

DOI: 10.19650/j.cnki.cjsi.J2108314

基于 DPC 指纹子空间匹配的室内 WiFi 定位方法^{*}

乐燕芬, 许远航, 施伟斌

(上海理工大学光电信息与计算机工程学院 上海 200093)

摘 要:针对无线接收信号强度 (RSS) 受传播环境突发噪声影响而引起指纹定位较大误差的问题, 本文提出了一种指纹子空间匹配结合密度峰值聚类 (DPC) 的定位算法, 有效避免大误差点。首先通过在线阶段目标 RSS 信号的接入点 (AP) 覆盖向量, 确定有效的参考位置点, 并划分多个指纹子空间, 利用改进的 WKNN 算法估计目标在每个子空间内的位置; 最后利用 DPC 算法选取决策值最大的 S 个估计位置确定目标。所提算法简单, 不需要离线阶段的学习过程训练定位模型, 尤其适合存在大量 AP 的大范围室内定位区域。实际环境中的定位实验表明, 基于 DPC 的指纹子空间匹配算法比 WKNN 算法的定位精度提升了 25% 左右, 且在参考点分布密度为 $1.8\text{ m} \times 1.8\text{ m}$ 的实验条件下基本消除了 4 m 以上的大定位误差, 有效提高了定位方法的整体性能。

关键词: 室内定位; WiFi 指纹; 指纹子空间匹配; DPC

中图分类号: TN92 TP391 TH70

文献标识码: A

国家标准学科分类代码: 510.40

WiFi fingerprint based indoor positioning with subspace matching and DPC

Le Yanfen, Xu Yuanhang, Shi Weibin

(School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: There is large error caused by the burst noise of the received signal strength (RSS) in the indoor fingerprint positioning. To address this issue, a subspace matching method combined with the algorithm of clustering by fast search and find of density peaks (DPC) is proposed. In this way, the large positioning error could be effectively avoided. Specifically, the coverage vector of the access points in the online RSS is used to select a subset of the reference points and the subspaces of the radio map. An improved weighted K-nearest neighbors (WKNN) approach is applied to achieve estimation. Then, the DPC algorithm is used to select S estimated positions with the largest decision values to determine the position of the target. The simple algorithm needs no learning process in the offline stage, which is especially suitable for large indoor areas with a lot of access points. Compared with the WKNN algorithm, experimental results show that the proposed method improves the positioning accuracy by about 25%. The large positioning error of more than 4 m is eliminated when the reference points is $1.8\text{ m} \times 1.8\text{ m}$. The overall positioning performance is improved effectively.

Keywords: indoor positioning; WiFi fingerprint; subspace matching; DPC

0 引 言

随着无线网络技术和智能设备的普及, 各类基于位置服务的应用也逐渐成为人们日常生活不可或缺的部分。对于位置信息的需求也不再局限于室外环境, 物联网、场馆管理、人群监测、公共服务等领域内的各类应用都使位置信息需求拓展到室内环境^[1-2]。由于全球定位

系统 (GPS) 信号在室内复杂环境下快速衰减, 无法实现有效定位, 国内外研究人员提出了多种不同技术以期实现室内环境中对用户位置信息的准确获取^[3-4]。目前广泛部署在建筑物内的无线局域网为基于位置指纹的定位技术提供了一种实施方便、成本低廉的方案, 也使得基于 WiFi 信号的 RSS (received signal strength, RSS) 指纹定位成为近年来的研究热点^[5-7]。

RSS 指纹定位的基本思想是无线信号传播的衰减特

收稿日期: 2021-07-26 Received Date: 2021-07-26

^{*} 基金项目: 国家自然科学基金 (51705324) 项目资助

性使得 RSS 信号在空间各个位置存在差异,而通常越邻近位置的 RSS 具有更大的相似性。通过在离线阶段采集监测区域内各参考位置点 (reference point, RP) 能接收的各个接入点 (access point, AP) 的 RSS 信号构建离线指纹库,在线阶段通过目标实时采集的 RSS 信号与指纹库内指纹 RSS 信号的匹配,估计目标与各参考位置的相对位置而获得定位。研究较早的是微软公司设计的 RADAR 系统,利用 RSS 信号之间的欧式距离,采用 k 近邻算法能得到房间级别的定位精度^[8];在此基础上又有了多种改进方案,采用不同的加权系数、优化各个近邻的权值或者消除 RSS 信号指数衰减带来的信号差值与物理距离的非线性来提高定位精度^[9-11]。而更多基于机器学习的算法也用于建立物理位置和 RSS 信号之间的映射模型,包括贝叶斯分类器、带正则化的线性回归模型^[12-13]、线性或非线性的支持向量机^[14]等,力求通过离线阶段指纹数据的训练,在存在多径效应、阴影效应、人员走动甚至 AP 关闭等复杂情况下,建立 RSS 信号分布的准确模型。上述基于位置指纹的定位方式,不管是基于指纹相似度匹配或是定位模型,要获得理想的定位效果,都基于在线采集的 RSS 信号与离线的指纹满足相似的分布模型这一前提。而离线或在线阶段的突发噪声都会改变 RSS 信号的空间分布,从而引起大的定位误差。

针对这一问题,本文利用目标在线接收的 RSS 信号,提出一种指纹子空间划分方案,利用多个子空间的位置估计来消除突发噪声的影响,并引入了密度峰值聚类算法 (clustering by fast search and find of density peaks, DPC) 来确定可信度最高的若干估计位置^[15],从而实现抑制大定位误差,提高定位性能的目的。

