

Updating Fingerprint Database by Crowdsourcing for Wi-Fi Indoor Localization^{*}

LI Yanjun^{*} XU Kaifeng SHAO Jianji

(School of Computer Science and Technology Zhejiang University of Technology Hangzhou 310023, China)

Abstract: Wi-Fi fingerprint localization has been well accepted to solve the indoor localization problem, while the most important part in this method is the construction and maintenance of the fingerprint database. A timely update of the fingerprint database is required due to the fluctuation of the received Wi-Fi signal strength. However, having professionals updating the fingerprint database is costly in terms of time and effort. In this paper, a crowdsourcing approach is proposed where the users themselves can evaluate and correct the localization result, thus training and using the localization system at the same time. Particularly a clustering based error detection method is adopted to detect the erroneous input location data, which can effectively avoid the contamination of the fingerprint database. Finally an indoor localization system is developed and practical experiments in real indoor environment show that the proposed approach achieves satisfying localization accuracy over a long time.

Key words: WLAN; indoor localization; crowd-sourcing; location fingerprint; clustering

EEACC: 6250F; 7230

doi: 10.3969/j.issn.1004-1699.2014.12.020

利用众包更新 Wi-Fi 室内定位指纹库的方法研究^{*}

李燕君^{*} 徐凯锋 邵剑集

(浙江工业大学计算机学院 杭州 310023)

摘 要: Wi-Fi 定位是目前较为主流的室内定位方法,而位置指纹库的建立和维护对 Wi-Fi 定位至关重要。Wi-Fi 信号时变性强要求指纹库及时更新。针对由专业人员更新指纹库的人力耗费问题,提出利用众包更新指纹库的方法,允许用户对定位结果进行评价和修正,使得用户在享受定位结果的同时参与到指纹库的维护更新中,特别针对用户的错误修正提出了基于聚类的错误检测方法,能有效避免指纹库被错误指纹污染。开发了室内定位系统,通过在真实室内环境的实验验证了本文提出的方法可以长时间保持较高的定位性能。

关键词: 无线局域网; 室内定位; 众包; 位置指纹; 聚类

中图分类号: TP393

文献标识码: A

文章编号: 1004-1699(2014)12-1692-07

随着大体量建筑的开发数量日益增多以及智能移动终端的普及,人们对室内位置服务的需求正迅速增加。而公共安全、应急救援、大型场馆管理、特殊人群监护、物联网和智慧城市建设等领域都需要准确的室内位置信息。传统的卫星定位导航系统,如 GPS 和北斗,只能提供室外定位导航服务,在室内环境,由于卫星信号受到阻隔加之室内地图的缺失,无法为用户提供精准的位置信息,产生卫星导航到手机用户的“最后一公里”难题。为此,国内外许多公司、高校和科研机构都围绕室内定位导航展开了研究。现有的室内定位技术种类繁多,有紫蜂(ZigBee)^[1]、超声波

定位^[2]、无线局域网(Wi-Fi)^[3-10]、射频标签(RFID)、蓝牙(Bluetooth)、超宽带无线电(UWB)、地磁场强、LED 定位、计算机视觉定位等。从技术成熟与大规模应用的现实角度考虑, Wi-Fi 定位成为当前主流、也是未来最具发展潜力的室内定位方法。

目前 Wi-Fi 室内定位方法中应用最普遍的是位置指纹定位法^[3-7],该方法分为离线训练和在线定位两个阶段。离线训练阶段在待定位区域选定一系列参考点,在这些参考点处采集来自不同无线接入点 AP (Access Point) 的信号强度 RSS (Received Signal Strength) 值,将参考点坐标和对应 AP 的 RSS

项目来源: 国家自然科学基金项目(61003264); 浙江省自然科学基金项目(LY13F020028)

收稿日期: 2014-07-20 修改日期: 2014-10-29

信息存储在数据库中,建立位置指纹数据库;在线定位阶段则依据一定的匹配算法将待测点上收到的相应 AP 的 RSS 信息与数据库中的已有信息进行比较,估计用户当前的位置。该方法的关键之一在于位置指纹数据库的建立和维护。实验表明,室内环境下射频信号传播非常复杂,墙壁、门窗和桌椅等设施以及人员走动会引发射频信号传播的多径和阴影效应,导致在室内固定位置处接收到各个 AP 的 RSS 时变性很强^[7-8];此外,部署的 AP 也可能发生故障或位置的变更。这些都意味着离线阶段建立的位置指纹库并不能一劳永逸,需要经常维护更新,否则会导致定位结果不准确。若采用专人定期更新指纹库的方法,非常耗费人力。为了降低维护更新指纹库的成本,本文提出了利用众包模式更新指纹库的方法,即让用户在享受定位结果的同时参与对指纹库的更新。本文的主要贡献在于:①提出了用户对定位结果的评价和更正机制,一方面利用定位结果更新指纹库,另一方面对用户可能错误的更正进行甄别,使得指纹库得到持续更新;②在智能手机终端实现了室内定位系统,并通过实验验证该系统随时间推移能保持较高的定位准确率。

