

DOI:10.19651/j.cnki.emt.2519025

基于 WiFi 的混合稳健 AP 定位技术研究*

贺卓悦 王伟 刘杰 马福生

(中北大学信息与通信工程学院 太原 030051)

摘要: 基于 WiFi 的指纹定位技术利用现有的无线接入点进行 WiFi 指纹定位时,存在 AP 质量不好掌控、指纹数据库冗余、实时定位计算量大等问题。为解决上述问题,提出一种混合稳健接入点选择算法。该算法首先在离线阶段通过综合信号稳定性指数分析 AP 的稳定性,然后利用互信息和相关系数分析 AP 的相似性,筛选出稳定、相似度低的 AP 构建新的轻量化指纹数据库;在线阶段利用对数距离路径损耗模型评估实时信号质量,选择高质量 AP 进行匹配定位。实验结果表明,与原始数据库相比,该算法有效筛选了冗余 AP,定位误差降低了 57.79%,同时定位精度在 1.5 m 以下的概率从 68.75% 提高到 93.75%。

关键词: AP 选择;指纹匹配;WiFi 室内定位

中图分类号: TN92 **文献标识码:** A **国家标准学科分类代码:** 510.5015

Research on hybrid robust AP positioning technology based on WiFi

He Zhuoyue Wang Wei Liu Jie Ma Fusheng

(School of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

Abstract: When WiFi fingerprint positioning technology utilizes existing access points for WiFi fingerprint positioning, there are problems such as uncontrollable AP quality, redundant fingerprint databases, and large real-time positioning calculation volume. To solve these problems, a Hybrid Robust Access Point Selection method is proposed. This algorithm first analyzes the stability of APs through comprehensive signal stability index in the offline stage, and then analyzes the similarity of APs using mutual information and correlation coefficient to screen out stable APs with low similarity to construct a new lightweight fingerprint database. In the online stage, the log-distance path loss model is used to evaluate the real-time signal quality and select high-quality APs for matching positioning. Experimental results show that compared with the original database, this algorithm effectively eliminates redundant APs, reduces positioning error by 57.79%, and increases the probability of positioning accuracy below 1.5 m from 68.75% to 93.75%.

Keywords: AP selection; fingerprint matching; WiFi indoor positioning

0 引言

在现代生活中,基于位置信息的服务已经成为人们日常生活必不可少的一部分^[1]。在室外环境中,全球定位系统(global positioning system, GPS)已经能够基本满足需求。在室内定位方面,由于环境复杂、多径效应以及障碍物遮挡等问题, GPS 无法满足室内定位需求^[2-4]。因此,随着无线通信技术和普及,基于无线局域网的室内定位方法被提出并发展迅速^[5-6]。其中,基于 WiFi 的室内定位技术由于其通过 WiFi 网络进行定位,覆盖范围广、不需要额外的定位设备、易实现、成本低、定位精度高等优点被

广泛应用^[7-9]。基于 WiFi 的室内定位技术中,基于位置指纹进行定位的方法使用 WiFi 信号强度进行定位^[10],不需要确定无线接入点的确切位置,最普遍且易于实现^[11]。该方法分为离线和在线两个阶段,离线阶段采集无线接入点(access point, AP)的接收信号强度指示值(received signal strength indication, RSSI)来建立位置指纹数据库;在线阶段利用采集到的数据数据库进行匹配来实现定位^[12-14]。

在现实的室内环境中,存在大量 AP,几乎每个位置都能收到多个 AP 信号,如果检测到的所有 AP 都参与定位,会大大增加计算复杂度。并且由于障碍物、多径效应、AP 分布不均匀等因素影响,采集到的 RSSI 值存在波动^[15-16],

收稿日期:2025-06-05

* 基金项目:山西省自然科学基金面上项目(202203021221117)资助

直接用采集的 AP 信号进行定位意味着 AP 质量不稳定,会影响定位效果。所以选择适合的 AP 是构建 WiFi 指纹数据库进行指纹定位的关键。

目前已有的 AP 选择方法包括文献[17]提出的最大均值算法,该算法主要思想为,将采集到的各个 AP 按照 RSSI 的平均大小排列,选择前 n 个 AP 组成数据库;文献[18]提出在该算法的基础上进行改进,除了按 RSSI 均值排列,还按方差排列,最后取均值排列的前 m 个和方差排列的前 n 个 AP 的交集;文献[19]通过稳定度选择 AP,利用不同 AP 的均值、方差和出现频率计算稳定度综合指标,保留稳定度高的 AP。他们都是在离线阶段选择 AP,这样能够有效筛选冗余 AP,减轻定位时的计算负担,但无法完全反映定位时 AP 的位置信息辨别能力。并且在线阶段环境如果发生变化,离线阶段筛选的 AP 可能不再是定位时最佳的 AP 组合。还有文献[20]在在线阶段将图像处理中的 RANSAC 算法引入 AP 选择;文献[21]提出计算互信息熵来进行在线 AP 选择;文献[22]使用 ACO 算法,利用 AP 划分区域再进行定位。这几种方法主要是在线选择 AP,能够实时调整选择的 AP,更好地适应在线定位的环境,但是对离线数据库没有进行筛选,匹配定位时计算复杂度较高,影响定位效率。综合离线在线的 AP 选择法有文献[23]提出利用标准差、皮尔逊相关系数和信息熵构建了综合 AP 选择策略,但是该方法在分析 AP 相似度时默认其均服从正态分布,而实际情况中的 AP 并不完全服从正态分布。