1 室内环境下 AP 信号的分布特性

在目前的实际室内环境中,通常布局有大量的 WiFi 接入点。一个确定位置点可检测到的 AP 数量在十几个到几十个不等,而在整个较大的环境内,可检测的接入点会多达几百个。如何选取携带有最多定位信息的接入点一直是定位研究的一个重要方面,包括随机选择、选择最强信号强度或最小信号方差的 AP,或者 Fisher 准则下的选择策略等,都试图从大量的可检测 AP 中寻找最利于定位的部分 AP,在降低离线或在线阶段定位算法计算复杂度的同时,期望保持甚至提高定位精度。

考虑到室内复杂环境下,尤其是墙壁遮挡等因素,使得 AP 的 RSS 信号的衰减呈现很强的空间性。只有处于同一局部空间的位置点,才可能接收到同一个 AP 子集。图 1(a) 分别用直线柱和虚直线柱给出了从实验指纹库随机选的 9[#] 和 180[#] AP 的分布情况,从图 1(a) 中可看出只有在 y 坐标 0.6 m, x 坐标在 35 m~50 m 附近的 10 个

位置点上能同时接受到这两个 AP 信号。而图 1(b) 给出了 80[#] AP 的分布情况。它与 9[#] AP 的信号分布无重叠,而能接收 8[#] AP 信号的全部参考位置点能同时接收到 180[#] AP 信号。因此本文提出的算法首先考虑到复杂室内环境中 AP 有限的信号覆盖范围,选取与目标能接收相似 AP 子集的参考点集作为可能的位置范围,在实现目标的粗定位的同时降低在线阶段需处理的指纹数目。

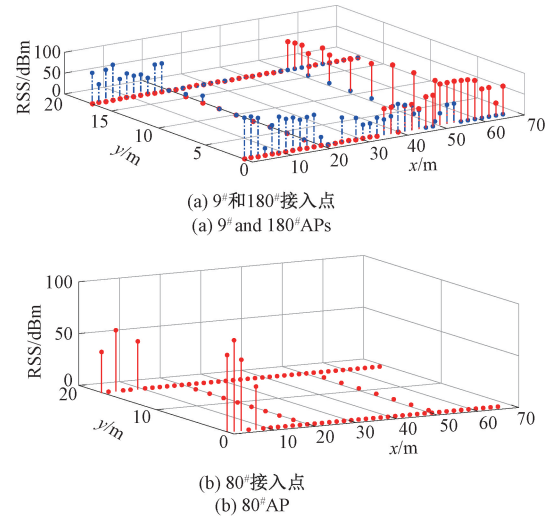


图 1 AP 的 RSS 信号分布

Fig. 1 Distribution of RSS signals from APs

此外,室内环境下人员的活动、接收设备的突然遮挡等各种干扰,使得 AP 信号并不呈现确定的传播衰减特性,信号强度常存在跳变。观察图 1(a) 中虚直线柱标识的 180[#] AP,在 y 坐标 0.6 m, x 坐标 20~30 m 附近的连续位置点存在信号断续传播的现象。图 1 中 9[#] 和 80[#] AP 也存在同样的情况。本文利用室内无线信号的局部传播模型对 AP 信号的路径损耗进行分析^[16]。 t 时刻从某确定 AP 到距离 d 处的信号损耗可表达为:

$$P(d)_t = P(d_0) + 10n\log\left(\frac{d}{d_0}\right) + P_{WF} + P_N + \Delta_t \quad (1)$$

式中: $P(d_0)$ 是在 d_0 处 (一般 $d_0 = 1$ m) 的路径损耗; $10n\log(d/d_0)$ 是从 AP 到接收位置点的传播损耗, n 是与具体传播环境相关的损耗因子; P_{WF} 是信号的穿墙/楼层损耗; P_N 为环境噪声,一般符合均值为 0 的高斯分布; Δ_t 表示 t 时刻的突发噪声。对于在同一参考位置点,不同时刻采集的同一 AP 的 RSS 信号的差值可表示为:

$$\Delta f(t_1, t_2) = f_{t1} - f_{t2} = P(d)_{t2} - P(d)_{t1} = (P_{N2} + \Delta_{t2}) - (P_{N1} + \Delta_{t1}) \quad (2)$$

式中:由于环境噪声符合 0 均值高斯分布,一般 RSS 信号的均值滤波处理可消除该噪声影响,则式 (2) 又可进一步表示为:

$$\Delta f(t_1, t_2) = \Delta_{t2} - \Delta_{t1} \quad (3)$$

从式(3)可看出,突发噪声会影响不同时刻采集的 RSS 信号,且由于其出现的随机性也很难完全消除对信号的影响。在实际定位过程中,为了避免突发噪声引起的对目标位置的模糊,不仅考虑能接收到信号的 AP 同时也随机提取部分不能接收信号的 AP 构建指纹子空间,来抑制离线或在线阶段突发噪声导致的 RSS 信号严重不匹配引起的重大定位误差。

2 基于 DPC 指纹子空间匹配的定位

本文所提的 DPC 指纹子空间匹配定位的算法框图如图 2 所示。算法在离线阶段完成整个定位区域 RSS 指纹的采集和处理,获得指纹数据库和每个 AP 的覆盖向量;在线阶段,目标获取 RSS 信号,并由此确定当前 AP 的覆盖向量,通过比较在线覆盖向量与指纹库 AP 覆盖向量的汉明距离,选取距离较小的若干 RP,作为后续所有指纹子空间的指纹坐标。通过这一匹配过程,使得目标定位在一定数量 RP 对应的局部空间内,可减少后续子空间的定位计算复杂度;通过固定 AP 和随机 AP 融合选择方式,构建 L 个子空间,并在每个子空间内利用改进的 WKNN (weighted K-nearest neighbors, WKNN) 算法估计目标位置;后续通过 DPC 算法从 L 个估计位置中选取决策值最大的 S 个确定目标的最终位置,该过程可有效排除由于 RSS 信号突发大噪声时引起的过大的定位误差。

