

DOI:10.19651/j.cnki.emt.2211486

# 基于投票机制的室内 WiFi 指纹定位算法<sup>\*</sup>

王开亮<sup>1</sup> 谢亚琴<sup>1</sup> 宦海<sup>2</sup> 周莉莉<sup>1</sup>

(1.南京信息工程大学电子与信息工程学院 南京 210044; 2.南京信息工程大学人工智能学院(未来技术学院) 南京 210044)

**摘要:** 针对传统室内 WiFi 指纹定位算法中单个距离度量的局限性且未考虑到 dBm 表示与功率之间的关系的问题,提出一种基于投票机制的室内 WiFi 指纹定位算法。在采集到接收信号强度(RSS)数据后,首先,对 RSS 数据进行预处理;然后,基于投票机制对每种距离度量选中的近邻点取交集组成公共近邻点,并统计每个公共近邻点出现的频率;最后,通过概率加权得到最终定位结果。实验结果表明,所提出方法的定位精度为 1.63 m,与 K 近邻(KNN)、斯皮尔曼(Spearman)和肯德尔相关系数(KTCC)方法的定位精度相比,平均定位精度分别提升了 10%、33%和 58%。此外,与 MAN2 数据集的最优定位精度 1.86 m 相比,定位精度提高了 12%。

**关键词:** 室内定位;WiFi 指纹;距离度量;KNN;投票机制

**中图分类号:** TN925 **文献标识码:** A **国家标准学科分类代码:** 510.5015

## Indoor WiFi fingerprinting algorithm based on voting mechanism

Wang Kai liang<sup>1</sup> Xie Ya qin<sup>1</sup> Huan Hai<sup>2</sup> Zhou Lili<sup>1</sup>

(1. School of Electronics &amp; Information Engineering, Nanjing University of Information Science &amp; Technology, Nanjing 210044, China; 2. School of Artificial Intelligence (School of Future Technology), Nanjing University of Information Science &amp; Technology, Nanjing 210044, China)

**Abstract:** Aiming at the limitation of a single distance metric in the traditional indoor WiFi fingerprinting algorithm and the relationship between dBm representation and power is not considered, an indoor Wi-Fi fingerprinting algorithm based on voting mechanism is proposed. After collecting the received signal strength (RSS) data, first, preprocess the RSS data. Then, based on the voting mechanism, the nearest neighbors selected by each distance metric are intersected to form common neighbors, and count each the frequency of common neighbor points. Finally, the final positioning result is obtained by probability weighting. Experimental results show that the proposed method achieves a localization accuracy of 1.63 m, and the average localization accuracy is improved by 10%, 33%, and 58%, respectively, compared with the localization accuracy of KNN, Spearman, and KTCC methods. Furthermore, the localization accuracy is improved by 12% compared to the optimal localization accuracy of 1.86 m in the MAN2 dataset.

**Keywords:** indoor location; WiFi fingerprint; distance metric; KNN; voting mechanism

## 0 引言

近年来,随着物联网的发展,人们对于基于位置的服务(location based service, LBS)需求越来越多,进而推动了定位技术的发展<sup>[1]</sup>。全球定位系统(global positioning system, GPS)<sup>[2]</sup>可以满足室外环境中的大部分定位需求,但由于建筑物内部难以接收到 GPS 信号,因此无法在室内环境中工作。为了克服这一缺点,室内定位技术应运而生并不断发展。

根据测量设备的不同,现有的室内定位技术可以分为

基于超宽带(ultra wide band, UWB)<sup>[3]</sup>、射频识别器(radio frequency identification, RFID)<sup>[4]</sup>、可见光<sup>[5]</sup>、蓝牙<sup>[6]</sup>和基于 WiFi<sup>[7]</sup>的定位技术等。其中,UWB 技术成本高昂,因为它需要专门的基础设施。RFID 技术使用射频信号的双向通信和数据交换进行定位,其成本低,但射频信号的传输距离短。可见光技术调制白色的发光二极管(LED)可以提供位置信息以获得高精度的位置,但如果光线被遮挡,则无法工作。蓝牙技术具有高安全性、低功耗和低成本的优点,但蓝牙的信号稳定性较差,通信范围较短。

由于 WiFi 接入点(access point, AP)被广泛部署在购

收稿日期:2022-09-21

<sup>\*</sup> 基金项目:国家自然科学基金(62001238)项目资助

物中心、办公楼、工厂等建筑中,同时 WiFi 信号很容易被智能手机等移动设备接收,因此 WiFi 指纹定位技术成为室内定位研究的热点。基于 WiFi 指纹的室内定位方法主要分为离线和在线两个阶段<sup>[8]</sup>。离线阶段,在每个参考点(reference point, RP)采集接收到的多个接入点的信号强度,与其对应的位置信息共同构成指纹数据库。在线阶段,当接收到用户的指纹时,利用定位算法检索指纹数据库,找到几个最匹配的指纹,并取出其对应的位置作为近邻点来预估用户位置。该方法不需要精确的时钟同步和额外的硬件设别,且易于实现,因此众多学者对其展开研究<sup>[9-12]</sup>。