针对以上问题,在上述方法的基础上,提出混合稳健接入点选择算法(hybrid robust access point selection, HRAPS),该算法通过更加多维的角度,多方面分析信号,结合离线和在线两个阶段进行 AP 筛选。离线阶段,综合运用信号接收情况、信号质量情况、时间稳定性、空间稳定性,得到综合信号稳定性指数,通过分析各个 AP 的稳定性;然后计算 AP 之间的互信息和相关系数,分成不同类型来判断 AP 之间的相似性,筛选出信号质量好、信号稳定、相似性低的若干 AP 组成轻量化指纹数据库。在线阶段,对实时采集到的 AP,依据对数距离路径损耗模型计算 AP 信号的信号强度指标,按照指标从大到小排列,筛选出信号强度指标较大的 AP 参与在线匹配定位。该方法能够有效提高室内定位精度。

1 基于稳定性和相似性的 HRAPS 算法

1.1 离线阶段

离线阶段的 AP 选择主要从稳定性和相似性两个方面来进行筛选。由于室内空间复杂,WiFi 信号会受到影响,AP 信号的 RSSI 值存在波动,如图 1(a)所示。在同一位置采集的同一 AP 信号的 RSSI 值明显存在波动。并且部分 AP 接收信号强度过弱,还有一些 AP 只能在非常有限的区域被接收到,这些 AP 不适合用于指纹定位。

现实环境中存在大量 AP,不同 AP 之间的相似度不

同,如图 1(b)所示,两个相似度大的 AP 的 RSSI 值变化高度同步,当一个 AP 的信号强时,另一个 AP 的信号也强,当一个 AP 的信号弱时,另一个 AP 的信号也弱,具有很强的相似性。在指纹数据库同时包含相似性高的 AP 对会造成冗余,增加计算复杂度。通过分析 AP 的稳定性以及判断不同 AP 之间的相似度,保留稳定性高、相似度低的 AP。

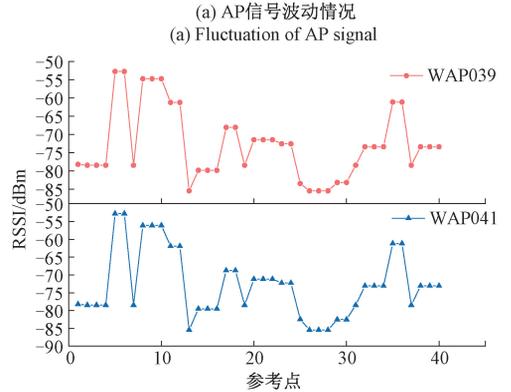
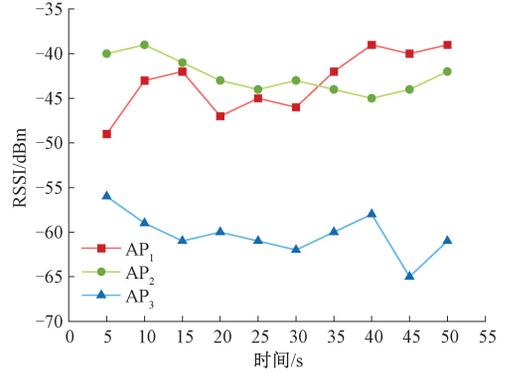


图 1 AP 信号波动及分布情况

Fig. 1 Fluctuation and distribution of AP signal

首先对稳定性进行分析,通过计算综合信号接收指数、综合信号质量指数、时间稳定性指数以及空间稳定指数得出信号稳定性指数和时间空间稳定性指数,进而得出综合信号稳定性指数,根据其选择受环境影响较小的稳定性更高的 AP。信号稳定性指数的计算公式如下:

$$AP_{S1} = aE_r + (1-a)Q_s \quad (1)$$

其中, E_r 表示信号综合接收指数,考虑距离因素对信号接收的影响; Q_s 表示综合信号质量指数,考虑信号质量的影响; a 为稳定性指数占比权重指标。

$$\begin{aligned} E_r &= (0.7R_{AP_i} + 0.3R_w) \times 100 \\ &= 0.7 \frac{N_{AP_i}}{S_{RP}} + 0.3 \frac{\sum (\omega_i \times I_{r_i})}{\sum \omega_i} \end{aligned} \quad (2)$$

$$\begin{aligned} Q_s &= 0.4Q_b + 0.3Q_d + 0.3Q_s \\ &= 0.4 \frac{N_{RSSI > RSSI_q}}{N_{AP_i}} + 0.3(1 - \frac{\sigma_h}{\mu_h}) + 0.3 \frac{N_{RSSI \in [Q_1, Q_3]}}{N_{AP_i}} \end{aligned} \quad (3)$$

其中, R_{AP_i} 表示基础信号接收比率, S_{RP} 表示所有的参考点个数, N_{AP_i} 表示 AP_i 能被所有参考点接收到的次数, R_w 表示加权接收比率, w_i 表示距离权重 ($w_i = 1/(d_i + 1)$), d_i 表示参考点到 AP_i 的距离)。 Q_b 表示信号质量比率, $N_{RSSI > RSSI_q}$ 表示 AP 的 RSSI 值大于阈值 $RSSI_q$ 的数量(当 RSSI 值大于 -70 dBm 时可以视该信号的强度良好、有较强的稳定性,所以在本文中 将 阈 值 设 为 -70 dBm); Q_d 表示信号强度分布均匀度, σ_h 表示信号强度直方图的标准差, μ_h 表示信号强度直方图的平均值; Q_s 表示信号稳定区间比例, $N_{RSSI} \in [Q_1, Q_3]$ 表示信号强度

$$S_s = 0.6 |SCI| + 0.4SCU = 0.6 \left| \frac{\sum [(D_{ij} - \bar{D}) (|RSSI_i - RSSI_j| - \bar{R})]}{\sqrt{[\sum (D_{ij} - \bar{D})^2]} \times \sqrt{[\sum (|RSSI_i - RSSI_j| - \bar{R})^2]}} \right| + 0.4(1 - \frac{\sigma_c}{100}) \quad (6)$$