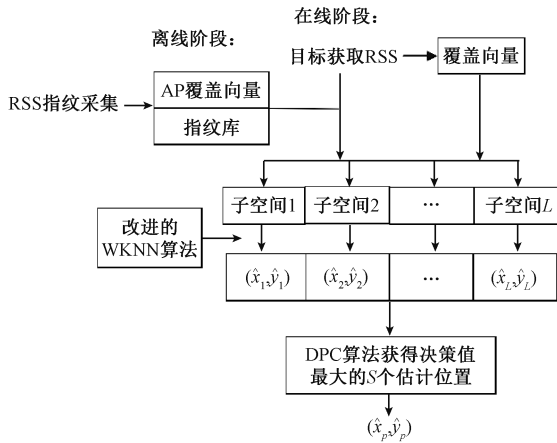


图 2 基于 DPC 指纹子空间匹配定位方法

Fig. 2 The positioning method with subspace matching and DPC

2.1 指纹子空间匹配定位

设离线阶段在整个定位区域可检测到 M 个 AP,每个 AP 由自身唯一的 MAC 地址标志并区分。指纹库里的每一条指纹可表示为:

$$\varphi_i = [\mathbf{P}_i \quad \mathbf{F}_i] = [x_i, y_i, f_{i1}, \dots, f_{iM}], \forall i \in \{1, 2, \dots, N\} \quad (4)$$

式中: $\mathbf{P}_i = [x_i, y_i]$ 表示第 i 条指纹也即第 i 个 RP 的空间位置,共 N 个参考点; $\mathbf{F}_i = [f_{i1}, f_{i2}, \dots, f_{iM}]$ 是第 i 个指纹, f_{ij} 表示在参考点 i 接收的第 j 个 AP 的 RSS 信号强度,常表示为采样时间内的信号均值;未接收到的 AP 信号则其 f_{ij} 为 0。根据指纹库可确定每个指纹的覆盖向量,第 i 个指纹的覆盖向量 $\mathbf{C}_i$ 表示为:

$$\mathbf{C}_i = [c_{i1}, c_{i2}, \dots, c_{iM}], \forall i \in \{1, 2, \dots, N\} \quad (5)$$

式中:若能接收到第 j 个 AP 的 RSS 信号,则 $c_{ij} = 1$,否则 $c_{ij} = 0$ 。覆盖向量 $\mathbf{C}$ 可以表示在某个参考点位置 AP 的分布情况,位置毗邻的参考点之间通常能共享一个较大的 AP 子集,换句话说,具有较相似覆盖向量的 RP,其物理位置通常较接近,这也是本文确定子空间参考点的依据。同时考虑到 $\mathbf{C}$ 是二进制数据类型,适合用汉明距离来表征两个向量的相似性。

在线定位时,设目标接收的 RSS 信号为:

$$\varphi_t = [f_{t1}, f_{t2}, \dots, f_{tM}] \quad (6)$$

同时也可确定相应的覆盖向量 $\mathbf{C}_t = [c_{t1}, c_{t2}, \dots, c_{tM}]$ 。为了确定目标所在的大致区域,计算 $\mathbf{C}_t$ 与每个指纹 $\mathbf{C}_i$ 的汉明距离 d_H ,

$$d_H(\mathbf{C}_t, \mathbf{C}_i) = \sum_{j=1}^M |c_{ij} - c_{tj}| \quad (7)$$

此距离越小,表明目标与该参考点能接收到的 AP 信号越一致,一般情况下其物理位置也较接近。通过式(7)对所有参考点根据汉明距离进行排序,并取前 N' 个最小的 d_H 所对应的参考点作为子空间的参考位置点 $\mathbf{P}' = \{\mathbf{P}'_1, \mathbf{P}'_2, \dots, \mathbf{P}'_{N'}\}$ 。

为了能有效利用目标接收的 RSS 信号,同时又避免因空间传播环境变化,如 AP 物理位置变动、AP 暂时关闭、新的 AP 加入或突发大量人员活动等情况引起的 RSS 信号突变引入的大定位误差,指纹子空间采用了固定 AP 与随机选择 AP 结合的方式。具体来说,第 l 个指纹子空间可表示为:

$$\Psi'_l = (\mathbf{P}' \quad \mathbf{F}'_l) = (\mathbf{P}' \quad \mathbf{F}'_m \quad \mathbf{F}'_s) \quad (8)$$

其中, $\mathbf{P}'$ 是选取的子空间参考位置点,均为 N' 个;子空间的指纹包含 $\mathbf{F}'_m$ 和 $\mathbf{F}'_s$ 两部分。 $\mathbf{F}'_m$ 作为子空间的固定 AP 部分,包含了从离线指纹库 Ψ 中提取的覆盖向量 $c_{ii} = 1$ 对应 AP 的指纹,携带了对应在线 RSS 的最多 AP 信息。 $\mathbf{F}'_s$ 作为子空间的随机 AP 部分,是从指纹库 Ψ 中 $c_{ii} = 0$ 所对应的 AP 中随机选取的 s 个。指纹 AP 的这种选择方式可有效避免突发噪声引起的在线 RSS 信号及覆盖向量的变化导致的参考点匹配错误。

根据 WKNN 算法,可通过式(9)计算 φ_t 与子空间内第 n 个指纹的欧式距离。

$$d_{on}(\varphi_t, \mathbf{F}'_n) = \frac{\sqrt{\sum_{j=1}^{m+s} ((f_{tj} - f'_{nj}) \cdot \theta_{nj})^2}}{w_{\theta n}}, \forall n \in \{1, 2, \dots, N'\} \quad (9)$$

$$w_{\theta_n} = \sum_{j=1}^{m+s} \theta_{nj}, \forall n \in \{1, 2, \dots, N'\} \quad (10)$$

