

基于 DBSCAN 子空间匹配的蜂窝网室内指纹定位算法

田增山 王向勇* 周牧 李玲霞
(重庆邮电大学移动通信重点实验室 重庆 400065)

摘要: 针对无线信道动态衰落特性引起的蜂窝网室内定位误差较大的问题, 该文提出基于密度的空间聚类 (Density Based Spatial Clustering of Applications with Noise, DBSCAN) 子空间匹配算法, 有效剔除大误差点, 提高定位精度。首先通过划分信号空间, 构建多个子空间, 在子空间中利用加权 K 近邻匹配算法 (Weighted K Nearest Neighbor, WKNN) 估计出目标位置; 然后利用 DBSCAN 对估计位置进行聚类以剔除异常点; 最后结合概率模型确定最终估计位置。实验结果表明, 基于 DBSCAN 的子空间匹配算法能有效剔除大误差点, 提高蜂窝网室内定位系统的整体性能。

关键词: 室内定位; 蜂窝网; DBSCAN; 子空间匹配

中图分类号: TN929.53

文献标识码: A

文章编号: 1009-5896(2017)05-1157-07

DOI: 10.11999/JEIT160768

DBSCAN Based Subspace Matching for Indoor Cellular Network Fingerprint Positioning Algorithm

TIAN Zengshan WANG Xiangyong ZHOU Mu LI Lingxia
(Chongqing Key Laboratory of Mobile Communication Technology, Chongqing University
of Posts and Telecommunications, Chongqing 400065, China)

Abstract: For the sake of reducing the indoor localization errors caused by dynamic signal fading in cellular network, this paper propose a novel Density-Based Spatial Clustering of Applications with Noise (DBSCAN) based subspace matching algorithm for indoor localization, which can improve the localization accuracy by eliminating the location with large errors. Specifically, the signal space is firstly divided into several subspaces, where a position estimation can be obtained respectively using the Weighted K Nearest Neighbors (WKNN) approach. Then, DBSCAN is applied to the position coordinates obtained from each subspace which eliminates the outliers. Finally, the location is estimated based on probability analysis. Experimental results show that the proposed approach can improve the location accuracy by eliminating the location with large errors.

Key words: Indoor localization; Cellular networks; Density Based Spatial Clustering Applications with Noise (DBSCAN); Subspace matching

1 引言

随着移动互联网、无线通信技术的不断发展, 人们对于位置服务的需求越来越高, 位置服务的基

础——定位技术成为国内外学者研究的热点。室内定位技术是公共安全、生产安全、应急救援、物联网、特殊人群监护、大型场馆管理、智慧城市建设等领域重点关注的内容^[1]。目前, 比较典型的室内定位技术主要有基于 RFID^[2], 可见光^[3], 蓝牙^[4], 红外^[5], WLAN^[6], MEMS^[7]定位等。它们在特定的条件下都有较高的定位精度, 但是也有各自的缺陷, RFID 受限于功率只能用于近距离定位; 可见光定位设备寿命短, 功耗高, 光源被遮挡后系统完全无法定位; 蓝牙定位和红外定位需要高密集布网, 且设备昂贵, 不适宜大规模商用; 基于 WiFi 的定位技术容易受接入点分布密度及外部信号的干扰等因素的影响; 基于 MEMS 的定位技术的定位误差随着时间累积, 导致长时间定位精度较差。此外, 还有基于扩频测距的定位方式^[8]和基于传感器网络的被动定位技

收稿日期: 2016-07-22; 改回日期: 2016-12-27; 网络出版: 2017-02-28

*通信作者: 王向勇 wangxiangyong1991@163.com

基金项目: 国家自然科学基金(61301126), 长江学者和创新团队发展计划(IRT1299), 重庆市基础与前沿研究计划(cstc2013jcyjA40041, cstc2015jcyjBX0065), 重庆邮电大学青年科学基金项目(A2013-31)。

Foundation Items: The National Natural Science Foundation of China (61301126), The Program for Changjiang Scholars and Innovative Research Team in University (IRT1299), The Fundamental and Frontier Research Project of Chongqing (cstc2013jcyjA40041, cstc2015jcyjBX0065), The Young Science Research Program of Chongqing University of Posts and Telecommunications (A2013-31)

术^[9], 它们虽然有较高的定位精度, 但是也需要在定位区域内布置复杂的 CCS 设备或者大量传感器节点。相比于这些定位系统, 蜂窝网定位技术结构简单, 能充分利用现有网络设施, 不需要增加额外硬件设备, 只要能接收到蜂窝网信号的地方, 都可以构建蜂窝网定位系统。