Bahl 等<sup>[13]</sup>提出的 KNN(K nearest neighbor)是一种经典的定位算法,其核心思想是通过计算用户与参考点之间指纹的欧氏距离来衡量其相似性,将具有最高相似性的  $K$  个参考点位置进行平均得到定位结果,但是 KNN 方法对不同近邻点都赋予了相同的权重,没有考虑到不同近邻点对于用户的贡献。Hou 等<sup>[14]</sup>对 KNN 进行了改进,提出了考虑不同参考点权重的 WKNN(weight KNN)方法。该方法将用户与前  $K$  个参考点指纹之间的欧氏距离的倒数作为每个近邻点的权重,再利用这些权重进行加权得到最终定位结果。但是,由于 RSS(received signal strength)指纹的多变和时变特性,基于 KNN 和 WKNN 的定位方法可信度不高。

在进行上述指纹匹配定位时,通常都是使用原始 RSS 值,单位为 dBm。Xie 等<sup>[15]</sup>对 RSS 值进行归一化处理,但是处理之后的值将会呈现线性关系,而实际功率与 dBm 表示之间是非线性的。因此,这种情况下,如果直接对 dBm 进行归一化就会进一步引入误差。为此,Torres 等<sup>[16]</sup>考虑到 dBm 表示与功率之间的关系,提出了两种备选方案来替代 RSS 值,实验结果表明,将 dBm 进行处理后,可进一步提高定位精度。

此外,大多数室内指纹定位技术都是通过单一的距离度量或相似度度量来衡量指纹之间的相似度,进而选取近邻点进行匹配定位。不同的距离度量方法会对定位结果产生影响。现有的距离度量方法包括欧氏距离、曼哈顿距离、Spearman 距离等。欧氏距离适用于比较空旷的场景,而曼哈顿距离则适用于办公场景,Spearman 距离适用于复杂多变的室内场景。在大多数情况下欧氏距离常被用来计算指纹之间的距离,但由于室内环境的不同,欧氏距离并不都是距离匹配中的最佳距离。Torres 等<sup>[17]</sup>在 16 个公共数据集使用 KNN 算法评估现有研究方法的定位精度和匹配时间,实验结果表明,采用 Sørensen 代替欧氏距离进行匹配将会提高定位精度。Xie 等<sup>[18]</sup>提出了一种改进的基于 Spearman 距离的  $K$  近邻算法,通过 Spearman 距离来衡量指纹之间的相似度,可以减少信号衰减和环境影响带来的干扰。Ma 等<sup>[19]</sup>提出 KTCC 相关系数来衡量用户指纹与参考点指纹之间的相似性,但是未考虑到 RSS 值,因此定位误差较大。Peng 等<sup>[20]</sup>将 RSS 的空间距离和物理距离用于

RP 选择,并基于这两个距离的融合加权算法进行位置计算,但这两个距离是成正比的,因而不能真正达到融合的效果。以前大多数的工作基本都是基于单个距离度量或者相似度度量来进行匹配,由于单个距离度量的局限性,通过其选中的近邻点的准确性未知。

为了解决上述问题,本文提出了一种基于投票机制的室内 WiFi 指纹定位算法。首先对 dBm 表示进行预处理。由于实际情况下,室内场景并不明确,因此,本文创造性地提出通过联合多个距离度量,对它们选中的近邻点进行交集,得到公共近邻点,最大程度保证了近邻点的准确性。同时根据公共近邻点出现的频次来计算权重,考虑了其对于用户位置的贡献程度。仿真结果表明,在室内具体场景不明确的情况下,多个距离度量联合的匹配精度高于单个距离度量的匹配精度。

## 1 系统模型

如图 1 所示,本文所提出的算法分为离线和在线两个阶段。离线阶段,通过采集每个参考点的指纹,与其对应的位置信息共同构成了指纹库。在线阶段,首先对指纹库和用户指纹进行预处理。然后,基于投票机制选出公共近邻点并统计它们出现的频率,通过概率加权得到最终定位结果。

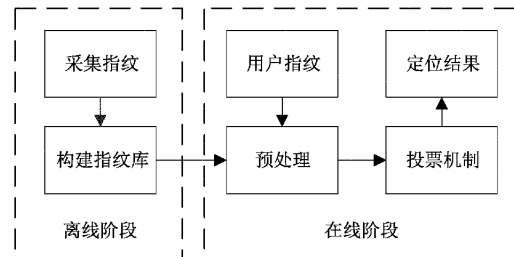


图 1 算法模型图

值得注意的是,Torres 等<sup>[16]</sup>对 51 种距离度量进行了研究,他们将所有距离度量分为了 8 组,每组内的距离度量虽然提供不同的距离或相似性值,但当它们用于基于距离的排序算法(如 KNN)时,它们是等价的。因此本文从这 8 组中任意选择一种,共形成了 8 种距离度量或相似度度量(Euclidean  $L_2$ 、Sørensen、City\_Block  $L_1$ 、Squared  $\chi^2$ 、Jaccard、Fidelity、Hellinger 和 Topsøe),这些度量构成了投票机制的基础,表达式分别如式(1)~(8)所示。

$$d_{ED} = \sqrt{\sum_{n=1}^{N_{AP}} |p_n - q_n|^2} \quad (1)$$

$$d_{sur} = \frac{\sum_{n=1}^{N_{AP}} |p_n - q_n|}{\sum_{n=1}^{N_{AP}} (p_n + q_n)} \quad (2)$$

$$d_{CB} = \sum_{n=1}^{N_{AP}} |p_n - q_n| \quad (3)$$

$$d_{sq} = \sum_{n=1}^{N_{AP}} \frac{(\mathbf{p}_n - \mathbf{q}_n)^2}{(\mathbf{p}_n + \mathbf{q}_n)^2} \quad (4)$$