式中: d_{on} 是活动 AP 的平均欧式距离; 定义 θ_{nj} 表示子空间内 AP 的活动性, 子空间内第 n 个指纹的第 j 个 AP 非 0, 则 $\theta_{nj} = 1$ 表示可参与指纹欧式距离的计算, 否则 $\theta_{nj} = 0$ 。这种改进型的欧氏距离可有效避免某个参考点指纹中固定 AP 部分的 RSS 信号距离较小而随机选取的 AP 其 RSS 信号距离较大使得总体欧式距离偏大, 而把该参考点从最近邻点剔除的情况, 有利于提高 WKNN 算法的定位精度。

第 l 子空间的估计位置是 K 个参考点位置的加权平均, 表示为:

$$(\hat{x}_l, \hat{y}_l) = \frac{1}{\sum_{j=1}^K \frac{1}{d_{oj}}} \sum_{j=1}^K \frac{(x_{lj}, y_{lj})}{d_{oj}}, \forall l \in \{1, 2, \dots, L\} \quad (11)$$

其中, (x_{lj}, y_{lj}) 是从子空间内选出的 RSS 信号平均欧式距离参考位置点的位置; d_{oj} 则是相应的平均欧式距离。

经过子空间匹配定位, 可以得到 L 个估计位置。由于子指纹库中某些 AP 是随机选择, 且目标定位时在线 RSS 采集时的传播空间与指纹库构建时的环境肯定存在差异, 使得估计位置存在一定的定位误差。本文利用基于密度峰值的聚类算法, 从这 L 个位置中选取可信度较高的几个来定位目标的最终位置。

2.2 密度峰值聚类算法选取可信估计位置

DPC 是一种基于密度的聚类算法, 通过计算样本的局部密度和距离, 自动高效完成核心点、边界点和噪声点的分类。相比较其他基于密度的聚类算法, 如 DBSCAN, DPC 算法同样可以处理任意形状的簇、不需要提前指定类的数目, 也有发现异常点的能力, 但它不只是依据样本点的密度进行聚类, 而是引入了一个综合样本密度和样本间距离的决策值参量对样本进行分类^[17]。

DPC 算法的核心思想是类簇中心点的局部密度总是比它的邻居点高; 且类簇中心点距离局部密度更高的点较远。对于任意样本点 i , 算法用两个参数来描述: 局部密度 ρ_i 、与密度更高的最近样本点的距离 δ_i 。算法提供了基于截断距离和核距离的两种方法来计算局部密度 ρ_i , 分别如式 (12)、(13) 所示:

$$\rho_i = \sum_{j \neq i} \chi(d_{ij} - d_c), \chi(x) = \begin{cases} 1, & x < 0 \\ 0, & x \geq 0 \end{cases} \quad (12)$$

$$\rho_i = \sum_{j \neq i} \exp \left[- \left(\frac{d_{ij}}{d_c} \right)^2 \right] \quad (13)$$

式中: d_{ij} 是样本点 i 与 j 之间的欧式距离; $d_c > 0$ 为截断距离, 也即样本点的邻居半径。研究表明, 对大样本数据,

d_c 值对聚类结果影响很小。算法的提出者用了所有样本间距离的前 2% 作为截断距离, 用户也可以自行定义。样本距离 δ_i 的定义如式 (14) 所示:

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \quad (14)$$

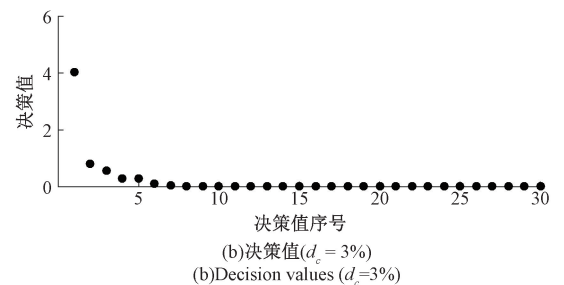
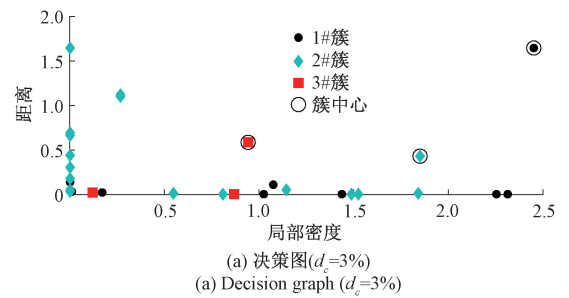
对于密度最大的样本点 i , 其距离 δ_i 则定义为:

$$\delta_i = \max_{i \neq j} (d_{ij}) \quad (15)$$

DPC 算法采用决策值 γ_i 来简化聚类中心的选择和确定。

$$\gamma_i = \rho_i \times \delta_i \quad (16)$$

对所有样本点的决策值排序, 可以观察到明显的跳跃点。跳跃值之前较大的决策值对应的样本点即为聚类中心, 同时也确定了聚类个数。后续则把其余的样本点分配到与它最近的密度更高的样本点所属的簇内则完成了聚类。图 3 给出了在某一测试点采用本文所提的指纹子空间匹配定位所得的 30 个估计位置的决策图。图 3(a)、(c) 中圆圈所标的为 3 个簇中心位置, 其他估计位置则用不同标记标识所属的不同类簇。图 3(b)、(d) 则给出了排序后的决策值 γ , 可以看到除了少数几个估计位置, 大部分的估计位置对应的决策值都很小。图 3(a)、(b) 和图 3(c)、(d) 由于选取的截断距离 d_c 不同, 决策图的分布和决策值的大小也有不同呈现, 但这对最终目标的估计位置并无大的影响。对于本测试点, 其实际位置为 (25.2, 0.6), d_c 取所有距离的 3% 时目标估计为 (25.24, 0.34), 而 d_c 取所有距离的 30% 时, 估计位置为 (25.15, 0.22); 采用 WKNN 算法时估计位置为 (23.39, 0.10); 定位误差分别是 0.26, 0.38 和 1.88 m。可见, DPC 算法的参数选择对定位结果并无大的影响, 这使得算法应用时较容易, 其有效性对参数的依赖性不强。