传统的蜂窝网定位算法, 主要有到达时间(Time Of Arrival, TOA)定位技术^[10]、到达时间差(Time Difference Of Arrival, TDOA)定位技术^[11]、到达角(Angle Of Arrival, AOA)定位技术^[12]。这些定位技术需要复杂的终端设备, TOA/TDOA 需要高精度的时钟同步系统, AOA 需要阵列天线。除此之外, 室内复杂的多径环境, 容易使到达时间与到达角判断不准确, 导致定位发生较大偏差。位置指纹是一种基于模式识别思想的定位技术, 它只需要一个可以接收蜂窝网信号的终端来采集信号, 通过匹配在线接收数据与离线指纹库来进行定位, 这种方式能克服多径的干扰。

关于蜂窝网指纹定位技术, 国内外的研究重点集中在室外^[13-15]而缺乏对室内指纹定位技术的研究, 主要原因是室内终端接收的信号易受无线传播环境的影响, 使信号在进行位置指纹匹配的时候发生偏差, 影响定位精度。基于此, 本文首先验证了蜂窝网室内指纹定位的可行性, 在此基础上本文结合基于密度的空间聚类^[16](Density-Based Spatial Clustering of Applications with Noise, DBSCAN)和子空间^[17]两种概念, 提出一种新的匹配定位算法。该算法充分利用丰富的蜂窝网小区信息, 提取多个信号子空间, 在子空间中进行加权 K 近邻(Weighted K Nearest Neighborhood, WKNN)匹配运算, 得到多个估计位置, 然后通过对所有估计位置进行 DBSCAN 聚类及概率分析, 最终确定待定位点坐标。文章最后利用室内环境下的蜂窝网实测数据验证了算法的有效性。

2 蜂窝网室内信号分析

位置指纹定位技术, 主要是通过匹配在线接收数据与离线指纹库, 计算实时接收信号与各指纹点的相似度来确定当前位置。系统的定位精度由以下两点决定: (1)信号的稳定性: 信号越稳定, 在同一个点不同时刻接收的信号越相似, 匹配越准确, 定位精度越高; (2)信号在不同指纹点之间的差异性: 不同指纹点之间的信号差异性越明显, 在线匹配出错的概率越低, 定位精度越高。

信号从基站到室内接收端, 服从无线信号传播模型^[18], 即

$$\text{rss}_t = P(d_0) + 10n \lg(d/d_0) + P_{\text{wf}} + \xi_t + \chi \quad (1)$$

其中, rss_t 表示 t 时刻接收信号强度(Received

Signal Strength Indication, RSSI); $P(d_0)$ 表示参考位置的接收信号强度, 它是一个常数; $10n \lg(d/d_0)$ 表示信号从参考位置到接收机的传播损耗, 其中 n 为与环境有关的损耗因子, 在同一环境下为常量; P_{wf} 表示信号在室内的穿墙损耗; ξ_t 表示 t 时刻的其它干扰, 可认为是突发噪声; χ 为环境噪声, 它服从零均值高斯分布, 即: $\chi \sim N(0, \sigma_n^2)$, 这里 σ_n^2 表示噪声功率。

对于同一个点两个不同的时间 t_1, t_2 接收的信号, 信号经历的路径是一致的, 所以其穿墙的损耗 P_{wf} 和传播损耗应该是相同的。因此其接收信号强度的差值可表示为

$$\begin{aligned} \nabla \text{rss}_i &= \text{rss}_{i_{t_1}} - \text{rss}_{i_{t_2}} \\ &= (P(d_0) + 10n \lg(d/d_0) + P_{\text{wf}} + \xi_{t_1} + \chi)_{t_1} \\ &\quad - (P(d_0) + 10n \lg(d/d_0) + P_{\text{wf}} + \xi_{t_2} + \chi)_{t_2} \\ &= \xi_{t_1} - \xi_{t_2} + \chi_{t_1} - \chi_{t_2} \end{aligned} \quad (2)$$