1 相关工作

基于位置指纹的 Wi-Fi 定位方法最早由微软的研究人员提出,他们开发了相应的室内定位系统 RADAR^[3]。由于 Wi-Fi 热点的广泛布置和智能移动终端的普及,该方法受到了广泛关注。其后,大多数研究工作都集中在对位置指纹匹配算法的设计和改进行上,两类典型的匹配算法是最近邻法^[3]和朴素贝叶斯法^[9]。其中,最近邻法是将定位建模为在 RSS 矢量空间的相似性判断问题,其本质是认为欧氏空间位置相近的点在 RSS 矢量空间的距离也相近,这类方法计算效率高,但结果容易受信号波动影响;而朴素贝叶斯法是在离线阶段采集参考点处的 RSS 统计分布信息,在定位阶段计算被测量点的 RSS 位置指纹在各个参考点处的后验概率,基于最大后验概率准则获得目标的位置估计。由于概率分布能描述 RSS 的随机特性,因此该方法比最近邻法更能适应信号波动,但要获得 RSS 的统计分布信息在离线训练时工作量较大,并且统计模型的选取对定位结果也有较大影响。相关文献中采用的统计模型有简单的高斯模型,也有较为复杂的混合高斯模型、柱状图表法和非参数化的高斯核方法等,这其中涉及到一些参数的选择,只能采用经验值,没有标准可依。

上述文献中不管是最近邻法还是贝叶斯法采用

的都是固定的位置指纹库。虽然基于贝叶斯法的匹配算法在一定程度上应对了 RSS 的时变特性,但是如果指纹库本身不进行更新,随着时间的推移,定位准确率仍然会下降。鉴于由专业人员定期维护指纹库的成本巨大,一些文献提出采用众包模式构建和维护指纹库。文献[10]提出了一种完全由用户自主构建指纹库的方法,基于已有参考点对整个区域的覆盖度决定是否提醒用户标记位置,几乎不需要专业人员参与建库和维护,但是对于较大的室内环境,用户最初使用该系统时会被频繁打扰,增加了用户的负担,会造成用户体验下降;文献[7]提出利用用户的闲散时间为系统上传参考点位置,但这种方法显然高估了一般用户的配合度,且他们没有考虑用户标记有误的情况。与这些方法不同的是,本文一开始构建数据库时,仍然由专业人员完成,以保证初始用户的体验,在更新阶段采用众包模式,让用户参与指纹库的更新。众包是一种分布式的问题解决和生产模式,在文献[8]中被用于构建室内地图。本文创新性地引入用户评价和更正机制,允许用户将定位结果直接作为参考点上传,也允许用户对定位结果进行修正,同时提出了鉴别机制防止用户由于粗心所带来的错误更正和防止用户的恶意修改。

2 利用众包维护指纹库的定位方法

2.1 研究动机

Wi-Fi 定位是利用接收到来自环境中 AP 的 RSS 序列相似程度来判断不同位置的邻近与否,但是固定位置的 RSS 值仍然具有时变性。图 1 显示了一段时间内在不同房间测得房间内 AP 的 RSS 变化情况,采样时间为 5 s,方差依次为 8.1、4.0 和 3.9。这些波动可能由墙壁、门窗和桌椅等设施以及人员走动引起,是无法避免的,并且如果 AP 位置和状态发生改变, RSS 的变化会更加明显。解决短时间 RSS 波动的问题可以在定位时多次采集 RSS 取其平均值,但指纹库若长时间得不到更新,定位精度仍然会随时间推移逐渐下降。为此,本文将围绕如何更新指纹库展开研究。