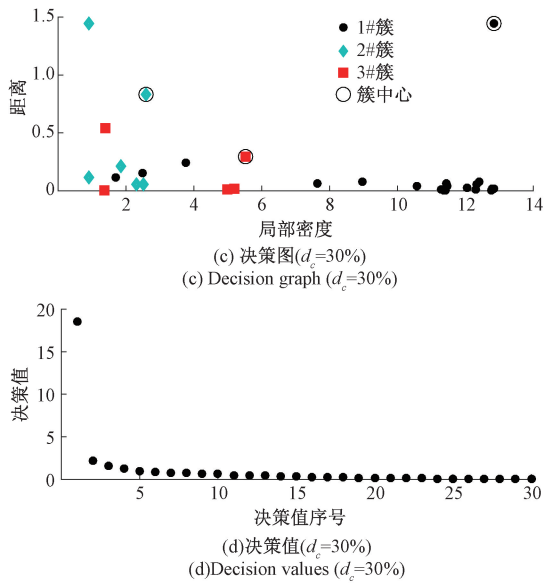


图3 DPC算法对估计位置的决策结果

Fig. 3 Decision results of DPC on estimated positions

不同于其他算法利用 DPC 算法剔除异常样本点或寻找聚类中心进行融合聚类,本文只利用决策值来选取可信度更高的估计位置,而不对估计位置进行聚类。经子空间匹配获得的 L 个估计位置,每个位置作为一个样本计算其决策值,选取最大的 S 个决策值对应的位置作为目标的最终估计位置,表示为:

$$(\hat{x}_p, \hat{y}_p) = \frac{1}{S} \sum_{j=1}^S (\hat{x}_j, \hat{y}_j) \quad (17)$$

3 实验结果及分析

3.1 实验环境建立

为验证本文所提的基于 DPC 指纹子空间匹配的室内定位方法的性能,在面积约为 $70 \text{ m} \times 14 \text{ m}$ 的室内办公环境进行了实际定位测验。实验环境具体布局如图 4,分别标记为地图 A 和地图 B。图 4 中圆点为参考位置点,三角表示在线测试点。实验中,利用移动智能终端搭载自主开发的微信小程序在多个工作日扫描采集了参考点和测试点的 WiFi 信号强度来构建离线指纹库及在线 RSS 信号。目标未能接收到信号的 AP 的 RSS 值设为 0。两个实验环境参数如表 1 所示。

表1 实验设置

Table 1 Setup of the experiments

场所	采集设备	采样间隔/m	指纹数量	AP 数量	测试点数量
A	Mate30	1.8×1.8	136	351	44
B	Honor20	1.8×1.8	92	199	42

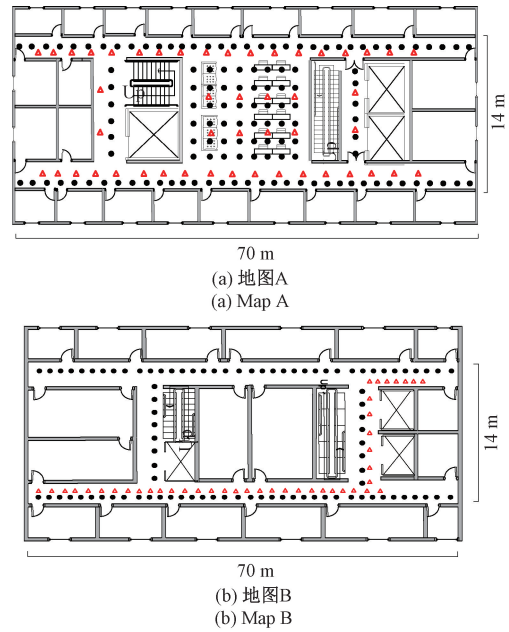


图4 实验布局图

Fig. 4 Maps of the experimental sites

3.2 基于 DPC 指纹子空间匹配定位方法性能分析

1) 定位方法对大误差的抑制

实验中首先分析了所提算法对大定位误差的抑制作用。实验分析中,本文以定位误差、平均定位误差 (average positioning error, APE) 和定位误差的累计分布函数 (cumulative distribution function, CDF) 作为衡量算法定位性能的指标。式(18)给出了 APE 的定义:

$$APE = \frac{1}{N_t} \sum_{i=1}^{N_t} E_i = \frac{1}{N_t} \sum_{i=1}^{N_t} \sqrt{(x_i - \hat{x}_{pi})^2 + (y_i - \hat{y}_{pi})^2} \quad (18)$$

式中: N_t 是测试点的数量; (x_i, y_i) 和 $(\hat{x}_{pi}, \hat{y}_{pi})$ 分别是第 i 个测试点的真实位置和算法的估计位置; E_i 则是相应的定位误差,也是两个位置间的欧式距离。

图 5 给出了 2 个地图中各测试点的定位误差。其中本文所提算法的参数设置如下:根据覆盖向量选取的子空间参考点数量 $N' = 20$,随机选择 AP 数 $s = 20$,子空间数 $L = 30$,子空间内改进 WKNN 匹配的最近邻点 $K = 3$,DPC 算法选取的决策值最大的估计位置数 $S = 5$;全指纹库匹配的 WKNN 算法最近邻点 $K = 3$ 。在地图 A 中,采用全局 WKNN 定位时,误差 4 m 以上的测试点有 6 个,最大误差是 5.42 m;采用本文所提的算法,所有测试点的定位误差都控制在 4 m 以下,最大的定位误差是 3.41 m。在地图 B 中也可观察到同样有效的抑制作用:全局 WKNN 的最大定位误差达到了 7.43 m,而采用本文算法,最大定位误差只有 3.97 m,所有测试点的误差也都