其中, ξ_{t_1}, ξ_{t_2} 表示 t_1, t_2 两个时刻的突发噪声, χ_{t_1}, χ_{t_2} 表示 t_1, t_2 两个时刻的环境噪声, 由于 χ_{t_1}, χ_{t_2} 为不相关的零均值高斯噪声, 所以 $(\chi_{t_1} - \chi_{t_2}) \sim N(0, \sqrt{2}\sigma_n^2)$, 即 $(\chi_{t_1} - \chi_{t_2}) = \sqrt{2}\chi$ 。式(2)可简化为

$$\nabla \text{rss}_i = \xi_{t_1} - \xi_{t_2} + \sqrt{2}\chi \quad (3)$$

同理, 对于不同的点 i, j , 其接收信号强度的差值可表示为

$$\begin{aligned} \nabla \text{rss}_{ij} &= \text{rss}_i - \text{rss}_j \\ &= \xi_{t_i} - \xi_{t_j} + 10n \lg(d_i/d_j) \\ &\quad - (P_{\text{wf}}^{(i)} - P_{\text{wf}}^{(j)}) + \sqrt{2}\chi \end{aligned} \quad (4)$$

其中, ξ_{t_i} 和 ξ_{t_j} 表示在采集第 i 个点和第 j 个点的接收信号强度时的突发噪声, d_i 和 d_j 表示基站到第 i 个点和第 j 个点的距离, $P_{\text{wf}}^{(i)}$ 和 $P_{\text{wf}}^{(j)}$ 表示信号到第 i 个点和第 j 个点的穿墙损耗。

当不考虑突发噪声 ξ_t 时, 相同的点仅受高斯噪声 χ 影响, 若一个点的采集时间足够长, 可以通过均值滤波抑制该误差; 不同的点, 因为信号从室外到这些点的路径不一样, 导致穿墙损耗 P_{wf} 也不一样, 再加上小区到接收端的距离 d 的影响, 使其信号存在差异性, 且相隔越远, 信号的差异性越大。这使得蜂窝网信号满足指纹定位的条件。

但是, 无线信道动态衰落特性和室内环境的复杂性决定突发噪声 ξ_t 不能被忽略。信号在室内的多径传播可能引起终端接收到的某些小区(频点号+基站识别码)信号强度值的波动, 特别是有人在靠近接收端的地方进行干扰时, 某些小区的 RSSI 会发生剧烈变化, 图 1 为实际测试中有人在接收端附近走动时, 几个典型小区的 RSSI 的变化情况。

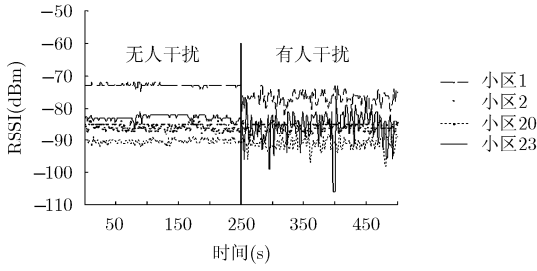


图1 人在接收端附近走动对 RSSI 的影响

图1中,前250 s是无人干扰时几个小区的RSSI,250 s以后是人靠近接收端走动时的几个小区的RSSI。图中显示,人的干扰会造成某些小区(如小区1, 小区23)的RSSI剧烈波动。这些突发噪声因为其出现的随机性,难以被消除。由式(3)和式(4)可以知道,当 ξ_t 变化剧烈时, ∇rssi 会变大, ∇rssi_{ij} 会变模糊,即信号的稳定性和指纹的差异性都会被破坏,导致在线接收数据匹配离线指纹库出错,影响系统的定位精度。基于此,本文提出基于 DBSCAN 的子空间匹配算法,旨在消除突发噪声的影响、提高系统的整体定位精度。

3 基于 DBSCAN 的子空间匹配算法

本文所提基于 DBSCAN 的子空间匹配算法流程如图2所示,它主要分为3个模块,分别为:子空间匹配定位模块,DBSCAN 聚类模块,概率模型模块。其中子空间匹配定位模块,主要是为了分离出没有出现突发噪声的小区信号,用其产生定位结果;DBSCAN 聚类模块是为了将各个子空间定位的结果形成一个个簇,其中心点作为备选位置,同时剔除异常点;概率模型模块是为了从备选位置中挑选出最有可能定位准确的点作为最终输出位置。下面将分别介绍这3个模块。

3.1 子空间匹配定位

通过离线时信号的采集,可以建立离线指纹库,其形式如式(5)所示:

$$\mathbf{FP} = \begin{pmatrix} x_1 & y_1 & \text{rssi}_{11} & \text{rssi}_{12} & \cdots & \text{rssi}_{1m} \\ x_2 & y_2 & \text{rssi}_{21} & \text{rssi}_{22} & \cdots & \text{rssi}_{2m} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ x_N & y_N & \text{rssi}_{N1} & \text{rssi}_{N2} & \cdots & \text{rssi}_{Nm} \end{pmatrix} \quad (5)$$