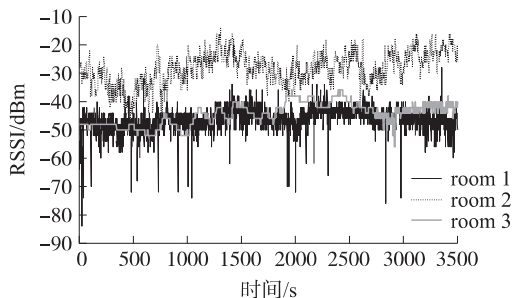


图 1 Wi-Fi 信号强度变化图

2.2 方法概述

本文提出的利用众包维护指纹库的定位方法分为离线训练、在线定位和指纹库更新 3 个阶段。在离线训练阶段,在待定位区域选定一系列参考点,在这些参考点处采集来自不同 AP 的信号强度值,将参考点坐标和对应 AP 的 RSS 信息存储在数据库中,建立初始位置指纹数据库,存于服务器端;在在线定位阶段,用户端将待定位点能扫描到的 AP 及其 RSS 上传至服务器,服务器根据定位算法估算出用户当前的位置返回给用户;在指纹库更新阶段,用户如果对定位结果满意,将定位结果与提交的 AP 及其 RSS 作为一条新的指纹加入指纹库中;如果用户对定位结果不满意,可以更正定位结果,服务器端采用错误检测方法甄别用户的更正是否有误,从而决定是否接受此次更正。如果服务器判定用户更正正确,则将更正的位置和定位时收集到的信息绑定形成新的位置指纹存入指纹库;否则丢弃本次用户的更正。此外,若用户在使用过程中指纹库一直只增不减,会出现指纹库过于臃肿的情况,降低在线定位的效率。为此,系统会根据采集时间定期删除过期的位置指纹,其中位置指纹的保存时长根据位置指纹库的使用频率进行设置。本方法可以使得指纹库可以在用户使用过程中得到不断更新,从而能够适应变化的 Wi-Fi 环境。

2.3 定位算法

本文采用文献[11]中提出的基于 RSS 相似度的定位算法。由于本文研究的重点是指纹库的更新方法,因此没有考虑较为复杂的基于概率模型的定位算法,基于 RSS 相似度的定位算法虽然简单但在实际应用中十分有效。本文的指纹库更新方法同样适用于其他定位算法。假设位置 i 扫描到的 AP 的集合为 A_i ,位置 j 扫描到的 AP 的集合为 A_j ,得到 AP 集合 $A=A_i \cup A_j$, $|A|$ 表示位置 i 和 j 所扫描到的 AP 的总个数;位置 i 扫描到 AP a 的 RSS 值用 $f_i(a)$ 表示,以 dBm 为单位, $|f_i(a)|$ 表示 $f_i(a)$ 的绝对值,若在位置 i 不能扫描到 AP a ,则令 $f_i(a)=0$ 。定义位置 i 和位置 j 的相似度为它们对应的 RSS 矢量的相似度,具体表达式为:

$$S_{ij} = \frac{1}{|A|} \sum_{a \in A} \frac{\min(|f_i(a)|, |f_j(a)|)}{\max(|f_i(a)|, |f_j(a)|)} \quad (1)$$

相似度 S 的取值范围是 $[0, 1]$, S 越大表征两组 RSS 越相似,进而认为位置 i 与位置 j 接近;反之则认为位置 i 与位置 j 远离。下面举例说明相似度的计算。假设环境中有 5 个 AP,编号从 1 到 5,位置 1 和位置 2 扫描到的 AP 物理地址和 RSS 如表 1 和

表 2 所示,根据式(1)得到位置 1 和位置 2 的相似度 $S_{12} = (0/80 + 55/75 + 60/75 + 0/85) / 4 \approx 0.383$ 。一种极端的情况,若 R_1 和 R_2 完全相同,则两个位置相似度为 1,即完全重合。

表 1 位置 1 扫描到的 AP 物理地址及 RSS

AP 编号	M_1	R_1 /dBm
1	14: e6: e4: d2: 35: 6a	-80
2	34: 08: 04: b9: 0f: 10	-75
3	84: 4b: f5: 8c: f6: 37	-75

表 2 位置 2 扫描到的 AP 物理地址及 RSS

AP 编号	M_2	R_2 /dBm
2	34: 08: 04: b9: 0f: 10	-55
3	84: 4b: f5: 8c: f6: 37	-60
4	56: f0: 6d: 7c: 57: 80	-85

指纹库中所有参考点的集合为 L ,将被测量点 j 与指纹库中每个参考点按照式(1)计算相似度,最终定位结果 \hat{P}_j 为与测量点相似度最高的参考点的位置,即:

$$\hat{P}_j = \operatorname{argmax}_{i \in L} S_{ij} \quad (2)$$

2.4 指纹库更新方法

利用众包更新指纹库的方法流程如图 2 所示。具体来说,当服务器按照 2.3 节的定位算法估计出用户当前位置返回给用户时,若用户对定位结果满意,定位结果将作为一条新的指纹存入指纹库;若用户对定位结果不满意,可以对结果进行更正,为了避免用户更正有误对指纹库造成污染,用户更正的位置需接受更正检测,更正检测的具体方法是通过聚类算法将指纹库中的点分成若干簇类,若用户更正的位置与定位结果在同一个簇类,则判定为更正正