$$s_{Jac} = \frac{\sum_{n=1}^{N_{AP}} \mathbf{p}_n \mathbf{q}_n}{\sum_{n=1}^{N_{AP}} \mathbf{p}_n^2 + \sum_{n=1}^{N_{AP}} \mathbf{q}_n^2 - \sum_{n=1}^{N_{AP}} \mathbf{p}_n \mathbf{q}_n} \quad (5)$$

$$d_{Jac} = 1 - s_{Jac} = \frac{\sum_{n=1}^{N_{AP}} (\mathbf{p}_n - \mathbf{q}_n)^2}{\sum_{n=1}^{N_{AP}} \mathbf{p}_n^2 + \sum_{n=1}^{N_{AP}} \mathbf{q}_n^2 - \sum_{n=1}^{N_{AP}} \mathbf{p}_n \mathbf{q}_n}$$

$$s_{Fid} = \sum_{n=1}^{N_{AP}} \sqrt{\mathbf{p}_n \mathbf{q}_n} \quad (6)$$

$$d_H = \sqrt{2 \sum_{n=1}^{N_{AP}} (\sqrt{\mathbf{p}_n} - \sqrt{\mathbf{q}_n})^2} \quad (7)$$

$$d_{Top} = \sum_{n=1}^{N_{AP}} \{ \mathbf{p}_n \ln[2\mathbf{p}_n / (\mathbf{p}_n + \mathbf{q}_n)] + \mathbf{q}_n \ln[2\mathbf{q}_n / (\mathbf{p}_n + \mathbf{q}_n)] \} \quad (8)$$

其中,  $d$  和  $s$  分别表示根据每种距离度量或相似度量来衡量指纹向量  $\mathbf{p}$  和  $\mathbf{q}$  的距离或相似度,  $N_{AP}$  表示 AP 的数量,  $\mathbf{p}$  和  $\mathbf{q}$  均表示各自接收到所有 AP 的信号强度值构成的指纹向量,  $\mathbf{p}_n$  和  $\mathbf{q}_n$  均表示各自接收到的来自于第  $n$  个 AP 的信号强度值。

## 2 算法描述

### 2.1 构建指纹库

离线阶段,通过在每个参考点采集指纹,它们的位置和指纹信息共同构成了指纹库。指纹库  $\boldsymbol{\varphi}$  可表示为:

$$\boldsymbol{\varphi} = \{ (\mathbf{FP}_1, \mathbf{P}_1), (\mathbf{FP}_2, \mathbf{P}_2), \dots, (\mathbf{FP}_{N_{RP}}, \mathbf{P}_{N_{RP}}) \} \quad (9)$$

其中,  $\mathbf{FP}_i$  和  $\mathbf{P}_i$  分别代表了第  $i$  个 RP 指纹和位置坐标,  $N_{RP}$  是所有参考点的总数。 $\mathbf{FP}_i$  可表示为  $1 \times N_{AP}$  的向量,如下式所示:

$$\mathbf{FP}_i = [RSS_i^1 \quad RSS_i^2 \quad \dots \quad RSS_i^{N_{AP}}] \quad (10)$$

其中,  $RSS_i^n$  表示第  $i$  个 RP 接收到来自于第  $n$  个 AP 的信号强度。

同理,用户的实时指纹  $\mathbf{FP}^{user}$  可表示为:

$$\mathbf{FP}^{user} = [RSS^1 \quad RSS^2 \quad \dots \quad RSS^{N_{AP}}] \quad (11)$$

其中,  $RSS^n$  表示用户接收到来自于第  $n$  个 AP 的信号强度。

### 2.2 预处理

指纹识别技术依赖于 RSS 值。RSS 值表示特定接收到的 WiFi(IEEE 802.11)无线信号的功率,用 dBm 表示。在这种特定情况下,根据式(12),它表示接收到的信号功率与 1 mW 参考功率之间的比值。

$$RSS = 10 \times \lg(\text{Power(mW)}) \quad (12)$$

图 2 显示了 RSS 值(dBm)与信号功率(mW)之间的关系。可以注意到,信号强度值的相同差异具有不同的功率差。例如,0 和 -10 dBm 对应于功率差  $0.9 \text{ mW}(9 \times 10^{-1} \text{ mW})$ , -50 和 -60 dBm 为  $0.000\,009 \text{ mW}(9 \times 10^{-6} \text{ mW})$ , -90 和 -100 dBm 为  $0.000\,000\,000\,9 \text{ mW}(9 \times 10^{-10} \text{ mW})$ 。然而,对于这 3 种情况,现有的 WiFi 指纹识别技术没有考虑到 dBm 表示与功率之间的关系。