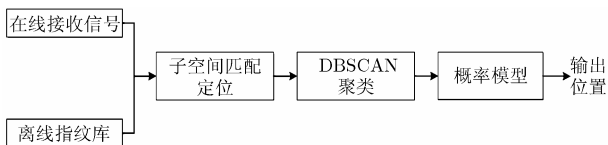


图2 基于 DBSCAN 的子空间匹配算法流程图

其中, (x_i, y_i) 表示第*i*个指纹点的坐标, rssi_{ij} 表示第*i*个指纹点处接收到的第*j*个小区的RSSI。一共有*N*个指纹点,采集到*m*个小区的信号。

为了分离出未受突发噪声影响的小区进行定位,本文构建原信号空间的子空间,在子空间中进行定位。设 v_j 表示第*j*个小区的RSSI,则所有*m*个小区的RSSI构成信号空间 $\mathbf{V}=[v_1, v_2, \dots, v_m]$ 。从这*m*个小区中任选*a*个小区的RSSI构成空间 $\mathbf{W}_e$, $\mathbf{W}_e$ 是 $\mathbf{V}$ 的子空间。设信号空间 $\mathbf{V}$ 的维度为*a*的子空间有*M*个,表示为

$$\left. \begin{aligned} \mathbf{W}_e &\subset \mathbf{V}, \quad e=1,2,\dots,M \\ \dim \mathbf{W}_e &= a \end{aligned} \right\} \quad (6)$$

$$M = C_m^a \quad (7)$$

其中, $\dim$ 表示维度, C_m^a 表示从*m*个数中取*a*个数的取法。

为方便计算,定义 h_{ej} ,用以标识第*j*个小区的RSSI是否在子空间 $\mathbf{W}_e$ 中:

$$\left. \begin{aligned} \mathbf{H}_e &= [h_{e1}, h_{e2}, \dots, h_{em}], \quad e=1,2,\dots,M \\ h_{ej} &= \begin{cases} 1, & v_j \in \mathbf{W}_e \\ 0, & v_j \notin \mathbf{W}_e \end{cases}, \quad j=1,2,\dots,m \\ \sum_{j=1}^m h_{ej} &= a \end{aligned} \right\} \quad (8)$$

假设当前接收信号为 $\mathbf{S}=[s_1, s_2, \dots, s_m]$,通过式(9)可计算 $\mathbf{S}$ 与各个指纹点在*a*维子空间 $\mathbf{W}_e$ 中的信号距离:

$$d_{ei} = \sqrt{\sum_{j=1}^m ((s_j - \text{rssi}_{ij}) h_{ej})^2}, \quad i=1,2,\dots,N \quad (9)$$

根据WKNN算法,对 d_{ei} 从小到大排序,得到:

$$d_{ei_1} \leq d_{ei_2} \leq \dots \leq d_{ei_N} \quad (10)$$

这里, $i_1, i_2, \dots, i_N$ 是 $1, 2, \dots, N$ 的一种重排。

取与 $\mathbf{S}$ 在子空间 $\mathbf{W}_e$ 中信号距离最近的*K*个点来估计当前位置坐标 $(\hat{x}_e, \hat{y}_e)$:

$$(\hat{x}_e, \hat{y}_e) = \frac{1}{\sum_{j=1}^K \frac{1}{d_{ei_j} + \varepsilon}} \sum_{j=1}^K \frac{1}{d_{ei_j} + \varepsilon} \times (x_{i_j}, y_{i_j}) \quad (11)$$