控制在 4 m 以下。从这 2 个实验场景的实际定位结果可看出,本算法通过有效的 AP 选择,改进的 WKNN 算法以及 DPC 算法的结合,能抑制在线定位时突发噪声引起的大定位误差。

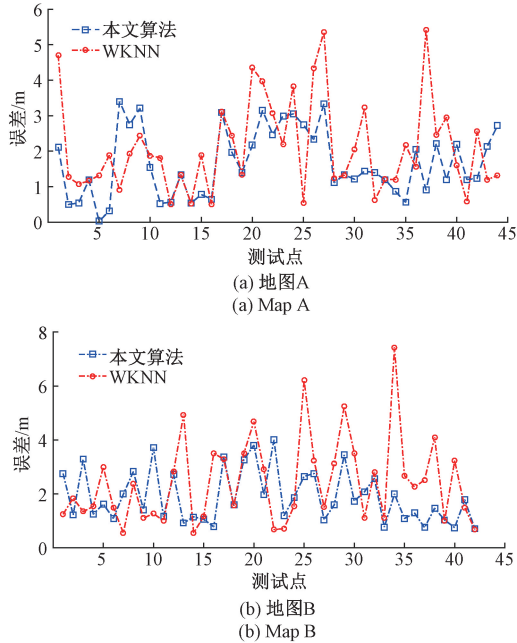


图 5 各测试点的定位误差

Fig. 5 Positioning error of the test points

实验中也分析了子空间内采用式(9)、(10)对 WKNN 算法改进后对定位性能的提升。图 6 的实验设置如下:子空间参考点数量 $N'=20$,位置指纹只包含固定 AP,不进行随机 AP 选择,也即每个测试点利用 20 个参考点和目标能接收的 AP 在线构建一个子空间。图 6 中 WKNN 表示在子空间内采用常规 WKNN 定位,A-WKNN 表示改进后的定位精度。可以看到在地图 B 中由于改进了目标 RSS 与指纹欧式距离的计算方式,定位性能有了很大的提升。而在地图 A 中性能提升不显著,但最大定位误差从 4.61 m 降到 4.08 m,平均定位误差从 1.99 m 降到了 1.81 m。可见在不增加计算复杂度的情况下,改进的 WKNN 有利于定位精度的提高。

2) 参数选择对性能的影响

本文所提算法在离线阶段不需要进行定位模型的训练,而是直接根据在线接收的 RSS 信号进行目标定位,但也需确定若干参数。实验中也分析了如何选取合适的参数保证定位性能。

(1) 子空间参考点数量 N'

子空间的参考点是依据在线 RSS 与指纹库中参考点的覆盖向量的距离选取的,它决定了子空间的大小。考虑本文采用的基于 WKNN 算法的指纹匹配,不涉及需要大量样本的模型训练,因此子空间参考点数无需太大,

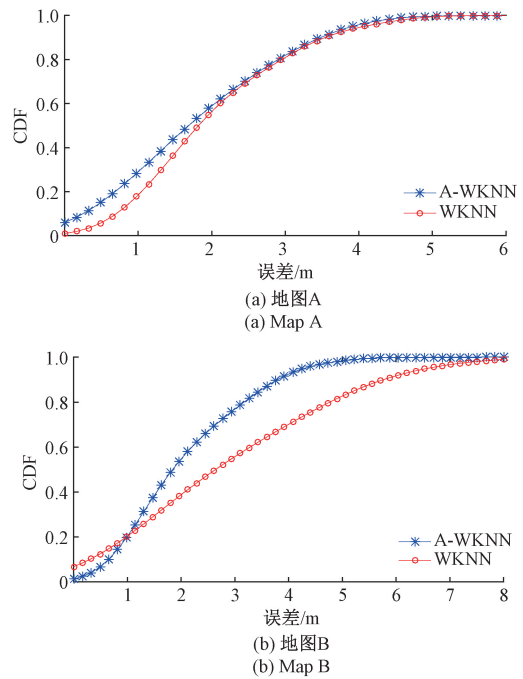


图 6 子空间内改进 WKNN 算法的定位性能

Fig. 6 Positioning performance of improved WKNN in subspace

表 2 的 1[#]、2[#]、3[#]给出了 N' 取不同值时的平均定位误差,选取 40 个参考点并没有比选取 20 个时有更好的定位精度。20~40 个足以包含目标真正所在位置附近的参考点, N' 太大,不仅会增大指纹匹配计算量,同时也会模糊目标在整个定位区域的局部位置。

(2) 随机选择 AP 数 s

在线目标不能接收信号的 AP 中随机选择 s 个,与在线 RSS 能接收的 AP 共同构成了子空间的位置指纹。表 2 的 1[#]、4[#]、5[#]实验结果表明, s 值在一定范围内的变化,对定位精度的影响不大。根据整个定位场所实际的 AP 数, s 可取 30~40 来平衡定位精度与计算复杂度。

(3) 子空间数 L

子空间数 L 决定了后续 DPC 算法的样本数。为保证 DPC 算法的有效性,原则上 L 不建议取太小。表 2 的 1[#]、6[#]、7[#]实验结果表明, L 与定位精度之间并没有确定的关系,10~30 都是可选的数目。不过值得一提的是, L 增大,使得子空间匹配和后续 DPC 算法定位的计算量都略增大。

(4) WKNN 的近邻数 K

在已有的各个文献中 WKNN 的近邻数 K 一般选取在 3~6。 K 值的最佳取值在不同应用场所略有不同。表 2 的 1[#]、8[#]实验结果表明,本定位方法对 K 的取值也与一般 WKNN 算法取值类似。考虑本实验场所有较长的狭窄走廊,尤其是地图 B,我们选取了 $K=3$ 。

