

文章编号: 2095-4980(2014)01-0076-06

低指纹密度下分段式高精度室内定位方法

谢代军, 孔范增

(信息工程大学 信息工程学院, 河南 郑州 450002)

摘要: 针对室内环境中传统定位方法在大定位区域、低指纹密度下定位精确度低、计算复杂度高的问题, 提出了一种基于线性插值法和分布重叠的分段式定位方法。该方法采用传统的最近邻法进行粗定位, 得到可信区域; 利用线性插值法更新可信区域内指纹数据库, 增加指纹密度; 在可信区域内, 采用基于分布重叠的指纹相似度匹配法实现精定位。实验结果表明, 在低指纹密度下, 该定位方法定位精确度较高, 算法复杂度适中, 具有一定的适用性。

关键词: 室内定位; 低指纹密度; 线性插值; 分布重叠; 指纹相似度

中图分类号: TN92

文献标识码: A

doi: 10.11805/TKYDA201401.0076

Segmented high accuracy indoor positioning under low fingerprint density

XIE Dai-jun, KONG Fan-zeng

(Institute of Information System Engineering, Information Engineering University, Zhengzhou Henan 450002, China)

Abstract: Aiming at low positioning accuracy and high computational complexity of the traditional indoor positioning method under large area and low fingerprint density, a new segmented high accuracy positioning algorithm based on linear interpolation and distribution overlapping is proposed. The algorithm uses the traditional nearest neighbor method for coarse positioning, and then updates the fingerprint database in the trusted area with linear interpolation method to increase the fingerprint density. Finally, fine positioning is realized by fingerprint matching method that based on distribution overlapping. Experimental results show that, even under low fingerprint density, the proposed method can get good positioning performance with high accuracy and suitable complexity.

Key words: indoor positioning; low fingerprint density; linear interpolation; distribution overlapping; fingerprint similarity

随着无线网络的普及, 基于位置的服务(Location Based Service, LBS)受到越来越多的关注, 在紧急救助、医疗保健、交通运输、治安消防、个性化信息传递等领域显示出巨大的活力。国内外针对基于 Wi-Fi、传感器网络、超宽带等的室内定位技术开展了大量研究^[1-4]。Wi-Fi 定位技术具有成本低、易部署、不受视线约束等优点, 基于 Wi-Fi 的位置指纹法定位是当前室内定位的重点。位置指纹法不需要专门的测量工具, 系统成本较低, 定位效果较好, 通过特定角度获取定位区域内的相关特征数据, 利用特征数据与预先建立的位置指纹信息匹配, 从而实现定位。定位精确度与离线阶段采集指纹的密度具有很大关系, 实际室内环境中, 各区域的指纹密度并不均等, 有的定位参考点较多, 指纹密度大, 而有些区域可能只有房间粒度(如 5 m 间隔)的指纹密度。因此, 研究在非均匀指纹密度, 尤其是低指纹密度情况下的 Wi-Fi 室内定位具有重要意义。

人们对基于 Wi-Fi 的指纹定位开展了大量研究, 较早的有 Bahl 等人提出的信号空间最近邻法^[3-4](Nearest Neighbor in Signal Space, NNSS)和信号空间 K 最近邻法(K-NNSS)。上述算法简单, 但定位精确度相对较低。针对该问题, 人们研究了基于概率分布的指纹匹配方法^[5-7]。T Roos 等^[7]采用柱状图表示接收信号的概率分布, 该方法需要较多存储空间; J Ledlie 等^[8]在指纹中保存分布的参数, 减少了存储空间的要求。赵方等^[9]采用基于贝叶斯准则的多元高斯混合分布模型, 相比单一高斯模型, 定位性能明显提高。徐玉滨等^[10]利用核方法进行似然运算, 提高了参考点指纹区分度, 适应能力较强, 但样本量要求较高。