其中, ε 是为了防止 d_{ei_j} 等于零而设置的正的极小量, (x_{i_j}, y_{i_j}) 表示第*i_j*个指纹点的坐标。

子空间匹配定位的算法流程图如图3所示。

通过子空间匹配算法,可以得到*M*个定位点,系统需要利用这*M*个点,输出一个位置,使定位误差最小。

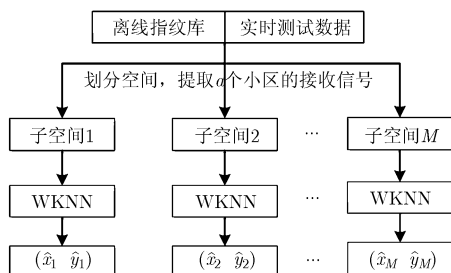


图 3 子空间匹配算法流程图

3.2 DBSCAN 聚类

假设第 l 个小区 RSSI 存在突发噪声, 在忽略高斯噪声的情况下, 根据式(9)-式(11), 若 $h_{el} = 0$, 则其不会影响子空间 $\mathbf{W}_e$ 的定位结果 $(\hat{x}_e, \hat{y}_e)$; 若 $h_{el} = 1$, 子空间 $\mathbf{W}_e$ 的定位结果 $(\hat{x}_e, \hat{y}_e)$ 会偏向于第 l 个小区信号跳变的方向, 通过聚类, 可以将没有发生偏离的定位点与发生偏离的点区分开。同理, 当其它小区信号存在突发噪声时, 也可以通过聚类区分开。

通过上面提到的子空间匹配定位, 一共可以得到 $M = C_m^a$ 个位置坐标, 通过上面的描述可知, 当子空间中包含有相同的突发噪声时, 其位置坐标更加接近。为了能从这 M 个坐标中找出正确的定位点, 本文对这些位置坐标进行聚类。

DBSCAN 是一种基于密度的聚类方法, 在本文算法中选择它主要是基于以下 3 点:

(1) 蜂窝网定位过程中, 会因为噪声使定位出的结果出现不与其它任何定位点邻近的异常点, DBSCAN 本身就具有发现异常点的能力, 它不会将异常点强行指派给某一类, 这样系统就可以在聚类的时候就将这些异常点剔除;

(2) DBSCAN 不需要提前指定类的数目, 它通过划分数据空间中稀疏区域分开的稠密区域来确定类。这点对本文所提蜂窝网定位的算法至关重要, 因为系统不能确定哪些小区信号存在突发噪声, 也就不能提前知道类的数量;

(3) DBSCAN 可以发现任意形状的簇, 也就是说, 不会因为定位出来的点的分布非常规而增加类的数量, 从而影响后面的判断。

本文所用 DBSCAN 算法伪代码如表 1 所示。

经过 DBSCAN 聚类, 可将子空间匹配算法估计出的 M 个位置坐标形成一个个簇, 假设在剔除异常点后, 一共有 k 个簇, 即通过 DBSCAN 聚成 k 类 $\{C_1, C_2, \dots, C_k\}$, 每一个类 C_i 内的点表示定位出该点的子空间包含相同的存在突发噪声的小区。计算这 k 个类的中心位置:

表 1 DBSCAN 算法伪代码

算法: DBSCAN

输入: D 为 M 个定位点的点集。 ϵ 为邻域半径。 Minpts 为邻域密度的阈值。

输出: 各点的类标签

算法步骤:

- (1) 将 D 中所有点标记为未观测点;
- (2) 初始化类标签 $b=1$;
- (3) do
- (4) 随机选择 D 中一个未观测的点 p ;
- (5) 将 p 标记为已观测点;
- (6) if p 的 ϵ -邻域内至少有 Minpts 个点
- (7) 创建一个新簇 C , 并将点 p 添加到 C 中;
- (8) 创建一个集合 Q , 它包含 p 的 ϵ -邻域内的所有点;
- (9) for Q 中的每个点 s
- (10) if s 是未观测点
- (11) 标记 s 为已观测点;
- (12) if s 的 ϵ -邻域内至少有 Minpts 个点, 把这些点添加到 Q 中;
- (13) if s 还不是任何簇的成员, 把 s 添加到 C ;
- (14) end for
- (15) 标记 C 中每个点为第 b 类;
- (16) $b=b+1$;
- (17) else 标记 p 为异常点;
- (18) until 没有未观测的点;

$$(ox_i, oy_i) = \frac{1}{\text{card}(C_i)} \sum_{e=1}^M u_{ie} \times (\hat{x}_e, \hat{y}_e), \quad i = 1, 2, \dots, k \quad (12)$$