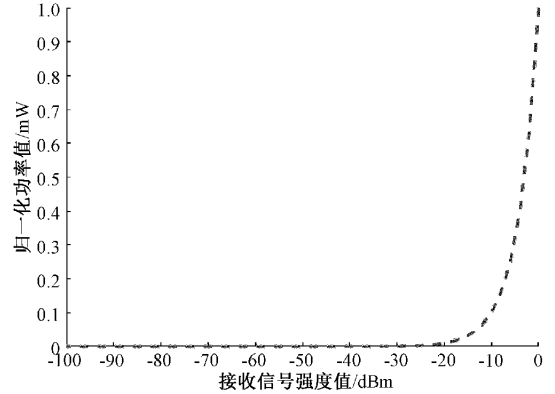


图 2 RSS 值(dBm)与信号功率(mW)之间的关系

因此,本文采用 4 种预处理方式来表示 RSS 值,它们分别是 ZeroToOneNormalized、Positive、Exponential 和 Powed。在 Positive 预处理方式中,只需减去最小值,如式(13)所示。

$$\text{Positive} : RSS - \min \quad (13)$$

其中,  $\min$  是最小的 RSS 值。因此,低值代表信号较弱,而更高的值表示信号更强。在此表示中,最低可能值为 0,用于表示未检测到 AP。类似地,ZeroToOneNormalized, 如式(14)所示。

$$\text{ZeroToOneNormalized} : \frac{\text{Positive}}{-\min} \quad (14)$$

该值与 Positive values 成正比,但强度值在 0~1 范围内归一化。以上两种预处理方法(Positive 和 ZeroToOneNormalized)均保持了原始值的线性。

除此之外,Exponential 和 Powed 预处理方法也被用来代替 RSS 值,以打破原始强度值的线性关系,分别如式(15)和(16)所示。

$$\text{Exponential} : \frac{\exp\left(\frac{\text{Positive}}{\alpha}\right)}{\exp\left(\frac{-\min}{\alpha}\right)} = \exp\left[\frac{1}{\alpha}(\text{Positive} + \min)\right] \quad (15)$$

$$\text{Powed} : \frac{(\text{Positive})^\beta}{(-\min)^\beta} = (\text{ZeroToOneNormalized})^\beta \quad (16)$$

其中,分母常数( $\alpha$ )设置为 24,后者的指数( $\beta$ )设置为数学常数  $e^{[16]}$ 。

为了更直观地表示 RSS 值与预处理之后的值之间的

关系,本文设置 RSS 值取值范围为  $-2 \sim -100$  dBm 之间,利用 3 种预处理方式对原始 RSS 值进行预处理,由于 Positive 和 ZeroToOneNormalized 预处理之后的值是等价的,故只考虑 ZeroToOneNormalized 方法,预处理之后的值与原始信号强度值关系如图 3 所示。

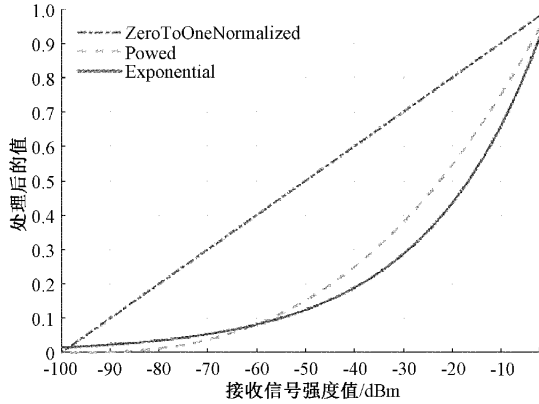


图 3 预处理之后的值与原始信号强度值关系图

如果接收机靠近 AP,则接收到的 RSS 值随时间变化较小。因此,当 RSS 值在较强范围内发生变化时,其预处理之后的值变化也随之增大。但是当 AP 距离接收机较远时,接收机端的信号强度值通常更容易发生信号波动。同时,AP 和接收机之间如果存在人或移动物体,那么信号也会随之波动。因此,当 RSS 值在较弱范围内发生变化时,其预处理之后的值变化也不大。Exponential 和 Powered 这两种预处理方法更倾向于按实际情况表示 RSS 值,例如  $-100$  和  $-90$  dBm (较低的信号)预处理之后的值表示相同,但  $-20$  和  $-10$  dBm (非常强的信号)预处理之后的值差异较大。而 ZeroToOneNormalized 预处理方法未考虑到 RSS 值和功率之间的关系,不管 RSS 值的大小如何,预处理之后的值始终与 RSS 值保持线性关系。

经过上述预处理过程,用户指纹  $FP^{user}$  替换为  $\vec{FP}^{user}$ ,指纹库  $\phi$  与指纹  $FP_i$  分别用  $\vec{\phi}$  和  $\vec{FP}_i$  替换。

### 2.3 投票机制

传统的 KNN 和 WKNN 算法采用参考点和用户之间的欧氏距离来选择  $K$  个最近的参考点进行位置匹配。KNN 算法将选中的  $K$  个近邻点坐标直接进行平均即可求得用户的位置,不同的是,WKNN 采用加权平均。考虑到近邻点到用户的距离不同,分配给近邻点的权重也不同,权重与距离成反比,如式(17)和(18)所示。

$$P_{user}^{KNN} = \frac{1}{K} \sum_{k=1}^K P_k \quad (17)$$

$$P_{user}^{WKNN} = \frac{\sum_{k=1}^K \frac{1}{E_k + eps} P_k}{\sum_{k=1}^K \frac{1}{E_k + eps}} \quad (18)$$

其中,  $P_{user}^{KNN}$  和  $P_{user}^{WKNN}$  分别表示利用 KNN 和 WKNN 计算出来的用户的预测位置,  $P_k$  表示第  $k$  个近邻点的位

置,  $E_k$  代表第  $k$  个近邻点与用户之间的欧氏距离,  $eps$  表示一个接近于 0 的正数,以防止分母等于 0 的情况。WKNN 算法相比于 KNN 算法获得更好的定位精度,但两者均根据单个距离度量来选取近邻点,都易受到环境影响,选取到异常近邻点,导致定位精度不高。

