_i / \sum_{i=1}^C \lambda_i \geq \eta \quad (9)$$

其中 $\{\lambda_1, \cdots, \lambda_C\}$ 为 \mathbf{S}_{BTW} 对应的特征值, η 为累积特征值比例阈值。参数 m 太小的话, 没有获取足够的类间离散度信息, 将降低定位精度; 参数 m 太大的话, 将引入对应特征值较小的, 几乎不含类间离散度信息的冗余定位特征和噪声, 容易产生学习机器的过拟合问题。通过离线阶段的训练可以发现, η 一般处于 85% 到 95% 之间时, 满足 $\rho(m) \geq \eta$ 的最小整数 m^* 为最佳值。

同理, 参数 M 表示保留矩阵 $\mathbf{U}^T \mathbf{S}_{\text{WTH}} \mathbf{U}$ 对应 M 个最小特征值的特征向量。特征向量对应特征值越小, 该方向对应类内方差越小。通过调整参数 M , 对应特征值较大的特征向量由于含有较大的方差, 被当作噪声去除。因此, 输入定位特征维数可以通过离线阶段的训练, 寻找最佳的累积特征值比例阈值而确定。

3.3 SVR 训练

本文利用 ϵ -不敏感 SVR [11] 算法来构造 RSS 定位特征与物理位置的非线性映射关系。给定训练数据 $(\mathbf{x}_i, z_i), i = 1 \cdots L, \mathbf{x}_i \in R^M$, \mathbf{x}_i 为第 i 个输入定位特征样本, z_i 为对应的输出位置坐标。通过非线性映射 ϕ 将输入定位特征映射到一个高维非线性空间, 然后在该空间构造回归函数:

$$z = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + b \quad (10)$$

其中 b 为偏置, \mathbf{w} 为权重系数。根据结构风险最小化原则, 它要解决的最优化问题形式为

$$\begin{aligned} \min J(\mathbf{w}) &= C' \sum_{i=1}^l (\xi_i + \xi_i^*) + \frac{1}{2} \|\mathbf{w}\|^2 \\ \text{s.t.} \quad &\begin{cases} (\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) - z_i \leq \varepsilon + \xi_i \\ z_i - (\langle \mathbf{w}, \phi(\mathbf{x}_i) \rangle + b) \leq \varepsilon + \xi_i^* \\ \xi_i^* \geq 0, \xi_i \geq 0, i = 1, 2, \dots, l \end{cases} \end{aligned} \quad (11)$$

其中 $\sum (\xi_i + \xi_i^*), i = 1 \dots L$ 体现经验风险, $\|\mathbf{w}\|^2$ 控制 VC 维的大小, C' 是经验风险与 VC 维的一个平衡参数。给定参数 ε, C' 和核函数参数, 对式(11)进行凸二次优化, 可以得出 \mathbf{w} 的全局唯一最优解。输出位置坐标有两维, 需要分别独立训练和建立对应的两个 SVR 函数。SVR 函数可以表达为

$$z = \sum_{i \in \text{SV}} \alpha_i k_1(\mathbf{x}, \hat{\mathbf{x}}_i) + b \quad (12)$$

其中 $\hat{\mathbf{x}}_i$ 为支持向量, α_i 为对应的系数, $k_1(\cdot, \cdot)$ 为 SVR 的核函数。

3.4 核函数选取

KDDA 变换和 SVR 学习机器都涉及到核函数的选取。合适的核函数选取意味着对于要解决的实际问题先验知识的合理编码利用^[11]。本文核函数的选取主要由 RSS 信号的统计分布特性和非线性决定。固定位置的来自某个 AP 的 RSS 值往往呈现高斯或类高斯分布, 且变化范围最高可达 10 dB。不同位置的 RSS 信号随着与 AP 距离的变化而非线性变化, 尤其在走廊拐角处或者有障碍物时, RSS 值将发生剧烈变化。高斯核是最常用的核函数之一, 已经被大量应用在模式识别问题中。鉴于高斯核良好的局部特征提取能力和平滑特性^[13], 可以有效挖掘 RSS 信号与物理位置的非线性关系, KDDA 变换和 SVR 学习机器都选择高斯核函数:

$$\hat{k}(\mathbf{y}_1, \mathbf{y}_2) = \exp\left(-\|\mathbf{y}_1 - \mathbf{y}_2\|^2 / 2\hat{\sigma}^2\right) \quad (13)$$