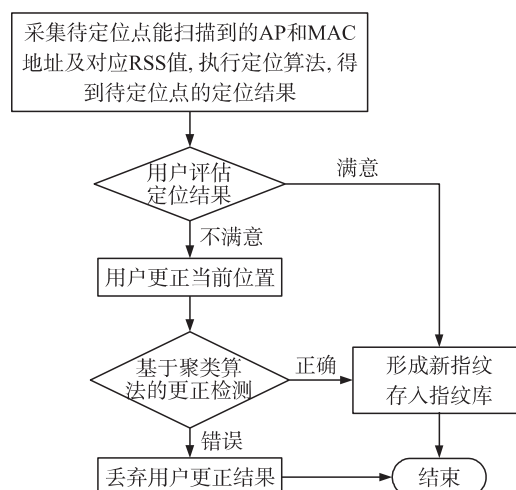


图 2 指纹库更新方法流程图

确,更正的结果作为一条新的指纹存入数据库,否则判定更正有误,更正位置不予接受。这样做的原因是我们认为室内环境能够完全被 AP 信号覆盖,并且在离线训练阶段由专业人员建立的指纹库是高质量的,能达到 90% 以上^[3,10]的定位准确率(定位准确率的定义在第 3 节给出)。但由于 RSS 具有时变性或个别 AP 的位置状态改变,随时间推移定位结果会发生一定偏差,与用户实际位置不符,但偏差不会太远,即使用户选择更正定位结果,更正后的位置应该仍然在系统给出的定位结果“附近”。如果更正位置偏离定位结果太远,该更正位置很可能是有误的。聚类算法则是用于判定两个位置是否邻近的标准。下面将具体介绍聚类算法和算法中涉及的参数选取问题。

2.4.1 聚类算法

室内环境下用户的活动区域不定导致聚类数目无法事先确定,并且房间形状不规则,所以基于划分和层次等聚类方法不能适用。而基于密度的聚类算法,如 DBSCAN^[12],事先不需要确定聚类数目,且可以形成任意形状的簇类,能满足我们的要求。我们采用 DBSCAN 对指纹库中的参考点进行聚类。具体步骤如下:考察指纹库中的某一个点 p ,以 p 为圆心,半径为 ε 的圆被称为 p 的邻域,若 p 的邻域内包含的点数超过阈值 MinP ,则认为点 p 是一个核心对象,邻域中的点和 p 同属于一个簇类,这些邻域中的点将作为下一轮考察对象,分别以这些点为圆心做相同的判断,并通过不断对未访问过的点进行区域查询来扩展它们所在的簇类,直至找出该簇类中所有能形成核心对象的点。算法包含 ε 和 MinP 两个参数,其中 ε 指明邻域半径的大小, MinP 指明做聚类时样本点的 ε 邻域内至少包含的点数。 ε 和 MinP 这两个参数的选取是本算法的关键。

2.4.2 参数选取

首先考虑 ε 的选取。以图 3 为例,假设用户的定位结果为黑色方块所示点 L ,用户将定位结果更正至点 L' 。由于在离线训练阶段由专业人员构建指纹库,定位准确率在 90% 以上^[3,10],此时定位偏差超过房间大小的概率小于 10%。若 ε 取值过小,如图 3(a) 所示,圆圈内的点形成一个聚类,该聚类范围小于房间大小。由于点 L' 不在聚类 A 中,系统会丢弃此次更正,造成用户的正确更正被丢弃的概率增大,不利于指纹库的有效更新。若 ε 过大,如图 3(b) 所示,聚类范围远大于房间大小,则一个聚类可能会覆盖多个房间,系统接受用户更正的概率增大,容易将错误指纹存入数据库。一种极端的情况是如果 ε 无穷大,那么

指纹库中所有点属于同一个簇类,不管用户的更正正确与否,系统都会接受,这样系统就不具备甄别能力。本文聚类的目标是同一房间内的指纹形成一个聚类,方法是将 ε 设置成室内环境中最小房间对角线的一半。在这种情况下,若两个房间的参考点相距较远,如图 3(c) 所示,用户将定位结果 L 更正为 L' 发生错误的概率较大,而此时聚类 C 不会覆盖另一个房间内的参考点,用户的更正会被拒绝;若两个房间的参考点相距较近,如图 3(d) 所示,两个房间的参考点都在定位偏差范围内,而聚类 D 可以将另一个房间内的参考点覆盖,此时用户将位置更正到另一个房间是很可能是正确的,系统会接受用户的更正。值得一提的是,此时有另一种情况,即定位结果本来是正确的,用户却更正到另一个房间,由于更正参考点与定位结果同处于一个聚类,系统也会接受此次更正,从而存入了一条错误的指纹。但众包模式有其独有的优越性,鉴于大部分用户是理智的^[13],即使此次存入了一条错误指纹,造成其他用户定位时结果有误,大部分用户还是会将会这个错误结果纠正过来,将正确的指纹存入数据库,先前错误的指纹会因为长时间不被使用而被删除。