为了进一步提高定位精度,本文利用投票机制对用户进行定位。首先,将  $FP^{user}$  和  $\vec{FP}_i$  分别代入  $p$  和  $q$ , 根据每个距离度量计算出用户与第  $i$  个参考点指纹之间的距离,如式(1)~(8)所示。然后,分别选取  $K$  个距离最小的参考点作为每个距离度量的近邻点。接着,将每种距离度量选出的  $K$  个近邻点取交集,构成公共近邻点  $P_{com}$ , 如式(19)所示。

$$P_{com} = [P_{com}^1 \ P_{com}^2 \ \dots \ P_{com}^{N_{com}}] \quad (19)$$

其中,  $P_{com}^c$  表示第  $c$  个公共近邻点,  $N_{com}$  表示公共近邻点的数量。

每个公共近邻点的权重由这个公共近邻点出现在所有距离度量选出的近邻点中的次数决定。出现的频率越高,其权重越大,反之亦然。最后,执行 WKNN 得到最终定位结果,用户的预测位置  $P_{user}^{proposed}$  可由式(20)得到:

$$P_{user}^{proposed} = \frac{\sum_{c=1}^{N_{com}} (count_{com}^c \times P_{com}^c)}{\sum_{c=1}^{N_{com}} count_{com}^c} \quad (20)$$

其中,  $count_{com}^c$  表示第  $c$  个公共近邻点的频次。

投票机制具体流程如下:

步骤 1) 根据式(1)~(8)计算每个距离度量中用户指纹  $\vec{FP}^{user}$  与第  $i$  个参考点指纹  $\vec{FP}_i$  之间的距离

步骤 2) 选取每个距离度量中前  $K$  个距离较小的参考点作为它们的近邻点

步骤 3) 取每个距离度量的近邻点的交集作为公共近邻点  $P_{com}$

步骤 4) 计算公共近邻点的权重

步骤 5) 执行 WKNN 操作

步骤 6) 输出定位结果  $P_{user}^{proposed}$

为了更形象地理解投票机制的过程,本文以表 1 为例,选取 5 种距离度量进行匹配。假设选取近邻点的数量  $K=6$ ,即每种距离度量将会选出 6 个近邻点,执行 KNN 操作可求得对应的预测位置。可以看出,每种距离度量预测的位置不同,此外选中的近邻点不尽相同。本文取这五种距离度量选中的近邻点交集,并统计它们的频次。如表 1 所示,公共近邻点(18,1)在每个距离度量中均被选中,故频次为 4。公共近邻点(18,3)、(19,1)和(17,1)在两个距离度量中选中,故频次为 2。最后对每个公共近邻点进行概率加权得到预测位置。结果表明,通过投票机制预测的位置与用户实际位置更接近,定位精度更高。

从表 1 中可以看出,影响定位精度的因素不仅有近邻



表 1 投票机制示例

距离度量	选中的近邻点(K=6)						预测位置	用户实际位置	定位误差/m
距离度量 1	(16,2)	(18,3)	(19,1)	(18,1)	(15,1)	(17,2)	(17.17,1.67)		1.22
距离度量 2	(18,5)	(18,1)	(20,5)	(19,5)	(16,3)	(19,1)	(18.33,3.33)		1.58
距离度量 3	(18,3)	(17,1)	(18,1)	(20,4)	(17,6)	(16,4)	(17.67,3.50)		1.89
距离度量 4	(18,6)	(17,1)	(18,1)	(17,2)	(19,3)	(20,6)	(18.17,3.17)	(18.39,1.75)	1.43
距离度量 5	(18,4)	(19,2)	(19,4)	(17,4)	(15,2)	(16,5)	(17.33,3.5)		2.04
提出的方法	[4×(18,1)+2×(18,3)+2×(19,1)+2×(17,1)]/						(18,1.4)		0.52
	(4+2+2+2)								

点数量 K, 还有距离度量的数量。因此本文将在下一章对此进行分析。

### 3 仿真分析

为了评估所提算法的定位性能, 本文基于 MAN2 数据集<sup>[21]</sup> (涵盖曼海姆大学(德国)校园办公楼 2 层的走廊)进行实验。如图 4 所示, 走廊区域的长和宽分别为 36 和 50 m, 参考点和测试点的数量分别为 130 和 46 个, 分别用蓝色和粉红色方块表示, 在采集到 RSS 后, 对 RSS 值进行了平均处理。处理之后的数据集中, 每个参考点和测试点均包含 10 个指纹。MAN2 数据集中共有 28 个 AP, 但其中参考点和测试点未接收到来自于部分 AP 的信号, 这是由于接收机接收灵敏度的限制所引起的。当信号强度小于 -100 dBm 时, 该信号将完全淹没在热噪声中, 接收机此时将无法正常工作。因此, 当接收到的信号强度值低于 -100 dBm 时, 数据集中均以 -100 dBm 来代替。