其中 $\mathbf{y}_1, \mathbf{y}_2$ 表示输入向量, $\hat{\sigma}$ 为高斯核宽度。

4 实验结果与分析

4.1 实验设置

实验环境如图 3 所示为典型的 WLAN 室内定位场景。阴影部分表示定位区域, 包括走廊和房间的一部分, 房间内布置有桌椅, 书柜等办公物品。定位区域可以接收 5 到 9 个 AP 的 RSS 信号, 大多数无线 AP 的信道环境为非视距的, 且存在墙体阻隔。在定位区域每隔 2 m 选取共计 16 个参考点。为了充分测试各个算法的定位性能, 除参考点之外随机选取 34 个测试点。利用终端驱动和我们开发的 RSS 信号采集软件, 在每个位置点采集 100 个 RSS 向量样本, 2 个样本/秒, 则共有 1600 个训练样本和 3400 个测试样本。为了设置 KDDA-SVR 算法的参数,

采用 10 倍交叉验证法^[7]: 将训练样本集分成 10 份, 轮流将其中 9 份用于试探性建立 RSS 定位模型, 剩下 1 份作为验证集合用于估计所建定位模型的性能; 平均 10 次结果作为算法的定位性能, 获得好的定位性能的参数设置为实际参数值。本文分别实现了 LDA 和 PCA^[9]特征提取结合 SVR 算法(LDA-SVR 与 PCA-SVR), 传统的 SVR^[4]算法, 人工神经网络^[14](Artificial Neural Network, ANN)和加权 K 近邻法(Weighted K-Nearest Neighbor, WKNN)^[5]。SVR 采用高斯核; ANN 采用隐藏层神经元数目为 12 的 3 层误差反向传播结构, WKNN 的近邻数 K 设置为 5。

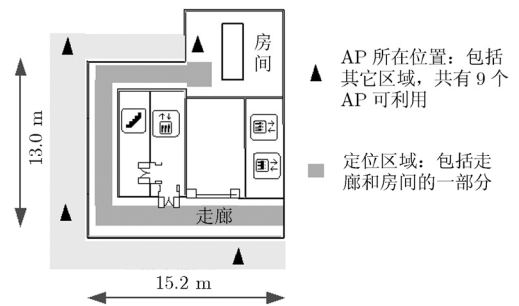


图 3 WLAN 室内定位实验环境

4.2 实验结果分析

如图 4 所示为参考点训练样本数目为 100 时, 各个算法在测试点的平均定位精度, 即定位误差累积概率分布。定位误差为 2 m 时, KDDA-SVR, LDA-SVR, PCA-SVR, SVR, ANN 和 WKNN 算法的定位精度分别为 71.2%, 64.9%, 59.8%, 57.3%, 49.2% 和 45.3%。定位误差为 3 m 时, 各个算法的定位精度依次为 88.4%, 82.9%, 80.2%, 78.5%, 72.1% 和 69.5%。本文所提出的 KDDA-SVR 算法的定位精度明显高于其它定位算法。LDA 定位特征提取对 SVR 算法有一定改进, 但是由于其本质的线性特性, 无法有效提取 RSS 信号的非线性定位特征, 限制了定位精度的进一步提高。PCA 定位特征提取对 SVR 定位算法的改进非常有限, 原因是 PCA 所提取的主成分方向可能是由于某个方向的 RSS 信号不确定性或者是噪声强度较大导致的, 而不是按判别力大小重组和提取定位特征。SVR 相对 ANN 有更优良的学习机器泛化能力, 可以避免 ANN 容易陷入局部最优的缺点, SVR 的定位精度高于 ANN。由于 WKNN 是确定性方法, 没有充分利用 RSS 信号的统计信息, 定位精度最低。