其中, T_s 表示时间稳定性指数, σ_{rssi} 表示 RSSI 值的标准差, μ_{rssi} 表示 RSSI 值的平均值; SCI 表示空间一致性指数, D_{ij} 表示参考点之间的物理距离, \bar{D} 表示所有距离的平均值, \bar{R} 表示所有 $|RSSI_i - RSSI_j|$ 的平均值; SCU 表示空间覆盖均匀度, σ_c 表示覆盖区域内 RSSI 值的标准差。

根据稳定性指数和时间空间稳定性指数计算信号综合稳定性指数,综合信号稳定性指数的计算公式如下:

$$AP_s = cAP_{s1} + (1 - c)AP_{s2} \quad (7)$$

其中 c 为综合信号稳定性指数占比权重指标。基于信号稳定性,设置合适的权重 (a, b, c),考虑信号稳定性指数、时间空间稳定性指数和综合信号稳定性指数,保留合适的 AP 用于定位。在保留 AP 时,并不是 AP 数量越少越好,AP 数量太少会导致指纹特征丢失;也不是 AP 数量越多越好,AP 数量太多会增加不必要的计算量。所以要通过分析选择合适数量的 AP。

接下来分析信号稳定性筛选后的 AP 的相似度,AP 之间的相似度可以通过互信息和相关系数来衡量,它们结合起来即能判断信号间的线性关系,又能分辨非线性关系。互信息越小表示 AP 对的相似度越小,两个 AP 越不容易相互影响。计算任意两个 AP (AP_i 和 AP_j) 信号之间互信息和相关系数。互信息计算公式为:

$$I(AP_i; AP_j) = H(AP_i) + H(AP_j) - H(AP_i AP_j) \quad (8)$$

其中, $H(AP_i)$ 表示 AP_i 的信息熵, $H(AP_j)$ 表示 AP_j 的信息熵, $H(AP_i AP_j)$ 表示 AP_i 和 AP_j 的联合信息熵。

$$H(AP_i) = - \sum_{rssi_i} p(rssi_i) \log p(rssi_i) \quad (9)$$

$$H(AP_i AP_j) = - \sum_{rssi_i, rssi_j} p(rssi_i, rssi_j) \log p(rssi_i, rssi_j) \quad (10)$$

其中, $p(rssi_i)$ 表示 AP_i 的接收信号强度指示值为 $rssi_i$ 时的概率, $p(rssi_i, rssi_j)$ 表示 AP_i 和 AP_j 的接收信号强度指示值分别为 $rssi_i$ 和 $rssi_j$ 时,它们同时发生的

值在区间 $[Q_1, Q_3]$ 的 AP 数, $[Q_1, Q_3]$ 表示信号强度的四分位区间。

时间空间稳定性指数计算公式如式(4)所示。

$$AP_{s2} = bT_s + (1 - b)S_s \quad (4)$$

其中, T_s 表示时间稳定性指数,反映信号在时间维度上的波动程度; S_s 表示空间稳定指数,反映信号在空间区域内的稳定性; b 为时间空间稳定性指数占比权重指标。

$$T_s = 1 - \frac{\sigma_{rssi}}{|\mu_{rssi}|} \quad (5)$$

概率。

相关系数计算公式为:

$$R(AP_i, AP_j) = \frac{\sum (rssi_i - \overline{rssi_i})(rssi_j - \overline{rssi_j})}{\sqrt{\sum (rssi_i - \overline{rssi_i})^2} \sqrt{\sum (rssi_j - \overline{rssi_j})^2}} \quad (11)$$

其中, $rssi_i$ 和 $rssi_j$ 分别为 AP_i 和 AP_j 的接收信号强度值。

互信息是一种可以用来衡量两个随机变量之间相互依赖程度的指标,它量化了两个信号共享的信息量。当互信息值为零时,表示两个信号相互独立,没有共享信息;而当互信息值接近上限时,意味着一个信号中包含较多关于另一个信号的信息,两个信号高度相关。对采集的 AP 计算其互信息,如图 2(a)所示为 AP 对的互信息分布情况,可以看出 AP 对的互信息在 2~4 数量最多,在 4 以上的数量较少。设置接近互信息上限的 4 为阈值。如图 2(b)所示为随机两对互信息大于 4 的 AP 对的相似度分布图。可以看出当互信息大于 4 时,两个 AP 信号的变化模式高度相关,并且可能存在覆盖范围重叠,AP 之间的相似度过高,会造成信息冗余。

相关系数能用来衡量两个变量之间的线性相关程度,取值范围为 $[-1, 1]$,越接近 1,说明相关性越强。当相关系数大于 0.8 时,说明两个 AP 极强相关,有非常强的相关性,并且呈线性相关;当相关系数在 0.4~0.8 时,说明 AP 对较强相关或中度相关;当相关系数小于 0.4 时,说明 AP 对弱相关或非线性相关^[24]。

对于呈波动性、不均匀分布的 RSSI 信号,结合互信息和相关系数,分析 AP 对所包含的信息量以及它们之间的线性相关性,综合判断 AP 之间的相似性。

通过计算 AP 之间的互信息和相关系数,将其分为 6 种类型:

- 1) $I(AP_i; AP_j) > 4$ 且 $R(AP_i, AP_j) > 0.8$: 强线性相关且强相似度;
- 2) $I(AP_i; AP_j) > 4$ 且 $0.4 < R(AP_i, AP_j) \leq 0.8$:

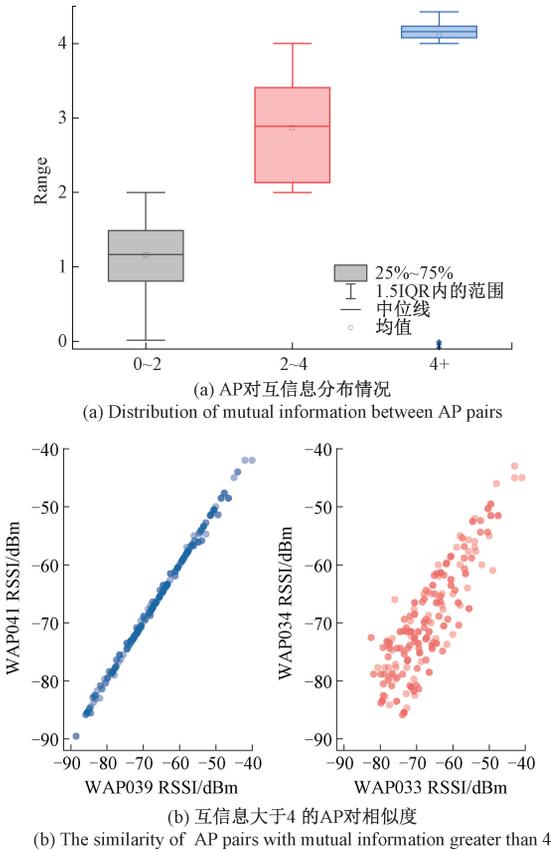


图 2 AP 对互信息情况

Fig. 2 AP for mutual information situation

中等线性相关；

- 3) $I(AP_i; AP_j) > 4$ 且 $R(AP_i, AP_j) \leq 0.4$: 非线性相关；
- 4) $I(AP_i; AP_j) < 4$ 且 $R(AP_i, AP_j) > 0.8$: 线性相关；
- 5) $I(AP_i; AP_j) < 4$ 且 $R(AP_i, AP_j) \leq 0.4$: 弱相关且弱相似度；
- 6) $I(AP_i; AP_j) < 4$ 且 $0.4 < R(AP_i, AP_j) \leq 0.8$: 其他类型。

在判断 AP 的相似性时, 首先将所有筛选保留下的 AP 两两组成一对, 计算所有 AP 对的互信息和相关系数, 然后按照上述分类方法将 AP 对分成 6 种类型, 剔除相似性高的 AP 对中的一个。最终保留的 AP 组成新的指纹数据库参与匹配定位。

1.2 在线阶段

经过离线阶段的选择, 指纹数据库的数量得到了缩减, 继续在在线阶段完成最后一步筛选。

研究表明, 室内路径损耗遵从对数距离路径损耗模型, 根据无线信号的传播规律, RSSI 随传播距离呈对数衰减。具体表达式如下:

$$RSSI(d) = RSSI(d_0) - 10n \lg\left(\frac{d}{d_0}\right) + x_\sigma \quad (12)$$

其中, $RSSI(d)$ 是距离为 d 处的接收信号强度, $RSSI(d_0)$ 是参考距离 d_0 处的接收信号强度; d_0 是参考距离, 通常为 1 m; d 是目标点的实际距离; n 是路径损耗指数, 通常是一个根据环境类型确定的常数, 一般在 2~4 之间; x_σ 是高斯噪声, 表示室内环境的复杂影响。

根据式(12)可以绘制出接收信号强度随距离变化曲线, 如图 3 所示。

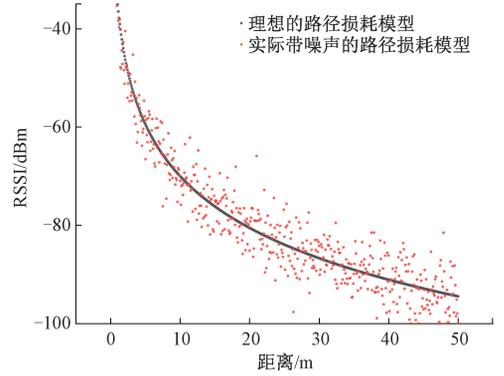


图 3 接收信号强度随距离变化曲线图

Fig. 3 Graph showing the variation of received signal strength with distance

从图 3 可以看出接收信号强度随着距离增加逐渐减小, 并且信号强度减小速度变慢。距离较小时减缓速度较快, 即曲线斜率较大, 此时信号强度较大; 随着距离越来越大, 即曲线斜率较小, 信号强度也较小。所以提出依据曲线斜率为判断依据的 AP 选择方法, 选择斜率大的 AP 进行定位。曲线斜率可以对式(12)通过一阶求导得出, 即:

$$RSSI(d)' = -\frac{10n}{d \ln(10)} \quad (13)$$

由式(13)可以得出, 斜率可由 d 判断。 d 由式(12)可以表示为:

$$d = d_0 \cdot 10^{\frac{RSSI(d) - RSSI(d_0) - x_\sigma}{-10n}} \quad (14)$$

假设 AP_i 和 AP_j 的 d 分别为 d_i, d_j , RSSI 值分别为 $RSSI(d)_i$ 和 $RSSI(d)_j$, 则:

$$d_i = d_0 \cdot 10^{\frac{RSSI(d)_i - RSSI(d_0) - x_\sigma}{-10n}} \quad (15)$$

$$d_j = d_0 \cdot 10^{\frac{RSSI(d)_j - RSSI(d_0) - x_\sigma}{-10n}} \quad (16)$$

$$\alpha = \frac{d_i}{d_j} = 10^{\frac{RSSI(d)_i - RSSI(d)_j}{-10n}} \quad (17)$$