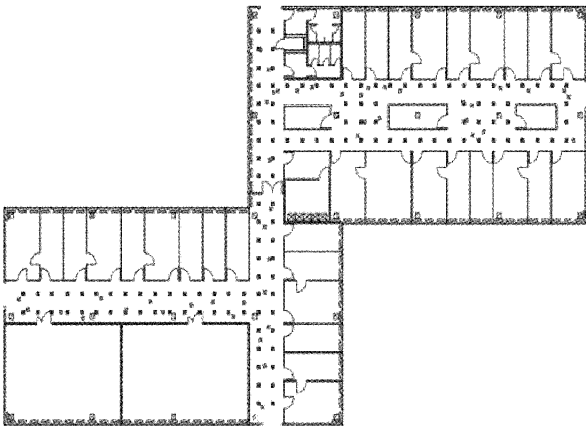


图 4 曼海姆大学(德国)校园办公楼 2 层走廊示意图<sup>[21]</sup>

#### 3.1 预处理对定位精度的影响

为了探究预处理方式对于定位精度的影响, 本文将每种预处理的定位结果和未使用预处理的定位结果进行对比, 如图 5 所示。

正如预期的那样, 使用非线性数据表示更适合于室内定位问题。图 5(a)展示了 3 种预处理方法和未使用预处理方法的累积分布函数图 (cumulative distribution function, CDF), 从图中可以看出, 当定位误差在 2 m 以内

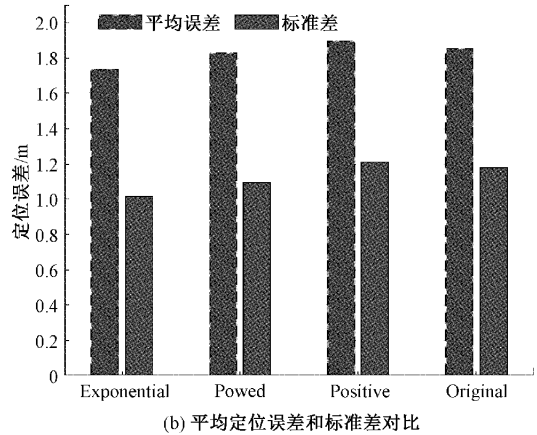
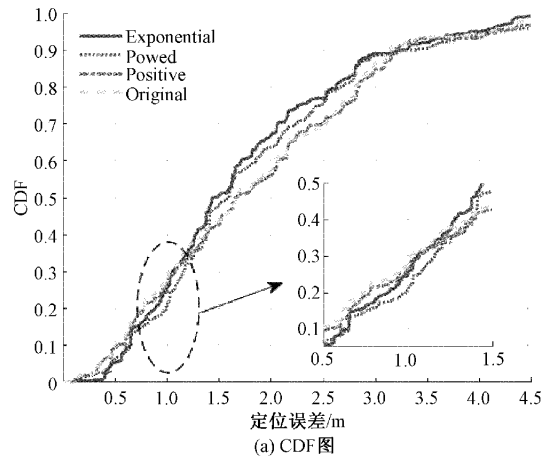


图 5 3 种预处理方法与未使用预处理的定位结果对比

时, Exponential 和 Powered 方法的概率分别为 67% 和 64%。而对于 Positive 方法而言, 它的概率略低于未使用预处理方法的概率。此外, 从图 5(b) 中可以看出, 这 3 种预处理方法和未使用预处理方法的平均定位误差和标准差分别为 Powered(1.826 0, 1.093 8)、Positive(1.895 0, 1.208 6)、Original(1.850 9, 1.175 9) 和 Exponential(1.735 1, 1.014 5)。Exponential 预处理方法均取得了较好的效果, 因此在下文中均用 Exponential 预处理方法。

#### 3.2 不同数量的距离度量对定位精度的影响

距离度量数量的多少对定位误差也有一定的影响。当距离度量数量较低时, 每个距离度量选中的近邻点交集

较少,甚至可能并没有交集,因此此时并不能较好地执行投票操作,将会导致定位精度降低。当距离度量数量较高时,一些异常近邻点被部分距离度量选中,导致这些异

常近邻点对最终位置有一定消极影响,也将导致定位误差增加。故本文分别从 8 个距离度量中,依次选取 1~8 个距离度量,分别计算它们的最低定位误差,如表 2 所示。

表 2 不同距离度量数量对定位误差的影响

数量	最低定位误差对应的距离度量	最低定位误差
1	Squared $\chi^2$	1.886 7
2	Sørensen + Euclidean $L_2$	1.777 2
3	Sørensen + Euclidean $L_2$ + Squared $\chi^2$	1.737 5
4	Sørensen + Squared $\chi^2$ + Jaccard + Topsøe	1.717 5
5	Sørensen + Squared $\chi^2$ + Jaccard + Topsøe + Euclidean $L_2$	1.692 7
6	Sørensen + Squared $\chi^2$ + Jaccard + Topsøe + Euclidean $L_2$ + City_Block $L_1$	1.707 8
7	Sørensen + Squared $\chi^2$ + Jaccard + Topsøe + Euclidean $L_2$ + City_Block $L_1$ + Fidelity	1.706 9
8	Sørensen + Squared $\chi^2$ + Jaccard + Topsøe + Euclidean $L_2$ + City_Block $L_1$ + Fidelity + Hellinger	3.491 1

从表 2 中可以看出,当距离度量选取 Sørensen、Squared $\chi^2$ 、Jaccard、Topsøe 和 Euclidean $L_2$  这五种时,平均定位误差最低,为 1.692 7 m。因此,在下文中,投票机制以这 5 种距离度量进行匹配。