如图 5 所示为不同的参考点采样样本数情况下, 各个定位算法的平均定位误差比较。在采样样

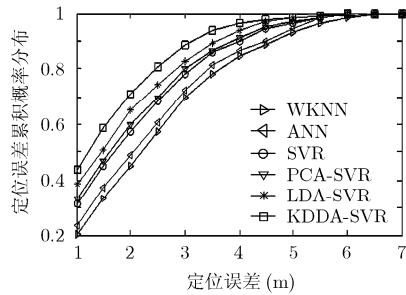


图 4 各个算法不同定位误差的累积概率分布

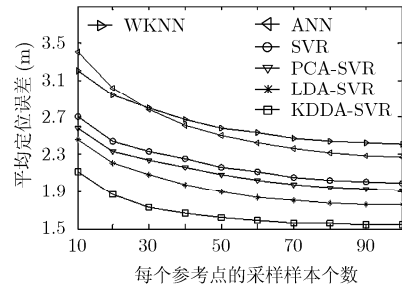


图 5 平均定位误差随着采样样本个数的变化情况

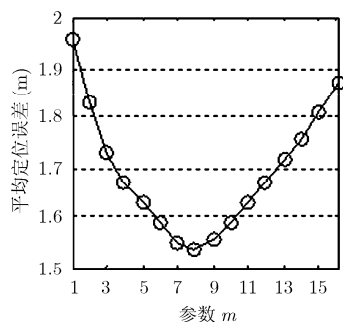
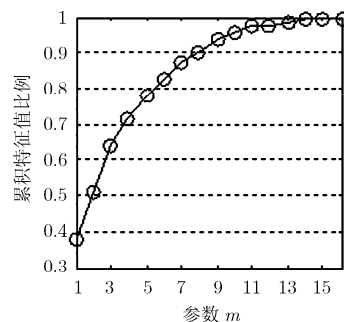
本数为 30 时, KDDA-SVR 算法的平均定位误差为 1.72 m, 仍然明显低于采样样本数为 100 时的其它所有算法, LDA-SVR, PCA-SVR, SVR, ANN 和 WKNN 算法平均定位误差依次为 1.75 m, 1.92 m, 1.99 m, 2.27 m, 和 2.41 m。当采样样本数为 10 时, KDDA-SVR 算法的平均定位误差为 2.14 m, 要想获得同等定位精度, SVR 算法的采样样本数至少需达到 60。同时可以看出, SVR 算法定位精度随着采样样本数的减少而下降的速度要明显小于 ANN 算法, 主要原因是 SVR 算法基于结构风险最小化原理, 对有限样本的非线性回归问题, 能较好地避免过拟合问题。因此, 在明显提高定位精度的同时, KDDA-SVR 算法可以大大减少离线阶段数据采集的工作量, 从而有利于定位系统的大规模部署应用。

如图 6 所示为 KDDA 变换中的参数 m 对 KDDA-SVR 算法平均定位误差的影响。参数 m 表示保留 S_{BTW} 的 m 个最大特征值对应的特征向量。特征向量对应特征值越大, 该方向对应类间距离越大, 也就意味着所提取特征越具有判别力。如图 6(a) 所示参数 m 为 8 时平均定位误差最低, 此时如图 6(b) 所示 $\rho = 90.5\%$, 即 90.5% 的类间离散度信息被保留, 而剩余的判别定位能力很弱的 9.5% 定位信息被当作冗余噪声去掉。参数 m 开始增加时, 随着所获取类间离散度信息的增加, 增强了定位特征的判别力, 平均定位误差逐渐减小; 当 $m=8$ 时到达拐点,

此后随着 m 增加所增加的特征向量的判别定位能力越来越弱, 且引入噪声和冗余定位特征, 平均定位误差逐渐增加。所有参数 M 设置为 m 的对应最佳值, 参数 $m=8$ 时, $M=5$ 。同理, 调整参数 M , 矩阵 $U^T S_{WTH} U$ 较大特征值对应的特征向量由于含有较大的方差, 被当作噪声去除。从 SVR 学习机器泛化能力的角度看, 输入特征的判别力和维数大小是两个关键因素。KDDA 变换可以有效提取最具判别力的定位特征, 增强 SVR 学习机器输入特征的判别力; 去除冗余定位特征和噪声, 降低 SVR 学习机器输入特征的维数。因此, KDDA 变换可以明显提高 SVR 学习机器的泛化能力, 从而提高定位精度。