其中, $\text{card}(C_i)$ 表示第 i 类元素的个数, $u(i, e)$ 用来标识点 $(\hat{x}_e, \hat{y}_e)$ 是否在 C_i 中。

$$u_{ie} = \begin{cases} 1, & (\hat{x}_e, \hat{y}_e) \in C_i \\ 0, & (\hat{x}_e, \hat{y}_e) \notin C_i \end{cases} \quad (13)$$

$$\text{card}(C_i) = \sum_{e=1}^M u_{ie} \quad (14)$$

计算出每一类的中心位置后, 这些位置作为备选输出位置, 下面通过概率模型推算出最有可能的定位准确的点。

3.3 概率模型

显然, 当子空间 $\mathbf{W}_e$ 中不包含有突发噪声的小区信号的时候, 定位结果最可信, 下面需要建立模型找出这一类定位点形成的簇。

假设有 r 个小区的接收信号存在突发噪声, 根据概率学, 随机抽取一个子空间来定位时, 正确点出现的先验概率为

$$P_{\text{Pr}} = \frac{\text{不包含突发噪声的子空间数量}}{\text{总的子空间数量}} = \frac{C_{m-r}^a}{M} \quad (15)$$

前面提到, DBSCAN 聚出的每一类 C_i 内的点表示定位出该点的子空间包含相同的存在突发噪声的小区, 于是 r 可由式(16)确定:

$$r = \lceil \log_2 k \rceil \quad (16)$$

这里的 $\lceil \cdot \rceil$ 表示向上取整, 它是为了防止某些偏离的点聚到一类, 导致类的数目减少。

在剔除异常点后, 随机选取一个点, 它是来自第 i 类的概率为

$$P_i = \frac{\text{card}(C_i)}{\sum_{i=1}^k \text{card}(C_i)} \quad (17)$$

假设第 c 类为定位准确的类, 则 P_c 应该最接近先验概率 P_{pr} , 于是 c 为

$$c = \arg \min_i \left| \frac{\text{card}(C_i)}{\sum_{i=1}^k \text{card}(C_i)} - \frac{C_{m-r}^a}{M} \right| \quad (18)$$

最终估计输出的定位结果为

$$(\hat{x} \ \hat{y}) = (ox_c \ oy_c) \quad (19)$$

4 实验结果分析

4.1 实验环境

为了验证蜂窝网无线信号用于室内定位的可行性以及本文提出的基于 DBSCAN 的子空间匹配算法的有效性和可靠性, 在面积为 $64.6 \text{ m} \times 18.5 \text{ m}$ 的室内大厅和走廊区域采集真实 GSM 蜂窝网无线信号, 共布置 122 个指纹点, 利用路测软件采集覆盖本区域的 25 个小区的 RSSI 值构建离线指纹数据库, 如图 4 所示。其中黑色点为指纹点, 相邻指纹点间隔为 1.2 m。实验时, 随机选取 93 个测试点位置, 采集其信号强度进行算法验证。算法运行软件平台为 Windows 7 系统, 硬件平台为: 处理器 Inter(R) Core(TM) i3-2120, 内存 4 GB。数据采集所用手机终端: 索尼爱立信 k800i, 路测软件: TEMS 9.1。

4.2 WKNN 匹配算法中 K 值的选择

根据选取的 K 值不同, 使用传统的 WKNN 匹配算法进行定位精度分析, 从而选择最佳的 K 值, 后面的算法都是基于此时确定的 K 值进行的验证分

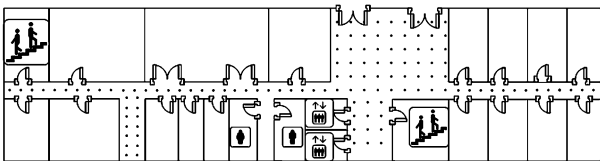


图4 实验环境平面图

析。本文分别选取 K 为 1, 3, 5, 7, 9, 对 93 个测试点用 WKNN 匹配算法进行定位, 根据测试点位置计算定位误差, 然后对这些定位误差进行概率密度估计, 得到的累积误差概率分布(CDF)曲线图如图 5 所示。

由图 5 可知, $K=1$ 时的性能最低, 这是因为此时的系统认为信号距离最近的点即是接收机所处位置, 容易受到异常点干扰, 一旦出现大误差点就无法通过其它邻近点拉回。同时从图中也可以看出, K 值越小, 系统对大误差的抑制效果越差, 这是因为 K 值较大时, 会出现较多的邻近点来纠正大误差点, 但是 K 值太大, 每次定位都会有离接收机较远的点参与, 导致系统的整体定位精度不高。结合图 5 与表 2 可知, $K=3$ 时除了出现少数的较大误差外, 整体性能均优于其它 K 值的性能, 所以, 在本文中 K 取值为 3。

表2 不同 K 值下的定位误差(m)

K 值	平均误差	67%误差	90%误差
1	3.9	4.5	9.2
3	3.3	4.0	6.6
5	3.5	4.0	6.8
7	3.5	4.2	6.9
9	3.6	4.3	7.2