3.3 K 值对定位精度的影响

在确定了预处理方式和投票机制中的距离度量后,通过匹配即可得到最终定位结果。但是,近邻点数量的选取也会对定位精度造成影响。近邻点数量过多的话,远离实际用户的参考点也有可能被选取进来;近邻点过少的话,有些对用户定位有贡献的近邻点则可能无法发挥作用。从图 6 可以看出,随着 K 值的增大,平均定位误差先减小后增大。当近邻点数量较小时,如 K=1 时,此时与最近邻法类似,即选中距离最小的近邻点作为目标位置。但是,由于测量误差的存在,距离最近的近邻点可能远离实际用户位置,从而导致较大的定位误差。当近邻点数量较大时,易选取到异常近邻点。由图可以看到,当 K=19 时,本文所提出的方法平均定位误差最低,为 1.638 3 m,因此下文中将 K 设置为 19。

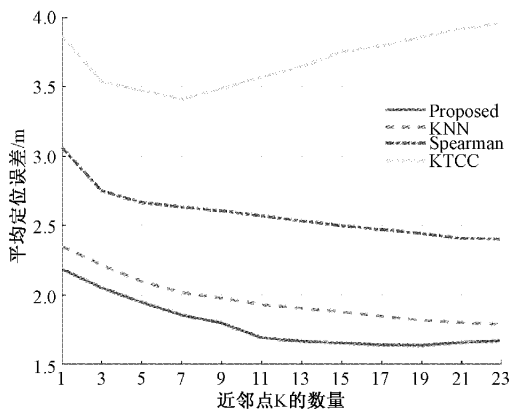


图 6 近邻点数量对定位精度的影响

3.4 不同方法的定位精度对比

图 7 展示了本文所提出的定位方法与现有方法的误差累积分布图。由图 7 可知,与其他方法相比,本文提出的定位方法性能最优,定位误差在 2 m 以内的概率为 70%,KNN 方法、Spearman 方法和 KTCC 方法分别为 62%、52%和 34%。这是因为 KNN 在进行匹配时易受环境影响,导致易选取到异常近邻点,最后进行加权得到的预测位置与用户的真实位置距离较远,因此定位误差较大。Spearman 和 KTCC 方法由于依赖 AP 的秩排序,没有考虑到 RSS 值,故导致了定位误差较大。

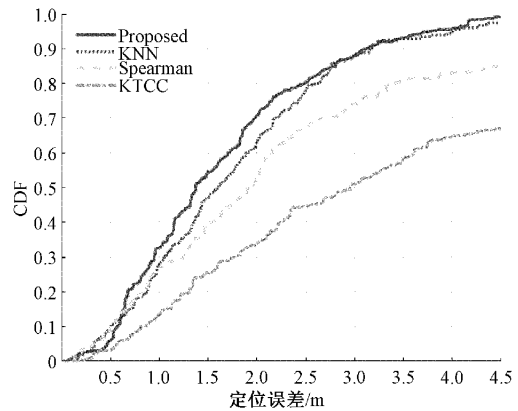


图 7 每种方法的 CDF 比较

为了更直观地对比每种方法的定位性能,图 8 比较了不同定位方法的平均误差。由图可以看出,本文提出的方法平均定位误差最低,为 1.638 3 m,而 KNN、Spearman 和 KTCC 方法的定位精度分别为 1.819 0、2.442 2 m 和 3.858 4 m。与 KNN、Spearman 和 KTCC 方法相比,本文所提出的方法的平均定位精度分别提升了 10%、33%和 58%。同时,为了与 MAN2 数据集的最优定位精度 (1.86 m) 对比,本文在图中 1.86 m 处设置基线,可直观地

看出提出的方法定位精度提高了 0.23 m,提升了 12%。

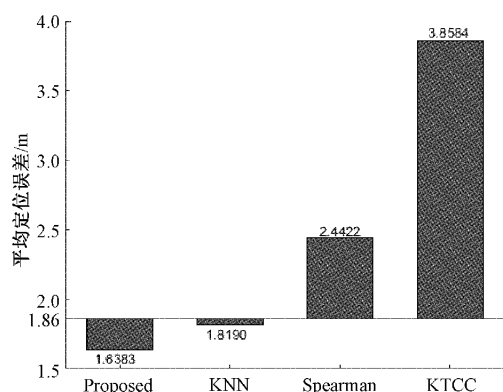


图 8 每种方法的平均误差比较

### 3.5 不同方法的定位时间对比

定位精度对于人们日常生活中导航和定位起着至关重要的作用,但是定位时间也不能忽视。为了对每种方法的定位效率进一步进行评估,本文探究了每种方法对单个用户定位的所需时间,如表 3 所示。

表 3 不同方法的定位时间对比

定位方法	定位时间/s
Proposed	0.147 1
KNN	0.003 2
Spearman	0.391 0
KTCC	0.265 1

由表 3 可知,所提算法的定位时间为 0.147 1 s,仅慢于 KNN 方法,而 Spearman 和 KTCC 方法所需定位时间相对较长。结合图 8 可得,所提算法无论在定位精度还是在定位时间上都表现出卓越的性能,可以满足实际生活场景中的定位需求。

## 4 结 论

本文提出了一种基于投票机制的室内 WiFi 指纹定位算法。考虑到 dBm 表示与功率之间的关系,首先对 RSS 测量数据进行预处理。同时基于投票机制对每种距离度量选中的近邻点取交集来确定公共近邻点,可以避免因单个距离度量的局限性而选取到异常近邻点,根据 5 种距离度量投票产生的公共近邻点以概率形式进行加权提高了定位精度。仿真结果表明,与 KNN、Spearman 和 KTCC 方法相比,本文提出的方法的定位精度均优于以上方法。此外,与 MAN2 数据集中最优定位精度相比,所提算法的定位精度提高了 12%。接下来的工作中,将在离线阶段添加指纹库的降噪处理以提高定位精度,并在其他公开数据集上进行验证。

### 参考文献

[1] SHIT R C, SHARMA S, PUTHAL D, et al.