表 2 不同参数下的平均定位误差

Table 2 Average positioning error with different parameters

实验序号	1 [#]	2 [#]	3 [#]	4 [#]	5 [#]	6 [#]	7 [#]	8 [#]	9 [#]
	$N' = 20$	$N' = 30$	$N' = 40$	$N' = 20$	$N' = 20$	$N' = 20$	$N' = 20$	$N' = 20$	$N' = 20$
	$s = 30$	$s = 30$	$s = 30$	$s = 20$	$s = 40$	$s = 30$	$s = 30$	$s = 30$	$s = 30$
地图	$L = 30$	$L = 30$	$L = 30$	$L = 30$	$L = 30$	$L = 10$	$L = 20$	$L = 30$	$L = 30$
	$K = 3$	$K = 3$	$K = 3$	$K = 3$	$K = 3$	$K = 3$	$K = 3$	$K = 4$	$K = 3$
	$S = 5$	$S = 5$	$S = 5$	$S = 5$	$S = 5$	$S = 5$	$S = 5$	$S = 5$	$S = 15$
A	1.71	1.70	1.72	1.74	1.69	1.74	1.75	1.78	1.74
B	2.10	2.11	2.18	2.11	2.03	1.90	2.00	2.12	2.09

(5) 最大决策值对应的估计位置数 S

S 对应的是从 L 个估计位置中选取可信度高的 S 个位置来完成最终目标的定位。表 2 的 1[#]、9[#] 实验结果表明该参数对定位精度影响不大。实际应用时根据 L 的大小, 选取合适值即可。

综合上述分析可知, 在实际应用中参数选择对本算法的性能影响不显著, 可根据实际应用环境选取合适值。本文的实验研究也提供了一些经验值供参考。

3.3 与已有定位方法的比较

为了进一步验证算法的定位性能, 图 7 给出了在两个实验场景中本文所提算法、DPC-Kmeans^[18] 和 PCA-KRR^[19] 算法的定位误差累计分布函数。其中 DPC-Kmeans 在离线阶段用 DPC 算法确定类簇中心点、再利用 Kmeans 算法对指纹聚类, 在线阶段通过簇匹配实现目标的粗定位, 再在簇内用 WKNN 实现定位, 是一种适合大区域的定位方式。本实验中地图 A 和地图 B 均形成 4 个聚类。而 PCA-KRR 是对指纹库中数量众多的 AP 信号提取主成分, 完成数据降维, 再利用核岭回归算法建立位置与 RSS 信号的模型。在本实验中, 主成分数分别选取了 11 和 22 个, 也即把包含 351 和 199 维的原始指纹降到了 11 和 22 维。由误差累计分布曲线可观察到, 本定位方法在两个场景中的最大定位误差均在 4 m 左右; 而另两种算法在地图 A 中的最大定位误差达到了 6 m, 在地图 B 中则超过了 6 m。CDF 的整体曲线表明本算法在抑制大误差和提高平均定位误差的性能上均优于其他两种方法。具体来讲, 上述三种算法在地图 A 中的平均定位误差分别是 1.69, 2.07 和 2.21 m; 在地图 B 中分别是 2.03, 2.48 和 2.34 m。而全局 WKNN 算法在这两个地图中的平均定位误差分别是 2.36 和 2.72 m。相比较传统 WKNN 算法, 所提算法的定位精度分别提升了 28% 和 25% 左右。

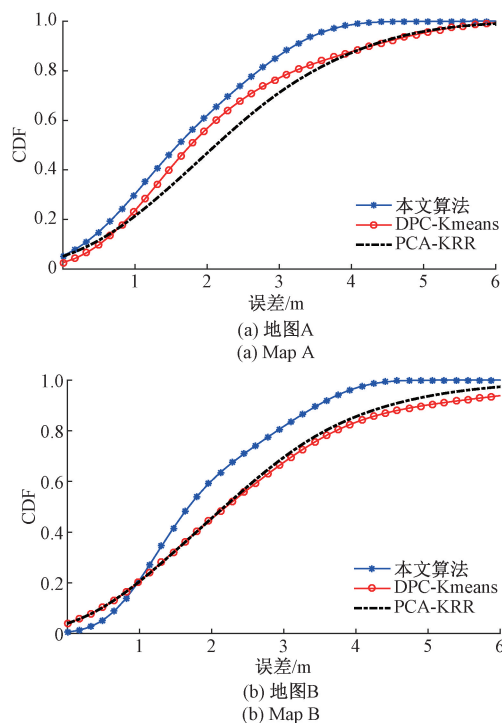


图 7 不同算法的定位性能比较

Fig. 7 Comparison of positioning performance of different methods

4 结 论

针对大区域室内定位中存在大量 AP, 而目标只能接收局部 AP 信号及 AP 信号存在突发干扰的实际情况, 提出了一种基于 DPC 的指纹子空间匹配的室内定位方法。该方法充分考虑 AP 信号在室内空间分布的局部性, 通过 RP 和 AP 的合理划分构建多个子空间, 利用一种改进的 WKNN 算法获取目标在每个子空间内的估计位置, 最

后利用 DPC 算法选取可信度最高的多个估计位置确定目标的最终位置,有效排除了突发干扰引起的大定位误差。实验研究和分析表明本定位方法实施简单,且能有效提高定位性能。