5 结束语

本文针对 RSS 信号的时变性引入冗余特征和噪声的问题, 提出了一种新的基于 KDDA 和 SVR 的 WLAN 室内定位算法。该算法采用 KDDA 对原始 RSS 信号进行定位信息重组, 从中提取最具判别力的定位特征, 去除冗余定位特征和噪声。基于 SVR 在有限样本非线性回归的特有优势, 结合 SVR 学习和建立定位特征与物理位置的映射关系。KDDA 变换有效增强了输入定位特征的判别力, 降低了输入定位特征维数, 从而明显提高了 SVR 学习机器的泛化能力。实验结果表明, 本文提出的 KDDA-SVR 算法明显提高了定位精度, 且参考点所需采样样本

(a) 平均定位误差随着参数 m 的变化情况(b) 累积特征值比例随着参数 m 的变化情况图 6 KDDA 变换的参数 m 与定位性能的关系

数大大减少,降低了人工数据采集工作量。采样样本数都为100时,KDDA-SVR算法的平均定位误差比传统的SVR和WKNN算法分别降低22.6%(0.45 m)和36.1%(0.87 m)。采样样本数减少到30,KDDA-SVR算法仍然比SVR和WKNN算法样本数为100时,平均定位误差降低13.6%(0.27 m)和28.6%(0.69 m)。

参考文献

- [1] Munoz D, Bouchereau F, and Vargas C, *et al.* Position Location Techniques and Applications[M]. Burlington, USA, Academic Press, 2009: 17–21.
- [2] Gu Y Y, Lo A, and Niemegeers I. A survey of indoor positioning systems for wireless personal networks[J]. *IEEE Communications Surveys & Tutorials*, 2009, 11(1): 13–32.
- [3] Mengual L, Marbán O, and Eibe S. Clustering-based location in wireless networks[J]. *Expert Systems with Applications*, 2010, 36(4): 7552–7561.
- [4] Brunato M and Battiti R. Statistical learning theory for location fingerprinting in wireless LANs[J]. *Computer Networks*, 2005, 47(6): 825–845.
- [5] Bahl P and Padmanabhan V N. RADAR: An in-building RF-based user location and tracking system[C]. Proc. IEEE INFOCOM, Tel Aviv, Israel, 2000: 775–784.
- [6] Carlos K F, Inmaculada M J, and Alicia G C, *et al.* Nonparametric model comparison and uncertainty evaluation for signal strength indoor location[J]. *IEEE Transactions on Mobile Computing*, 2009, 8(9): 1250–1264.
- [7] Chen K Y, Yang Q, and Yin J, *et al.* Power-efficient access-point selection for indoor location estimation[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2006, 18(7): 877–888.
- [8] Kushki A, Plataniotis K, and Venetsanopoulos A. Intelligent dynamic radio tracking in indoor wireless local area networks[J]. *IEEE Transactions on Mobile Computing*, 2010, 9(3): 405–419.
- [9] Fang S H, Lin T N, and Lin P. Location fingerprinting in a decorrelated space[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2008, 20(5): 685–691.
- [10] Lu J W, Plataniotis K N, and Venetsanopoulos A N. Face recognition using kernel direct discriminant analysis algorithms[J]. *IEEE Transactions on Neural Networks*, 2003, 14(1): 117–126.
- [11] Vapnik V N. An overview of statistical learning theory[J]. *IEEE Transactions on Neural Networks*, 1999, 10(5): 988–999.
- [12] Theodoridis S and Koutroumbas K. Pattern Recognition[M]. 4th Edition. Burlington, USA, Academic Press, 2009: 213–253.
- [13] Keerthi S S and Lin C J. Asymptotic behaviors of support vector machines with Gaussian kernel[J]. *Neural Computation*, 2003, 15(7): 1667–1689.
- [14] Edgar A M, Raul C, and Jesus F. Estimating user location in a WLAN using backpropagation neural networks. Lecture Notes in Computer Science, Berlin, Germany, Springer-Verlag, 2004, 3315: 737–746.

徐玉滨：男，1954年生，教授，博士生导师，研究领域包括通信网及专用移动通信系统、导航定位、多媒体通信。

邓志安：男，1985年生，博士生，研究方向包括无线定位、人工智能、模式识别。

马琳：男，1980年生，博士，讲师，研究领域包括宽带无线传输、专用通信网络、无线定位技术。