上述定位方法多采用全局搜索方法, 参考点指纹密度高时, 搜索开销大, 定位延迟长; 指纹密度低时, 定位精确度则较低。针对上述问题, 本文提出一种低指纹密度下分段式高精度定位方法, 该方法采用全局粗定位和

收稿日期: 2012-11-29; 修回日期: 2013-01-10

局部精定位相结合的模式进行定位,用传统的最邻近法在低密度指纹库中进行匹配,得到目标粗定位估计位置;根据粗定位估计位置,构建一个可信区域,在该区域内通过插值法生成新指纹,在可信区域内采用基于分布重叠的指纹匹配法得到目标的精估计位置。

1 定位算法

1.1 算法基本流程

分段式定位基本流程如图 1 所示。通过全局参考点指纹选择,构建一个低密度指纹库,大大减少指纹数据下载到本地网络的通信时间以及定位时全局搜索匹配时间,采用传统的最邻近法得到目标的估计位置,即粗定位。精定位则是在由粗定位估计位置决定的可信区域内,通过对可信区域内原有指纹进行线性插值,生成新的指纹,提高区域内指纹密度,再用基于分布重叠计算指纹相似度的方法,进行指纹匹配,得到目标最终估计位置。



Fig.1 Flow chart of the segmented positioning
图 1 分段式定位流程图

1.2 粗定位

1.2.1 构建低密度指纹数据库

典型室内环境如宾馆,其相邻 2 个房间的间距一般在 4 m 左右,因此,以 4 m 为间距设置参考点即可达到房间粒度的定位要求。用户在线定位时,按照一定距离(如 4 m)进行参考点指纹选择,由从全局范围内筛选出来的指纹构成指纹数据库,即为低密度指纹数据库。

1.2.2 向量空间最邻近法

在向量空间法中,使用该参考点所有信号强度样本向量的中心作为该参考点的代表向量。设指纹数据库中第 i 个参考点 C_i 的中心为 $\mu_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{im})$, m 为向量空间的维数,也是 Wi-Fi 接入点(Access Point, AP)的数目。

定位目标点 O 的向量为 $s = (s_1, s_2, \dots, s_m)$, 是在线阶段对 m 个 AP 信号强度观测值的均值。目标点与定位参考点的欧氏空间距离为:

$$L(O, C_i) = \sqrt{\sum_{j=1}^m |s_j - \mu_{ij}|^2} \quad (1)$$

根据式(1),逐个计算目标点与指纹数据库中各参考点的空间距离后,取与目标点信号空间距离最短的 K 个点(如取 $K=3$)的坐标均值为目标估计位置。即:

$$\hat{O}_{\text{粗}} = \frac{1}{3} \sum_{i=1}^3 O_i = \frac{1}{3} \sum_{i=1}^3 (x_i, y_i) \quad (2)$$

式中 O_1, O_2, O_3 分别为与目标点欧氏空间距离最短的 3 个参考点。

1.3 精定位

如图 1 所示,粗定位得到目标的大致位置后,以该位置为圆心,设定可信区域。更新可信区域内的指纹数据库后,采用基于分布重叠的指纹匹配方法进行精定位。

1.3.1 确定可信区域与更新指纹数据库

可信区域设定为以粗定位估计位置为圆心,以一定距离 R 为半径的圆。半径 R 的大小影响搜索时长和定位精确度。 R 过大,扩大了精定位搜索范围,增加了运算复杂度; R 过小,可能导致精定位结果误差偏大。根据粗定位参考点间距和实际环境中房间大小,设定 $R=5$ m 比较合适。在用户得到定位结果后,如认为估计不准确,可进行二次定位,增加 R 值,扩大搜索范围。精定位的所用参考点包括 2 种:一种是离线阶段设定的原参考点(其中一部分粗定位时已经使用);另一种是线性插值生成的新参考点,原参考点和新参考点的位置如图 2 所示。新参考点有两类:

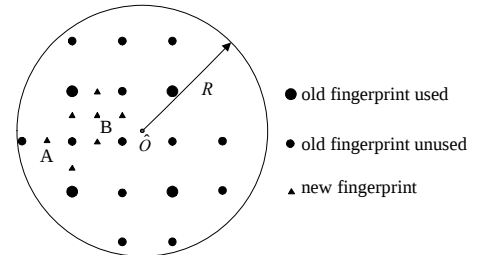


Fig.2 Location of reference points in the trusted zone
图 2 可信区域内参考点位置示意图

一类如图 2 中的 A 点,该类点的物理坐标、信号强度样本值分别为两相邻原参考点坐标、信号强度样本值的均

值；另一类如图 2 中的 B 点，该类点的坐标为其周围 4 个原参考点坐标的均值，其信号强度样本值为周围 4 个参考点样本的均值。由可信区域内原参考点指纹和新生成参考点的指纹构成了精定位的指纹数据库，完成定位数据库的更新。

1.3.2 基于分布重叠的指纹匹配法

基于分布重叠的指纹匹配法基于室内信号分布，服从对数高斯分布，位置指纹里存储各 AP 在该位置处信号高斯分布的均值和标准差，以各 AP 信号的概率分布来有效表示信号特征，用 2 个高斯分布的重叠度来计算分布的相似度，进而计算 2 个指纹的相似度，可有效反映 2 个位置在物理空间上的相近程度。根据以上匹配准则，将目标点与指纹库里的指纹逐个进行匹配，计算指纹相似度，取最大的 K 个位置并取其平均，得到目标点的位置估计。

假设可信区域内共有 l 个参考点 $\{L_1, L_2, \dots, L_l\}$ ，相应地，指纹数据库里有 l 个位置指纹 $\{F_1, F_2, \dots, F_l\}$ ，每个指纹与参考点有一对一的映射关系，第 i 个参考点的位置指纹为：

$$F_i = (s_1, s_2, \dots, s_m) \quad (3)$$

式中： m 为整个定位区域内部署的 AP 总数； s_j 表示第 j 个 AP 在第 i 个参考点处的信号特征。该信号特征为：

$$s_j = (\mu_j, \delta_j) \quad (4)$$

式中 μ_j, δ_j 分别表示第 j 个 AP 在第 i 个参考点处信号强度的均值和标准差。

位置 L_a 和 L_b 的指纹分别是 F_a, F_b ，每个指纹都有 m 个特征点，定义 2 个指纹的相似度为 2 个指纹在各个特征点上的相似度之和，即：

$$S(F_a, F_b) = \sum_{i=1}^m \delta_i = \sum_{i=1}^m O(\mu_{ia}, \delta_{ia}, \mu_{ib}, \delta_{ib}) \quad (5)$$

式中： δ_i 表示单个接入点 AP 的相似度贡献值； $O(\cdot)$ 是第 i 个接入点 AP 在位置 L_a 和 L_b 处信号强度分布的重叠度计算函数(简称重叠度函数)，如图 3 所示。

位置指纹相似度反映了 2 个位置处无线电信号分布的相似程度，以及空间上的接近度。在大尺度范围内(10 m)，两位置越近，指纹相似度越大；小尺度范围内(1 m)，相邻的两参考点，其与目标的指纹相似度区别性不大。因此，本文取与目标点指纹相似度最大的 K 个参考点构成候选点集 H_K ，排除候选点集中离散度最大的参考点后，取候选点集的几何中心作为目标点的最终估计位置。

候选点集 $H_K = \{X_1, X_2, \dots, X_K\}$ ，其中 $X_i = (x_i, y_i)$ 是候选点集中第 i 个参考点的坐标，则候选点集的几何中心为：

$$X_{\text{centre}} = \frac{1}{K} \left(\sum_{i=1}^K x_i, \sum_{i=1}^K y_i \right) = (\bar{x}, \bar{y}) \quad (6)$$