最终可得出信号强度指标 α 。由式(17)计算 α , 如果 $\alpha > 1$, 则 $d_i > d_j$, 则 AP_i 的信号强度值大于 AP_j ; 反之, 如果 $\alpha < 1$, 则 $d_i < d_j$, 则 AP_i 的信号强度值小于 AP_j 。以此类推, 将在线阶段采集的 AP 按照从大到小的顺序进行排列, 选取高质量 AP 参与匹配定位。当参与实时定位的 AP 减少, 计算复杂度也随之减少。

同时为了增强在线匹配定位阶段的环境适应能力, 对信号传播特性、信号空间分布特征、以及空间障碍物影响

三类环境因素进行分析。

在线定位阶段,对测试集中的每个 AP 采集其当前 RSSI 值,并与所有 AP 中的最大 RSSI 进行比较,估算路径损耗作为信号传播参数。路径损耗的大小可以反映 AP 信号在当前环境中的衰减程度,间接体现环境对信号传播的影响^[25];对测试集所有有效 RSSI 值,计算其方差和均值作为空间分布特征参数。方差越大,表明当前环境下信号分布越分散,可能存在较大的空间复杂度或多径效应^[26];通过分析测试集 AP 信号序列中的突变点数(RSSI 突变大于均值加标准差的点数)作为障碍物参数,估计环境中可能存在障碍物,突变点越多,表明环境障碍或动态干扰越严重^[27]。

基于上述方法构建自适应在线 AP 选择系统。在在线定位阶段,依据信号强度指示 α 、信号传播特性、信号空间分布特征、以及空间障碍物影响自适应选择 AP 子集参与匹配定位。

自适应在线 AP 选择系统步骤如下:

1) 构建基础测试集,按照 α 对 AP 进行排列,计算不同保留比例时的平均定位误差,将平均定位误差最小时对应的 AP 保留比例下保留的 AP 作为参考 AP 子集;

2) 对参考 AP 子集分析信号传播特性、信号空间分布特征、以及空间障碍物影响,得到信号传播参数、空间分布参数和障碍物参数;

3) 依据信号传播参数、空间分布参数和障碍物参数调整下次 AP 保留比例,实现自适应动态调整。

通过自适应在线 AP 选择系统,动态选择 AP 参与匹配定位,这样既能降低在线定位的计算量,又能使 HRAPS 算法更加有效应对环境的变化,提高定位精度。

2 实验环境

为了验证 HRAPS 算法的性能,在中北大学科学楼 B 座 6 层中,选取人员走动随机、拐角处存在木门或玻璃门遮挡的走廊、电梯间和楼梯间进行实验,该区域长约 62 m,宽约 12 m,面积约 744 m²,所有墙体均为常见混凝土墙体,厚度为 200~300 mm。设置 42 个均匀分布的参考点,使用 OPPO K11 移动终端进行数据采集。实验环境的平面图如图 4 所示。

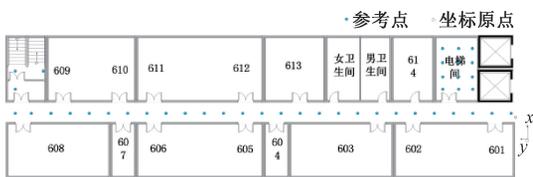


图 4 实验环境图

Fig. 4 Experimental environment diagram

在每个参考点使用采集软件 NetSpot 采集对应的 RSSI 值、MAC 地址、信道等信息,构建 WiFi 指纹数据库。

在现实环境中 WiFi 信号强度不断变化,采集时随机行人走动、走廊门随机开关,所以在每个参考点处保持设备高度为 110 cm,以 10 s 为采样间隔采集 20 次 RSSI 值,并求平均值作为该参考点的 RSSI 值。最大程度排除干扰并保留现实环境中的基本信号特征。

3 实验结果

3.1 离线阶段

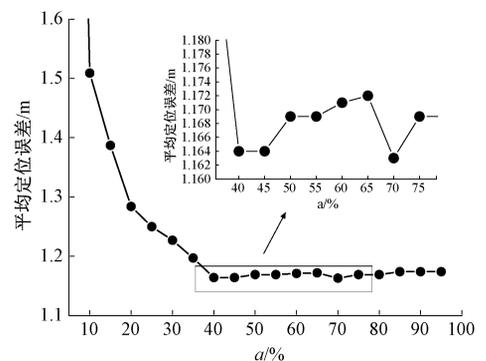
1) 稳定性筛选

进行 AP 选择时,首先筛选信号强度大于 -80 dBm 的信号,然后计算 AP 的综合信号接收指数、综合信号质量指数、时间稳定性指数以及空间稳定指数,接着按照合适的权重计算信号稳定性指数、时间空间稳定性指数以及综合信号稳定性指数。

为了选择合适的权重,通过控制变量法,控制其他因素保持一致分别对权重 a, b, c 进行测试。本文采用 WKNN 算法进行定位验证,经过实验,在其他条件相同时,WKNN 算法当 $K=3$ 时定位效果最好,在本文实验时默认取 $K=3$ 。

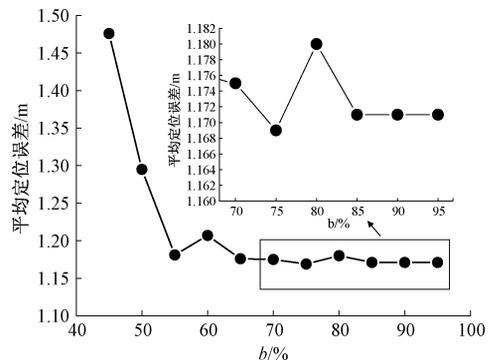
如图 5 所示为各个权重取不同数据时对应的平均定位误差分布图。

从图 5(a)、(b)、(c) 分别可以看出,当权重 a, b, c 分别为 70%、75% 和 90% 时平均定位误差最小,所以选择综合信号接收指数、时间稳定性指数和信号稳定性指数对应的