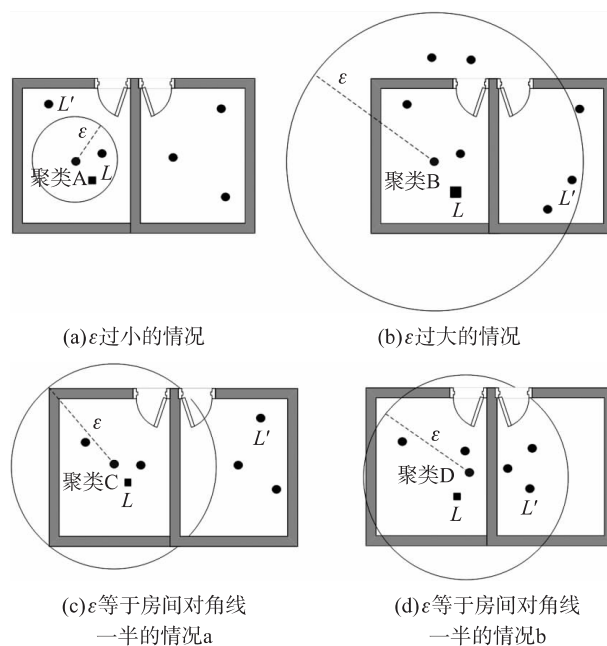


图 3 ε 取值对聚类结果的影响

然后考虑参数 MinP 的选取。 MinP 的取值也会影响聚类的大小。如果聚类过大,一种极端的情况是地图上所有点都在一个聚类中,这样用户的每次更正结果都会被服务器接受,若有用户恶意更正,指纹库的质量将会大大下降;相反,如果聚类过小,比如一个点就是一个聚类,这样用户的每次更正都会被服务器端拒绝,导致数据库得不到有效更新。我们为了确定一个合适的 MinP 阈值,我们将阈值

选择问题转化为 Neyman-Pearson 判决问题。假设 L_s 是服务器给出的定位结果, L_c 是用户更正的位置, 错误更正的检测问题转化为检验以下两个假设:

$$H_1: L_s \text{ 和 } L_c \text{ 来自不同的房间}$$

$$H_0: L_s \text{ 和 } L_c \text{ 来自相同的房间}$$

我们认为一个合适的 MinP 值能达到如下效果: 如果 L_s 和 L_c 来自不同的房间, 系统尽可能将它们分到不同的簇类, 判断为用户更正有误; 而如果 L_s 和 L_c 来自相同的房间, 系统尽可能将它们分到相同的簇类, 判断为用户更正正确。令 x 表示 L_s 和 L_c 不在同一个聚类的情况下 MinP 的取值分布, 令 $P(x|H_1)$ 表示在 L_s 和 L_c 来自不同房间的情况下, 系统通过聚类算法判断出它们不属于同一个聚类的概率, 即检测概率 P_D , $P(x|H_0)$ 表示在 L_s 和 L_c 来自相同房间的情况下, 系统通过聚类算法判断出它们不属于同一个聚类的概率, 即虚警概率 P_{FA} 。在 Neyman-Pearson 模型下, 我们期望找到一个 MinP 值使得在给定的虚警概率下检测概率最大化。由于无法直接获得 x 的分布函数, 我们将在第 3 节通过实验找到最优的 MinP 值。

3 实验结果及分析

为了评估本文提出方法的性能, 我们在浙江工业大学某办公楼 4 层的部分区域进行了实验。具体实验环境如图 4 所示, 区域大小为 $95.4 \text{ m} \times 26.4 \text{ m}$ 。经测试发现, 实验区域共部署了 14 个 AP, 分别位于图中标记的房间内, 它们在办公室内自主放置, 并非为本次实验特意部署。终端采用三星 Galaxy S7562 手机, 操作系统为 Android 4.0.4。离线训练阶段, 共采集 40 个参考点的 RSS 指纹, 参考点均匀分布在走廊、楼梯和各个房间。本文采用如下指标来评价定位性能:

$$\text{定位准确率} =$$

$$\frac{\text{估计位置所在房间与实际所在房间一致的数目}}{\text{定位总数}}$$

其中, 走廊区域也进行了划分, 走廊和房间标号如图 4 所示。我们首先通过实验找到最优的聚类参数 MinP , 然后将本文提出的更新指纹库的方法与不更新指纹库的方法和无条件更新指纹库的方法进行了比较。