候选点集中各点到几何中心的距离为：

$$d(X_{\text{centre}}, X_i) = \sqrt{(\bar{x} - x_i)^2 + (\bar{y} - y_i)^2} \quad (7)$$

候选点集各点到几何中心的平均距离为：

$$L_{\text{average}} = \frac{1}{K} \sum_{i=1}^K L(X_{\text{centre}}, X_i) \quad (8)$$

最大距离为：

$$L_{\text{max}} = \arg \max_{1 \leq i \leq K} L(X_{\text{average}}, X_i) \quad (9)$$

定义点集离散度为最大距离与平均距离的比值：

$$D = \frac{L_{\text{max}}}{L_{\text{average}}} \quad (10)$$

根据式(10)，计算候选点集 H_K 的离散度，如 $D > 1.5$ 超出规定值，则将 L_{max} 对应参考点从候选点集中排除。根据式(6)，重新计算候选点集的几何中心。取候选点集的几何中心为目标点精定位的估计位置：

$$\hat{O}_{\text{精}} = X_{\text{centre}} = (\bar{x}, \bar{y}) \quad (11)$$

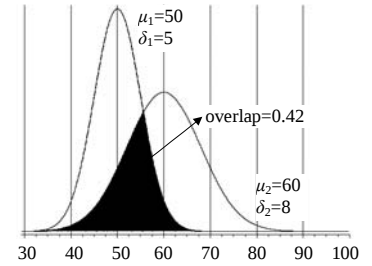


Fig.3 Overlap of Gauss distribution
图 3 高斯分布重叠

1.4 低指纹密度下分段式定位的步骤

综上所述,在离线阶段得到指纹数据库之后,按照以下流程确定目标位置:

Step 1: 对各个 AP 进行若干次扫描,得到定位目标点的指纹;

Step 2: 在本地构建或者从定位服务器下载低密度指纹库;

Step 3: 利用式(1)计算参考点指纹与目标点指纹的距离,利用式(2)得到粗定位估计位置 $\hat{O}_{\text{粗}}$;

Step 4: 根据粗定位估计位置及设定的半径值 R ,确定可信区域范围;

Step 5: 利用可信区域内原有参考点,线性插值生成新的参考点及新指纹,用原有参考点和新生成参考点的指纹构建精定位指纹数据库;

Step 6: 根据式(5)计算目标点与可信区域内各参考点的指纹相似度;

Step 7: 取与目标点指纹相似度最大的 K 个参考点构成候选点集,利用式(6)~式(10),计算候选点集的离散度,并将使候选点集离散度值超出规定值的参考点排除;

Step 8: 由式(11),确定目标点的精定位的估计位置 $\hat{O}_{\text{精}}$ 。

2 定位实验

2.1 实验环境设置

实验环境如图 4 所示,共设置 89 个定位参考点,部署 10 个 AP。相邻定位参考点间距为 2 m,AP 离地面 78 cm。采用华硕 F8H8SN 笔记本的 Intel(R) Wireless Wi-Fi Link 4965AGN 网卡对 AP 进行扫描,采集 AP 信号强度。每个参考点以 5 s 间隔扫描 20 次。为充分验证算法的性能,在图 2 中的走廊和房间内共设定 50 个定位测试点,逐个进行定位测试,检测算法定位性能。低密度指纹库,则由图中左上角的第一个参考点开始,按 $d=4$ m 的间距进行抽取,共构成 25 个参考点指纹。