参考文献

- [1] HARROUD H, AHMED M, KARMOUCH A. Policy-driven personalized multimedia services for mobile users[J]. IEEE Transactions on Mobile Computing, 2003, 2(1): 16-24.
- [2] 王坚, 刘纪平, 韩厚增, 等. 应急救援无缝定位关键技术研究[J]. 武汉大学学报(信息科学版), 2020, 45(8): 1126-1136.
WANG J, LIU J P, HAN H Z, et al. Key technologies of seamless location in emergency rescue[J]. Geomatics and Information Science of Wuhan University, 2020, 45(8): 1126-1136.
- [3] YU N, ZHAN X H, ZHAO S N, et al. A precise dead reckoning algorithm based on Bluetooth and multiple sensors[J]. IEEE Internet of Things Journal, 2018, 5(1): 336-351.
- [4] 孙伟, 李亚丹, 黄恒, 等. 基于级联滤波的建筑信息/惯导室内定位方法[J]. 仪器仪表学报, 2021, 42(3): 10-16.
SUN W, LI Y D, HUANG H, et al. A location method of building structure information/inertial navigation combination based on the cascade filtering[J]. Chinese Journal of Scientific Instrument, 2021, 42(3): 10-16.
- [5] 米伟娟, 李娜. 基于变分自编码器的 WLAN 定位方法[J]. 电子测量与仪器学报, 2020, 34(12): 101-108.
MI W J, LI N. Variational auto encoder based WLAN positioning[J]. Journal of Electronic Measurement and Instrumentation, 2020, 34(12): 101-108.
- [6] WU Z, FU K, JEDARI E, et al. A fast and resource efficient method for indoor positioning using received signal strength[J]. IEEE Transactions on Vehicular Technology, 2016, 65(12): 9747-9758.
- [7] QIAN W, LARRI F, GECHTER F. Supervised and semi-supervised deep probabilistic models for indoor positioning problems[J]. Neurocomputing, 2021, 435(3): 228-238.
- [8] PARAMVIR B, VENKATA N P. RADAR: An in-building RF-based user location and tracking system[C]. Infocom Nineteenth Joint Conference of the IEEE Computer & Communications Societies, IEEE, 2000: 775-784.
- [9] FANG X M, JIANG Z H, NAN L, et al. Optimal weighted K-nearest neighbour algorithm for wireless sensor network fingerprint localisation in noisy environment[J]. IET Communications, 2018, 12(10): 1171-1177.
- [10] 王博远, 刘学林, 蔚保国, 等. WiFi 指纹定位中改进的加权 k 近邻算法[J]. 西安电子科技大学学报, 2019, 46(5): 41-47.
WANG B Y, LIU X L, WEI B G, et al. Improved weighted k-nearest neighbor algorithm for WiFi fingerprint positioning[J]. Journal of Xidian University, 2019, 46(5): 41-47.
- [11] XUE W X, YU K, HUA X H, et al. APs' virtual positions-based reference point clustering and physical distance-based weighting for indoor Wi-Fi positioning[J]. IEEE Internet of Things Journal, 2018, 5(4): 3031-3042.
- [12] ALI K, NIKOLAOS G, DANIEL J P, et al. A joint indoor WLAN localization and outlier detection scheme using LASSO and Elastic-Net optimization techniques[J]. IEEE Transactions on Mobile Computing, 2017, 16(8): 2079-2092.
- [13] LE Y F, ZHANG H N, SHI W B, et al. Received signal strength based indoor positioning algorithm using advanced clustering and kernel ridge regression[J]. Frontiers of Information Technology & Electronic Engineering, 2021, 22(6): 827-838.
- [14] CHEN C, WANG Y, ZHANG Y, et al. Indoor positioning algorithm based on nonlinear PLS integrated with RVM[J]. IEEE Sensors Journal, 2018, 18(2): 660-668.
- [15] ALEX R, ALESSANDRO L. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492-1496.
- [16] DING G, TAN Z, ZHANG J, et al. Regional propagation model based fingerprinting localization in indoor environments[C]. 2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications, IEEE, 2013: 291-295.
- [17] LIU R, WANG H, YU X. Shared-nearest-neighbor-based clustering by fast search and find of density peaks[J]. Information Sciences, 2018, 450: 200-226.

- [18] 周向前, 张峰, 刘叶楠, 等. 基于融合聚类的蓝牙室内定位系统算法优化[J]. 自动化与仪表, 2020, 35(1): 80-85.
ZHOU X Q, ZHANG F, LIU Y N, et al. Algorithm optimization of Bluetooth indoor positioning system based on fusion clustering[J]. Automation & Instrumentation, 2020, 35(1): 80-85.
- [19] LE Y, JIN S, ZHANG H, et al. Fingerprinting indoor positioning method based on kernel ridge regression with feature reduction[J]. Wireless Communications and Mobile Computing, 2021: 1-12.

作者简介



乐燕芬(通信作者), 2001 年于厦门大学获得学士学位, 2004 年于中国科学院大学获得硕士学位, 2016 年于上海理工大学获得博士学位, 现为上海理工大学副教授, 主要研究方向为室内定位及物联网相关技术。

E-mail: leyanfen@usst.edu.cn

Le Yanfen (Corresponding author) received her B. Sc. degree from Xiamen University in 2001, received her M. Sc. degree from University of Chinese Academy of Sciences in 2004, and received her Ph. D. degree from University of Shanghai for Science and

Technology in 2016. She is currently an associate professor at University of Shanghai for Science and Technology. Her main research interests include indoor localization and internet of things.



许远航, 2020 年于江苏大学京江学院获得学士学位, 现为上海理工大学研究生, 主要研究方向为基于无线网络的室内定位。

E-mail: xyh2298429306@163.com

Xu Yuanhang received his B. Sc. degree from Jiangsu University Jingjiang College in 2020. He is currently a graduate student at University of Shanghai for Science and Technology. His main research interest is indoor positioning based on wireless networks.



施伟斌, 2002 年于复旦大学获得博士学位, 现为上海理工大学副教授, 主要研究方向为无线传感器网络与物联网。

E-mail: shiweibin@usst.edu.cn

Shi Weibin received his Ph. D. degree from Fudan University in 2002. He is currently an associate professor at University of Shanghai for Science and Technology. His main research interests include wireless sensor network and internet of things.