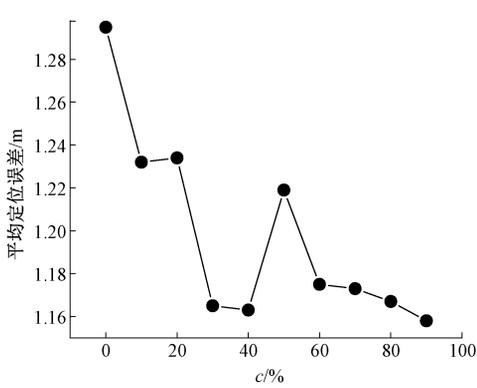
(a) 不同权重 a 的定位精度

(a) The positioning accuracy with different weights a



(b) 不同权重 b 的定位精度

(b) The positioning accuracy with different weights b



(c) 不同权重 c 的定位精度
(c) The positioning accuracy with different weights c

图 5 不同权重对应的定位精度

Fig. 5 The positioning accuracy corresponding to different weights

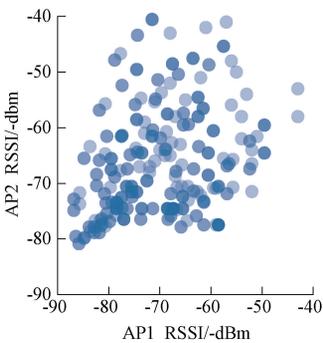
权重为 0.7、0.75、0.9。

对应的权重确定之后,计算 AP 的综合信号稳定性指数,设置合适的阈值筛选 AP。如图 6 所示为取不同综合信号稳定性指数阈值时对应的平均定位误差。从图 6 可以看出,采用综合信号稳定性指数阈值为 70%时定位精度最好,所以采用 70%的综合信号稳定性阈值进行 AP 选择。

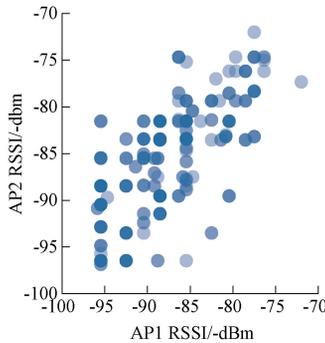
通过分析信号稳定性,采用综合信号稳定性指数阈值对 AP 进行筛选,既缩减了指纹数据库,去除了含有位置信息较少的信号,保留了信号质量相对稳定的信号。

2) 相似性筛选

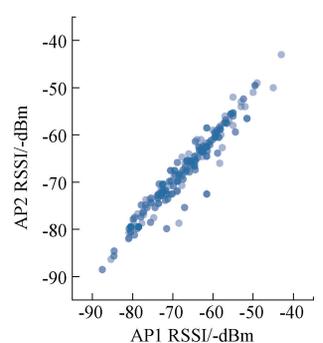
通过稳定性对 AP 筛选完成后,再使用互信息和相关



(a) 非线性相关
(a) Nonlinear correlation



(b) 其他类型
(b) Other types



(c) 强线性相关且强相似度
(c) Strong linear correlation and strong similarity

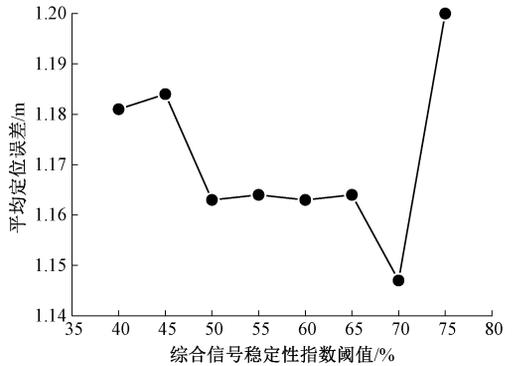
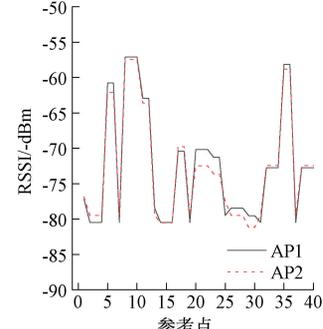
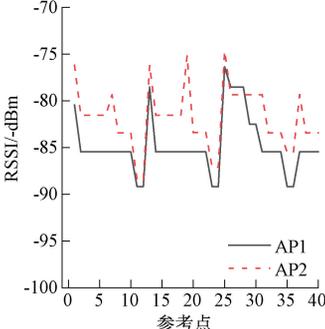
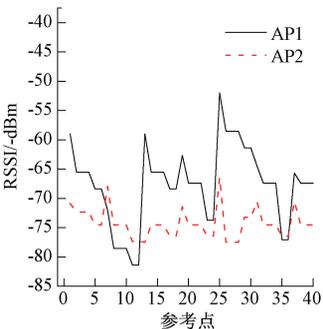


图 6 不同综合信号稳定性指数阈值的平均定位误差
Fig. 6 The average positioning error of different thresholds of comprehensive signal stability indices

系数分析信号相关性。如图 7 所示为 6 种相似性类型随机 AP 对的相似性分布情况。从图中可以看出,图 7(c)和(e)存在较为明显的线性相关情况,相似性较大,所以筛选强线性相关且强相似度和线性相关两种类型 AP 中的一个 AP。

如图 8 所示为 AP 对 6 种类型的分布情况。其中,弱相关且弱相似度占 54.7%,其他类型占 39.2%,线性相关占 3.7%,强线性相关且强相似度、中等线性相关和非线性相关占比较小,分别为 0.8%、0.9%和 0.8%。线性相关和强线性相关且强相似度两种类型共占 4.5%。