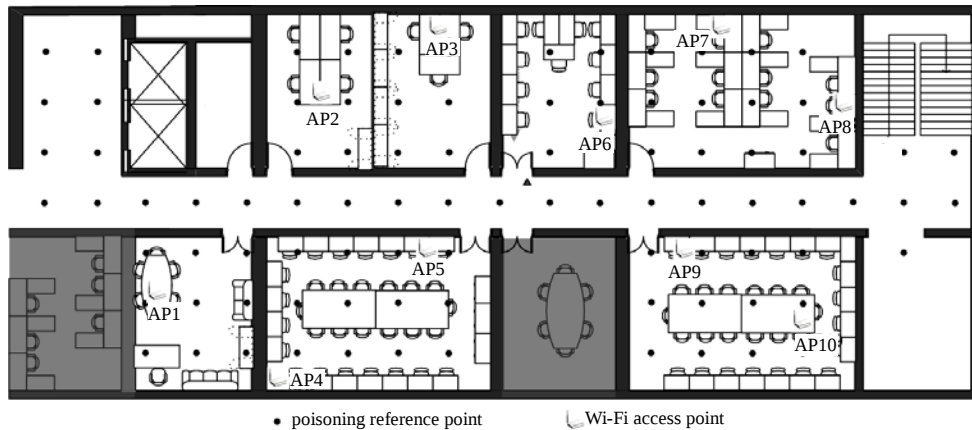


Fig.4 Experiment environment

图 4 实验环境

2.2 实验结果及分析

利用数据对算法进行测试,性能分析如下。

2.2.1 粗估计与精估计性能比较

图 5 为定位误差累积概率分布曲线图,从图中可以看出,参考点间距为 $d=2$ m, 1 m 时,定位效果均比粗估计的定位效果好,这与预期相符。 $d=2$ m 表示精定位时,可信区域内的参考点间距为 2 m,且所有参考点指纹为离线阶段采集的; $d=1$ m 表示精定位时,可信区域内的参考点一部分是通过插值法生成的,参考点间距为 1 m。从图 5 可以看出,插值后的平均定位误差为 1.45 m,优于未插值时的 1.75 m,更远远优于粗定位估计时的 2.3 m。可见,本文所提算法在定位精确度上有较大的提高。

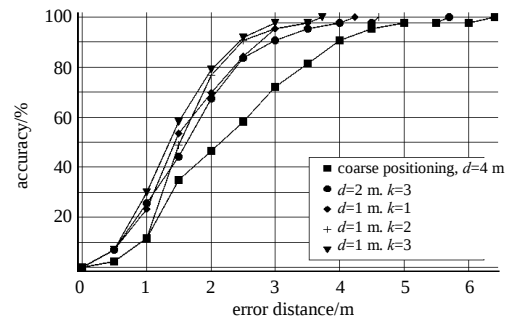


Fig.5 Cumulative probability distribution of positioning error

图 5 定位误差累积概率分布

2.2.2 可信区域半径 R 的选择

可信区域半径决定了精定位阶段算法需要匹配的参考点指纹数, 对算法的定位精确度和复杂度有较大影响。从图 6 可看出, 随着半径 R 的增大, 精定位平均定位误差逐渐减少, 当 R 大于 8 m 后, 平均定位误差变化很小, 而算法在精定位阶段需要匹配的参考点随 R 的增大一直在增大。分析 2 条曲线可看出, 当 $R=5$ m 时, 算法平均定位误差为 1.45 m, 稍大于 $R=7.5$ m 时的定位误差, 但其需要匹配的参考点数为 33 个, 只有 $R=7.5$ m 时的一半。综合平均定位精确度和算法复杂度两方面分析, 可信区域半径 R 为 5 m 时, 算法综合性能最好。

2.2.3 K 的选择

当 K 取不同值时, 算法在粗定位与未插值精定位、插值后精定位 3 种情况下的平均定位误差如图 7 所示。从图 7 可看出, 插值后精定位情况下, $K=3$ 时, 平均定位误差最小。另, 从图 5 可看出, $K=3$ 时, 算法在各定位精确度下的累积概率均大于 $K=1, K=2$ 时的累积概率, 且 $K=3$ 时, 算法的最大定位误差也小于 $K=1, K=2$ 时的最大定位误差。因此, 综合算法复杂度、平均定位误差、累积概率及最大定位误差等指标, $K=3$ 时, 算法综合性能最好。