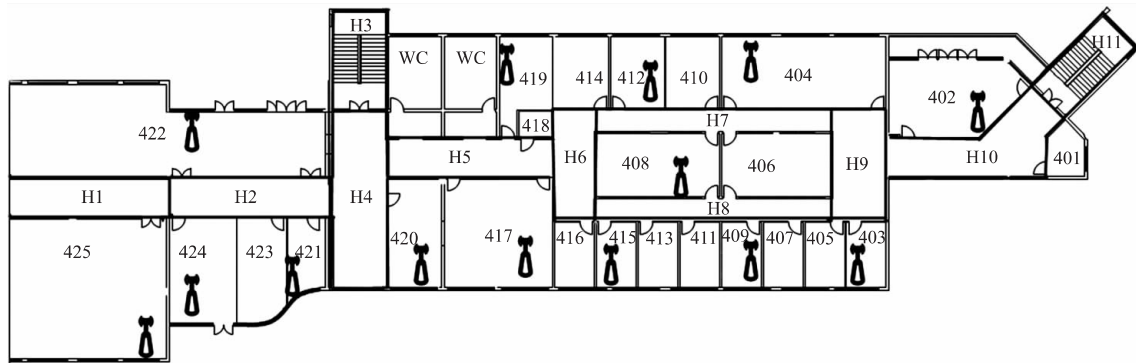


图 4 实验环境平面图

3.1 系统概述

系统由终端和服务端两个部分组成, 系统结构如图 5 所示。终端负责收集 Wi-Fi 信息, 并提供交互界面。服务端提供了 3 种服务, 分别是指纹库共享, 响应用户的定位需求和甄别用户更正。服务端存储来自各个终端提交的 Wi-Fi 指纹, 形成一个指纹数据库, 无论哪个用户提交了新指纹, 该数据库都会得到更新, 从而使其他用户也能享受到最新的指纹库。

为了克服短时间内 RSS 不稳定对定位的影响,

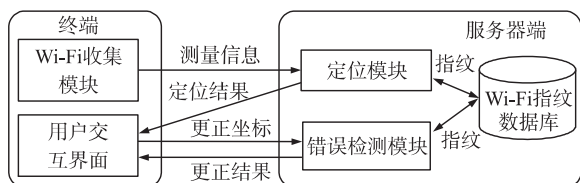


图 5 系统结构

终端在收集 Wi-Fi 信息时, 连续扫描 3 次, 选取 3 次扫描中 RSS 均超过 -80 dBm 的热点, 计算其 RSS 均值上传给服务端。为了让用户乐于帮助完善指纹库, 我们在终端设计了用户友好的交互界面, 如图 6 所示。区别于以往用文本方式提供定位结果的方法

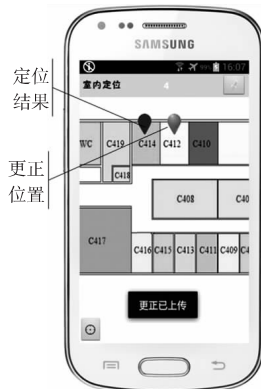


图 6 室内定位系统终端界面

(例如只提供用户当前所在房间编号^[4]) ,本系统具备室内地图加载和显示功能,用户可以加载所在建筑物的楼层平面图,直观的从室内平面图上查看自己当前的位置,也可以通过拖拽位置标注更正定位结果。

3.2 聚类参数选取

为了找到最优的 $\text{Min}P$ 取值,我们对一系列不同的 $\text{Min}P$ 取值进行了测试实验。图 7 和图 8 分别显示了随着 $\text{Min}P$ 取值从 1 变化到 4,虚警概率 P_{FA} 和检测概率 P_{D} 的变化情况,对于每个 $\text{Min}P$ 取值,我们做了 50 次测试。当 $\text{Min}P=1$ 时,虚警概率 P_{FA} 和检测概率 P_{D} 均为 100%,这意味着用户所有的更正都会被系统接受。随 $\text{Min}P$ 值增大,虚警概率 P_{FA} 减小,当 $\text{Min}P=4$ 时, P_{FA} 接近于 0,这意味着当 $\text{Min}P$ 取值足够大时,系统不会接受用户的任何更正;同样,检测概率 P_{D} 也随 $\text{Min}P$ 值增大逐渐减小,但下降的幅度明显小于虚警概率 P_{FA} 。结合图 7 和图 8,我们可以发现,对于给定的虚警概率上限 P_{FA}^* ,该 P_{FA}^* 所对应的 $\text{Min}P$ 值能使检测概率达到最大。例如,给定系统容许的虚警概率为 15%,要使得系统的检测概率最大, $\text{Min}P$ 应取值为 3,此时检测概率可以达到 80%。在后续的实验中, $\text{Min}P$ 均取值为 3。