通过 HRAPS 算法在离线阶段的 AP 选择完成后,AP 数量从最开始的 163 减少到了 105 个,指纹数据库规模也随之减少了 35.58%。有效减轻了在线匹配定位阶段的计

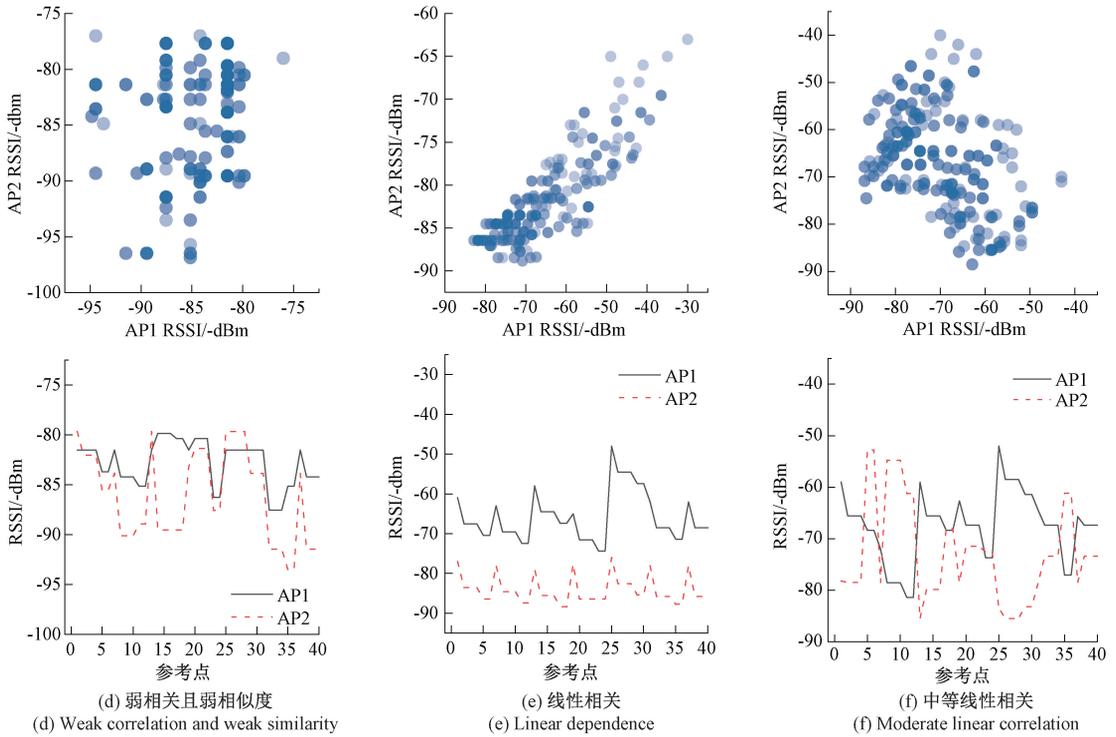


图 7 六种相似性类型随机 AP 对的相似性分布情况

Fig. 7 The distribution of similarity among six types of random AP pairs of similarity types

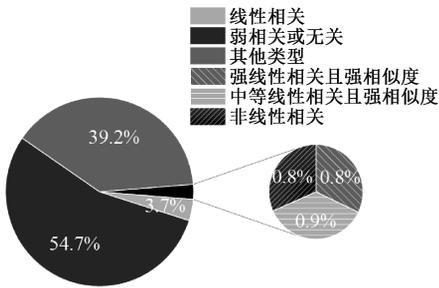


图 8 AP 对不同类型分布情况

Fig. 8 AP for different types of distribution scenarios

算负担。

3.2 在线阶段

在理想的环境中,用于定位的 AP 数量越多,定位效果越好,定位越准确。然而在现实环境中,AP 数量越多,定位精度可能会下降,同时定位复杂度提高。如图 9 所示为在本文实验环境下,构建基础测试集利用信号强度指标 α 从大到小对在线阶段的 AP 进行排列,取不同比例 AP 时对应的平均定位误差。从图 9 可以看出,当保留前 35% 的 AP 时平均定位误差最小,所以选择保留前 35% 的 AP 作为参考 AP 子集。利用该子集计算信号传播参数、空间分布参数和障碍物参数参与在线 AP 自适应调节,根据不同的测试集调整保留合适的 AP 数参与匹配定位。

如图 10 所示为不同测试集使用自适应在线 AP 选择系统后保留的 AP 数以及对应的平均定位误差,可以看出

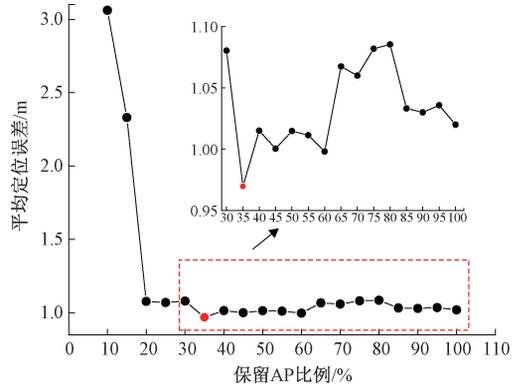


图 9 保留不同比例 AP 对应的定位精度

Fig. 9 Maintain the positioning accuracy corresponding to different proportions of AP

自适应在线 AP 选择系统可以根据不同的测试集保留不同数量的 AP 参与定位,且能获得较好的定位精度。

构建 3 个不同的新测试集对自适应在线 AP 选择系统的普适性进行验证。

测试集 A:与原始测试集间隔一个月采集构建的测试数据集。

测试集 B:仅在原始测试集全部空间中的楼梯间采集数据构建的测试数据集。该测试集较原始测试集所包含的 AP 数减少,位置信息较少,环境也有所不同;