4.3 基于 DBSCAN 的子空间匹配算法

在对文中第 3 节提到的基于 DBSCAN 的子空间匹配算法进行验证之前, 需要选取子空间的维度 a 。考虑到若 a 取值太小, 会因为每次参与定位的小区太少而影响定位精度; a 过大可能使每个子空间都包含发生突发噪声的小区, 而人员干扰一般不会造成太多的小区信号发生跳变, 综合考虑选取子空间的维度 $a=21$ 进行定位, 得到的定位结果与传统的匹配定位算法的对比图如图 6 所示。

图 6(a)给出了 93 个测试点用 WKNN 算法和本文提出的基于 DBSCAN 的子空间匹配定位算法的定位误差, 可以看出, 对于大部分点, 本文算法均在一定程度上对 WKNN 定出的位置进行了优化。结合图 6(b)和表 3 可以看出, 本文所提算法相比传统的 NN, KNN, WKNN 匹配算法, 成功地减少了大误差点的数量, 提高了系统的整体定位性能。

5 结束语

为了实现蜂窝网室内定位, 本文首先对蜂窝网室内信号进行分析, 然后搭建平台采集数据, 证明

蜂窝网室内指纹定位的可行性。针对接收端干扰和数据丢包等情况造成的蜂窝网室内定位精度低的问题,本文提出一种基于 DBSCAN 的子空间匹配算法,该算法充分利用蜂窝网信号丰富的特征,通过划分,构造多个子空间,分离出没有突发噪声的小区信号产生定位点,再利用 DBSCAN 聚类算法在剔除异常点的同时将各子空间定出的位置形成多个类并求出类中心,最后通过概率模型找到最有可能定位准确的类,输出其类中心作为最终定位结果。

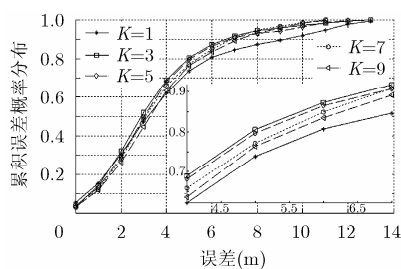
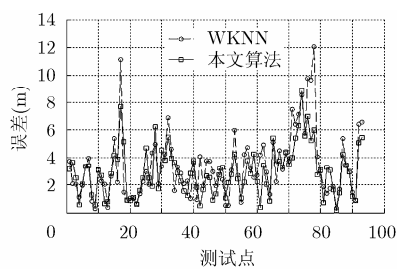
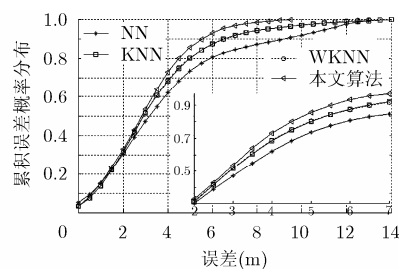


图5 不同 K 值下的累积误差概率分布



(a) 各个测试点用 WKNN 与本文算法的定位误差对比图



(b) 传统算法与本文算法的累积误差分布对比图

通过与传统的匹配算法进行比较,本文算法有效地减少了大误差,提高了整体定位精度。

表3 4种算法的定位误差(m)

算法	平均误差	67%误差	90%误差
NN	3.9	4.5	9.2
KNN	3.4	4.0	6.6
WKNN	3.3	4.0	6.6
本文算法	3.0	3.8	5.6