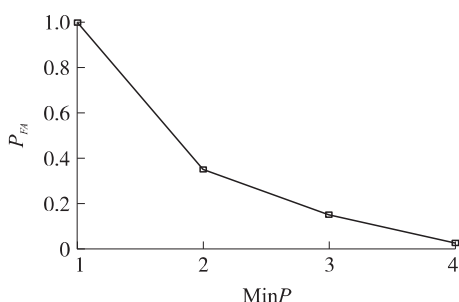


图 7 P_{FA} 随 $\text{Min}P$ 值增大的变化情况

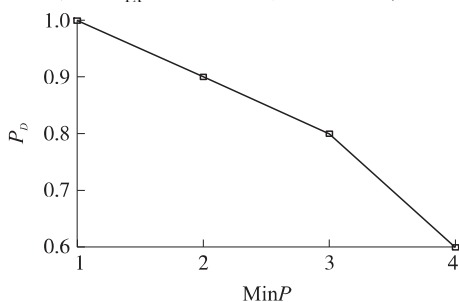


图 8 P_{D} 随 $\text{Min}P$ 值增大的变化情况

3.3 与不更新和人工更新指纹库的方法比较

为了验证本文提出的更新指纹库方法的有效性,我们分别基于众包模式更新的指纹库、不更新的指纹库和人工更新的指纹库做定位。实验时间持续一周,得到的定位准确率在一周内的变化情况如图 9 所示。从图中可以看出,不更新指纹库会使定位准确率迅速下降,每隔 24 h 进行一次人工更新得到的定位准确

率最高,每隔 48 h 进行一次人工更新得到的定位准确率有明显的波动,而本文提出的利用众包更新指纹库的方法在一周内定位准确率能保持在 80% 以上,且无需专业人员参与指纹库的更新。这说明本文提出的众包更新指纹库的方法能够适应变化的 Wi-Fi 环境,一定程度上保持指纹库的质量。值得一提的是,由于本次实验中参与定位实验的志愿者人数不多,且随时间推移对位置指纹的更正积极性下降,所以定位准确率随时间推移略有下降,其定位准确率不及由专业人员每隔 24 h 进行指纹库更新的准确率。但是,我们相信在实际应用中,如果有更多的用户参与指纹库的更新,系统定位准确率仍能保持较高的水平。未来我们也将进一步研究提高用户参与指纹库更新积极性的激励机制。

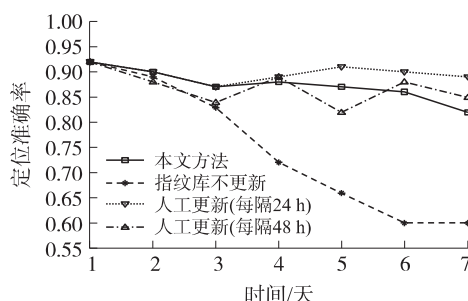


图 9 本文提出方法与不更新和人工更新指纹库方法的比较

3.4 与无条件更新指纹库的方法比较

甄别用户对定位结果的更正是否有误是本系统的关键,为了验证本文提出方法的有效性,我们测试了让志愿者恶意更改定位结果的情况,将本文提出的方法与无条件接受用户更新指纹库的方法进行比较,得到定位准确率随用户恶意更改错误率增大的变化情况,如图 10 所示。从图中我们可以看出,如果无条件接受用户的更正,定位准确率随着用户更正错误率的增加而快速减小,当用户更正错误率为 40% 时,准确度已经下降到 50% 以下;而用本文提出的聚类算法对用户的更正做了进一步检测,能丢弃一部分错误的更正,维持指纹库的质量,使得定位准确率随用户更正错误率的增加下降较为缓慢。从图 10 中可以

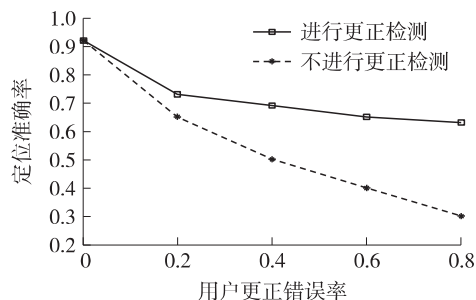


图 10 本文提出方法与无条件更新指纹库方法的比较

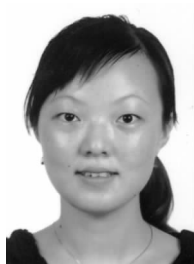
看出,即使更正错误率高达 80%,我们提出方法的准确率依然可以保持 60% 以上。