2.2.4 算法平均定位时间

可信区域半径 R 设定为 5 m, K 值设定为 3, 对本文所提分段式定位方法进行测试。进行 25 批, 每批对 50 个不同地点进行定位测试。对 25 批, 共 1 250 次定位时间进行统计, 所提算法的平均定位时间约为 16 ms。可见所提算法定位时间短, 具有较高的定位实时性。

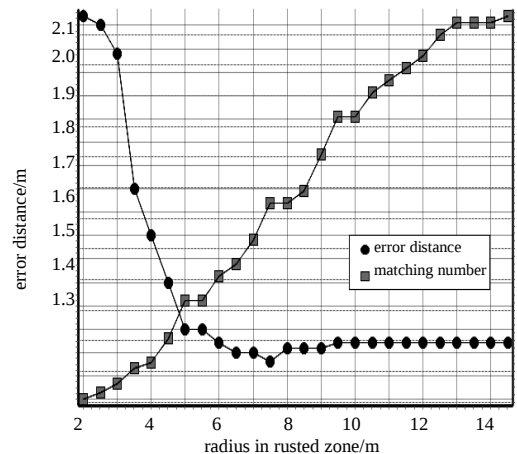


Fig.6 Radius selection in trusted zone

图 6 可信区域半径 R 的选择

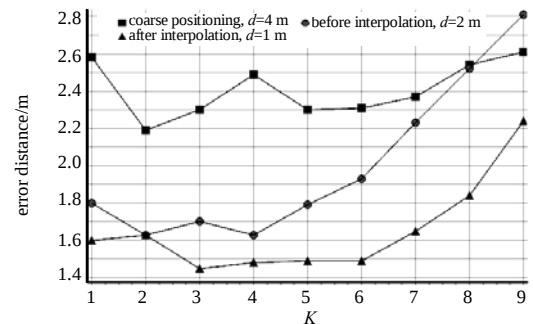


Fig.7 Average localization error versus K

图 7 平均定位误差与 K 的关系

3 结论

低指纹密度下, 现有算法定位效果较差, 本文提出了一种采用全局低指纹密度下粗定位与局部高指纹密度下精定位相结合的分段式定位的新方法。粗定位采用运算简单的最邻近法, 精定位则在基于粗定位估计结果设定的可信区域内, 利用插值法生成新参考点的办法, 增加局部指纹密度, 进而采用基于分布重叠的匹配方法估计目标位置。

本文比较了粗定位、插值前及插值后算法的定位性能, 分析了定位参数 R 和 K 的不同选择对算法定位性能的影响。测试结果表明, 插值后采用基于分布重叠的匹配法定位效果好, 可信区域半径 $R=5$ m, $K=3$ 时, 算法在定位精度、算法复杂度等方面综合性能最好。

参考文献:

- [1] 宋婉甜, 李智. 一种新的无线传感网络三维定位方法[J]. 太赫兹科学与电子信息学报, 2012, 10(3): 257-261. (SONG Wantian, LI Zhi. A new three-dimensional localization method for WSN[J]. Journal of Terahertz Science and Electronic Information Technology, 2012, 10(3): 257-261.)
- [2] 杨萍, 李智. 一种基于 RSSI 的权重值选择及加权定位算法[J]. 太赫兹科学与电子信息学报, 2012, 10(2): 148-151. (YANG Ping, LI Zhi. Weight-select and location algorithm based on RSSI[J]. Journal of Terahertz Science and Electronic Information Technology, 2012, 10(2): 148-151.)
- [3] Bahl P, Padmanabhan V N. RADA: an in-building RF-based user location and tracking system[C]// Proceeding of the IEEE INFOCOM 2000. Tel-Aviv, Israel: [s.n.], 2000: 775-784.
- [4] Bahl P, Padmanabhan V N, Balachandran A. Enhancements to the RADAR user location and tracking system[EB/OL]. (2002-02). <http://research.microsoft.com/~bahl/papers/pdf/msr-tr-2000-12.pdf>.