参考文献

- [1] THI H N V, KEUN H R, and NAMKYU P. A method for predicting future location of mobile user for location-based services system[J]. *Computers & Industrial Engineering*, 2009, 57(1): 91-105. doi: 10.1016/j.cie.2008.07.009.
- [2] KUNG H Y, CHAISIT S, and PHUONG N T M. Optimization of an RFID location identification scheme based on the neural network[J]. *International Journal of Communication Systems*, 2015, 28(4): 625-644. doi: 10.1002/dac.2692.
- [3] 吴楠, 王旭东, 胡晴晴, 等. 基于多LED的高精度室内可见光定位方法[J]. *电子与信息学报*, 2015, 37(3): 727-732. doi: 10.11999/JEIT140725.
WU Nan, WANG Xudong, HU Qingqing, *et al.* Multiple LED based high accuracy indoor visible light positioning scheme[J]. *Journal of Electronics & Information Technology*, 2015, 37(3): 727-732. doi: 10.11999/JEIT140725.
- [4] WANG Yixin, YE Qiang, CHENG Jie, *et al.* RSSI-based bluetooth indoor localization[C]. International Conference on Mobile Ad-hoc and Sensor Networks (MSN), Shenzhen, 2015: 165-171.
- [5] YANG Bo, LEI Yiqun, and YAN Bei. Distributed multi-human location algorithm using naive bayes classifier for a binary pyroelectric infrared sensor tracking system[J]. *IEEE Sensors Journal*, 2015, 16(1): 216-223. doi: 10.1109/JSEN. 2015.2477540.
- [6] 陈兵, 杨小玲. 一种基于概率密度的 WLAN 接入点定位的算法[J]. *电子与信息学报*, 2015, 37(4): 855-862. doi: 10.11999/JEIT140661.
CHEN Bing and YANG Xiaoling. A WLAN access point localization algorithm based on probability density[J]. *Journal of Electronics & Information Technology*, 2015, 37(4): 855-862. doi: 10.11999/JEIT140661.
- [7] WILLEMSSEN T, KELLER F, and STERNBERG H. Concept for building a MEMS based indoor localization system[C]. International Conference on Indoor Positioning and Indoor Navigation (IPIN), Busan, 2014: 1-10.
- [8] WANG Jie, GAO Qinghua, YU Yan, *et al.* Toward robust indoor localization based on Bayesian filter using chirp-spread-spectrum ranging[J]. *IEEE Transactions on Industrial Electronics*, 2012, 59(3): 1622-1629.
- [9] WANG Jie, GAO Qinghua, WANG Hongyu, *et al.* Device-free localization with multi-dimensional wireless link information[J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(1): 356-366.
- [10] HE Jie, LI Shen, PAHLAVAN Kaveh, *et al.* A realtime testbed for performance evaluation of indoor TOA location system[C]. 2012 IEEE International Conference on Communications (ICC), Ottawa, 2012: 482-486.
- [11] HARA Shinsuke, ANZAI Daisuke, YABU Tomofumi, *et al.* Analysis on TOA and TDOA location estimation performances in a cellular system[C]. 2011 IEEE International Conference on Communications (ICC), Kyoto,

- 2011: 1-5.
- [12] LIU Congfeng, YANG Jie, and WANG Fengshuai. Joint TDOA and AOA location algorithm[J]. *Journal of Systems Engineering and Electronics*, 2013, 24(2): 183-188. doi: 10.1109/JSEE.2013.00023.
- [13] TIAN Zengshan, LIU Xindi, ZHOU Mu, *et al.* Mobility tracking by fingerprint-based KNN/PF approach in cellular networks[C]. 2013 IEEE Wireless Communications and Networking Conference (WCNC), Shanghai, 2013: 4570-4575.
- [14] MAHER P S and MALANEY R A. A novel fingerprint location method using ray-tracing[C]. Global Telecommunications Conference (GLOBECOM), Honolulu, 2009: 1-5.
- [15] DRAWIL N, AMAR H, and BASIR O. Cellular network fingerprint localization simulation: A soft computing approach[C]. 2014 IEEE 80th Vehicular Technology Conference, Vancouver, 2014: 1-5.
- [16] 范明, 孟晓峰. 数据挖掘概念与技术[M]. 第3版, 北京: 机械工业出版社, 2012: 307-309.
- FAN Ming and MENG Xiaofeng. Data Mining Concepts and Techniques[M]. Third Edition, Beijing, China Machine Press, 2012: 307-309.
- [17] HONG J and OHTSUKI T. Device-free passive localization from signal subspace eigenvectors[C]. 2014 IEEE Global Communications Conference, Austin, 2014: 430-435.
- [18] DING Genming, TAN Zhenhui, ZHANG Jinbao, *et al.* Regional propagation model based fingerprinting localization in indoor environments[C]. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, London, 2013: 291-295.
- 田增山: 男, 1968年生, 教授, 博士生导师, 研究方向为移动通信、个人通信、GPS及蜂窝网定位系统及其应用技术.
- 王向勇: 男, 1991年生, 硕士生, 研究方向为蜂窝网室内定位技术、人体行为识别.
- 周牧: 男, 1985年生, 教授, 博士, 研究方向为无线定位技术、机器学习与人工智能、凸优化理论.
- 李玲霞: 女, 1976年生, 高级工程师, 硕士生导师, 研究方向为未来移动通信理论与技术、宽带无线接入技术.