4 结束语

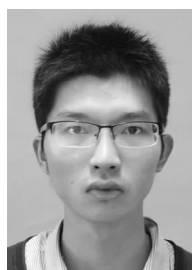
指纹库的维护更新对 Wi-Fi 室内定位非常重要,直接影响到 Wi-Fi 定位技术的应用推广。本文提出的利用众包更新指纹库的方法允许用户对定位结果进行评价和更正,使得用户在享受定位结果的同时参与到指纹库的维护更新中,特别针对用户的错误更正提出了基于聚类的错误检测方法,能有效避免指纹库被普通用户上传的错误指纹污染。开发了室内定位系统,通过在真实室内环境的实验验证了本文提出方法的有效性。本文的研究解决了由专业人员维护指纹库耗费巨大人力成本的问题,使得定位系统可以在较长时间内保持较高的定位性能,如果在实际运行中有更多用户使用该系统并参与指纹库的更新,有望突破 Wi-Fi 定位大规模应用的壁垒。

参考文献:

- [1] 邵丽鹏,朱梅冬,杨丹. 基于 ZigBee 的加权质心定位算法的仿真与实现[J]. 传感技术学报, 2010, 23(1): 149-152.
- [2] 韩霜,罗海勇,陈颖,等. 基于 TDOA 的超声波室内定位系统的设计与实现[J]. 传感技术学报, 2010, 23(3): 347-353.
- [3] Bahl P, Venkata N P. Radar: an in-Building RF-Based User Location and Tracking System[C]//The 19th Annual Joint Conference of the IEEE Computer and Communications Societies (Infocom). Tel Aviv: IEEE, 2000: 775-784.
- [4] Bruno L, Robertson P. WiSLAM: Improving FootSLAM with WiFi [C]//International Conference on Indoor Positioning and Indoor Navigation (IPIN). Guimaraes: IEEE, 2011: 1-10.
- [5] 宋震龙,蒋刚毅,黄晔,等. 基于偏度-峰度检验的无线局域网室内定位算法[J]. 通信学报, 2012, 33(5): 99-105.
- [6] 赵方,罗海勇,马严,等. 基于公共信标集的高精度射频指纹定位算法[J]. 计算机研究与发展, 2012, 49(2): 243-252.
- [7] Bolliger P, Partridge K, Chu M, et al. Improving Location Fingerprinting through Motion Detection and Asynchronous Interval Labeling [C]//The 4th International Symposium on Location and Context Awareness (LoCA). Heidelberg: Springer, 2009: 37-51.
- [8] Wu C, Yang Z, Liu Y, et al. WILL: Wireless Indoor Localization without Site Survey [J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(4): 839-848.
- [9] Seshadri V, Zaruba G V, Huber M. A Bayesian Sampling Approach to in-Door Localization of Wireless Devices Using Received Signal Strength Indication [C]//The 3rd IEEE International Conference on Pervasive Computing and Communications (PerCom). Hawaii: IEEE, 2005: 75-84.
- [10] Park J, Charrow B, Curtis D, et al. Growing an Organic Indoor Location System [C]//The 8th International Conference on Mobile Systems, Applications, and Services (MobiSys). San Francisco: ACM, 2010: 271-284.
- [11] Wang H, Sen S, Elgohary A, et al. No Need to War-Drive: Unsupervised Indoor Localization [C]//The 10th International Conference on Mobile Systems, Applications, and Services (MobiSys). Low Wood Bay: ACM, 2012: 197-210.
- [12] Ester M, Kriegel H P, Sander J, et al. A Density-Based Algorithm for Discovering Clusters in Large Spatial Database with Noise [C]//The 2nd International Conference on Knowledge Discovery and Data Mining (KDD). Oregon, Portland: ACM, 1996: 226-231.
- [13] Jenkins H, Deuze M. Convergence Culture [J]. The International Journal of Research into New Media Technologies, 2008, 14(1): 5-12.



李燕君(1982-),女,江苏南通人,博士,副教授,主要研究领域为无线网络的协议与算法, yjli@zjut.edu.cn;



徐凯锋(1990-),男,研究生,主要研究领域为无线网络,室内定位, zjutxkf@163.com;



邵剑集(1991-),男,硕士生,主要研究领域为无线网络协议,算法仿真, shaojianji@126.com。