Location of things (LoT): A review and taxonomy of sensors localization in IoT infrastructure[J]. IEEE Communications Surveys & Tutorials, 2018, 20(3): 2028-2061.

[2] STEINER L, MEINDL M, MARTY C, et al. Impact of GPS processing on the estimation of snow water equivalent using refracted GPS signals[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(1):123-135.

[3] 张斌飞,靳伍银. 基于 UWB 和 IMU 紧组合的室内定位导航算法[J]. 电子测量技术, 2022, 45(10): 67-73.

[4] 谢良波,李宇洋,杨小龙,等. 基于多频点相位距离/角度联合估计的 RFID 室内定位算法[J]. 仪器仪表学报, 2022, 43(7): 112-122.

[5] LI H, HUANG H, XU Y, et al. A fast and high-accuracy real-time visible light positioning system based on single LED lamp with a beacon[J]. IEEE Photonics Journal, 2020, 12(6): 1-12.

[6] FARAGHER R, HARLE R. Location fingerprinting with bluetooth low energy beacons[J]. IEEE Journal on Selected Areas in Communications, 2015, 33(11): 2418-2428.

[7] DUBEY A, SOOD P, SANTOS J, et al. An enhanced approach to imaging the indoor environment using WiFi RSSI measurements[J]. IEEE Transactions on Vehicular Technology, 2021, 70(9): 8415-8430.

[8] LI H, ZENG X, LI Y, et al. Convolutional neural networks based indoor WiFi localization with a novel kind of CSI images[J]. China Communications, 2019, 16(9): 250-260.

[9] XIE Y, WANG K, HUAN H. BPNN based indoor fingerprinting localization algorithm against environmental fluctuations[J]. IEEE Sensors Journal, 2022, 20(12): 12002-12016.

[10] 朱清山,王伟. 基于 RSSI 的指纹地图室内定位算法[J]. 国外电子测量技术, 2020, 39(10): 6-9.

[11] 郑自兴,姚金杰. 基于信号强度的指纹定位匹配算法研究[J]. 国外电子测量技术, 2020, 39(9): 31-35.

[12] 杨帆,赵东东. 基于 Android 平台的 WiFi 定位[J]. 电子测量技术, 2012, 35(9): 116-119.

[13] BAHL P, PADMANABHAN V N. RADAR: An in-building RF-based user location and tracking system[C]. Proceedings IEEE INFOCOM 2000. Conference on Computer Communications. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies, IEEE, 2000, 2: 775-784.

[14] HOU C, XIE Y, ZHANG Z. An improved convolutional neural network based indoor localization

- by using Jenks natural breaks algorithm[J]. China Communications, 2022, 19(4): 291-301.
- [15] XIE Y, WANG T, XING Z, et al. An improved indoor location algorithm based on backpropagation neural network[J]. Arabian Journal for Science and Engineering, 2022: 1-13.
- [16] TORRES-SOSPEDRA J, MONTOLIU R, TRILLES S, et al. Comprehensive analysis of distance and similarity measures for WiFi fingerprinting indoor positioning systems [J]. Expert Systems with Applications, 2015, 42(23): 9263-9278.
- [17] TORRES-SOSPEDRA J, RICHTER P, MOREIRA A, et al. A comprehensive and reproducible comparison of clustering and optimization rules in WiFi fingerprinting[J]. IEEE Transactions on Mobile Computing, 2022, 21(3): 769-782.
- [18] XIE Y, WANG Y, NALLANATHAN A, et al. An improved K-nearest-neighbor indoor localization method based on spearman distance[J]. IEEE Signal Processing Letters, 2016, 23(3): 351-355.
- [19] MA Z, WU B, POSLAD S. A WiFi RSSI ranking fingerprint positioning system and its application to indoor activities of daily living recognition [J]. International Journal of Distributed Sensor Networks, 2019, 15(4): 1550147719837916.
- [20] PENG X, CHEN R, YU K, et al. An improved weighted K-nearest neighbor algorithm for indoor localization[J]. Electronics, 2020, 9(12): 2117.
- [21] KING T, HAENSELMANN T, EFFELSBERG W. On-demand fingerprint selection for 802.11-based positioning systems[C]. International Symposium on a World of Wireless, Mobile and Multimedia Networks, IEEE, 2008:1-8.

### 作者简介

王开亮, 硕士研究生, 主要研究方向为室内定位和机器学习。

E-mail: 20201249326@nuist.edu.cn

谢亚琴, 博士, 副教授, 主要研究方向为室内定位、卫星导航、路径规划和人工智能。

E-mail: xyq@nuist.edu.cn

宦海(通信作者), 博士, 副教授, 主要研究方向为计算机视觉、自然语言处理和超分辨率重建。

E-mail: haihuan@nuist.edu.cn

周莉莉, 硕士研究生, 主要研究方向为室内定位。

E-mail: 20201249367@nuist.edu.cn