测试集 C:将原始测试集中的 AP 数随机减少 30% 构建的测试数据集。该测试集与原始测试集 AP 发生变化。

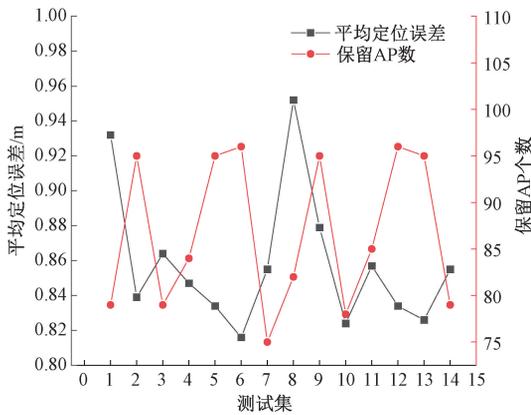
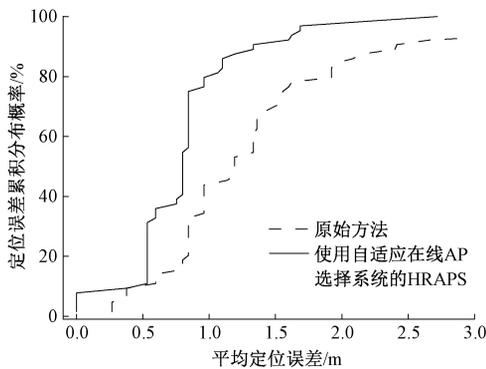


图 10 自适应在线 AP 选择系统效果

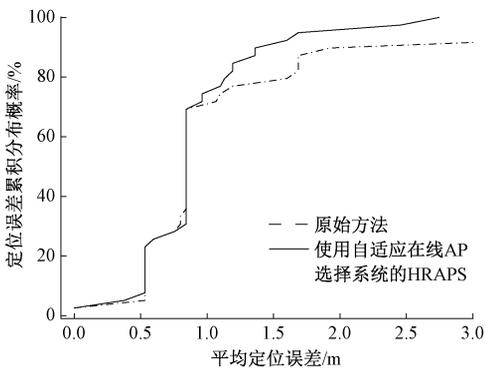
Fig. 10 Effect of the adaptive online AP selection system

如图 11 所示为测试集 A、B、C 使用包含自适应在线 AP 选择系统的 HRAPS 算法进行匹配定位与原始方法直接定位的 CDF 对比图。可以看出,自适应在线 AP 选择系统在不同环境、不同 AP 下的定位精度依然表现良好,能比较有效地适应环境变化。

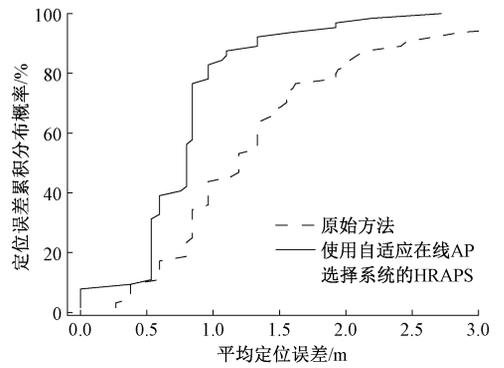
如表 1 所示为测试集 A、B、C 使用包含自适应在线 AP 选择系统的 HRAPS 算法参与匹配定位前后参与定位的 AP 数、匹配定位耗时以及平均定位误差。可以看出自适应在线 AP 选择系统能够有效适应环境的时间变化、空间变化和 AP 变化,HRAPS 算法能有效减少计算复杂度,



(a) 测试集 A
(a) Test set A



(b) 测试集 B
(b) Test set B



(c) 测试集 C
(c) Test set C

图 11 测试集 A、B、C 使用自适应在线 AP 选择系统的 HRAPS 算法与原始方法的 CDF 对比图

Fig. 11 Shows the CDF comparison between the

HRAPS algorithm of the adaptive online AP selection system and the original method for test sets A, B, and C

缩减定位时间,提高定位精度。

表 1 测试集 A、B、C 使用 HRAPS 前后的定位效果

Table 1 Shows the positioning effects of test sets A, B, and C before and after using HRAPS

测试集	使用 HRAPS		在线阶段		HRAPS 前后	
	前后参与在线		匹配耗时/		平均定位	
	前	后	前	后	前	后
A	163	89	900	854	1.952	0.824
B	78	26	862	223	1.645	0.941
C	114	62	820	608	1.957	0.813

3.3 算法比较

如表 2 所示为使用 WKNN 算法直接进行定位、使用文献[17]~文献[23]中的 AP 选择算法以及本文 HRAPS 算法进行定位的平均定位误差。可以看出 HRAPS 算法的平均定位误差最小,明显优于其他算法。

表 2 不同方法的平均定位误差

Table 2 The average positioning errors of different methods

定位算法	平均定位误差/m
WKNN	1.952
文献[17]	1.905
文献[18]	1.515
文献[19]	1.444
文献[20]	2.270
文献[21]	1.962
文献[22]	1.206
文献[23]	1.285
HRAPS(本文算法)	0.824

如图 12 所示为原始数据直接使用 WKNN 算法进行定位和使用本文 HRAPS 算法后再定位以及使用其他 AP 选择算法后进行定位的定位误差累积分布概率对比图。从图 12 可以看出,使用本文 HRAPS 算法在定位误差概率分布上优于直接使用 WKNN 进行定位以及其他算法。

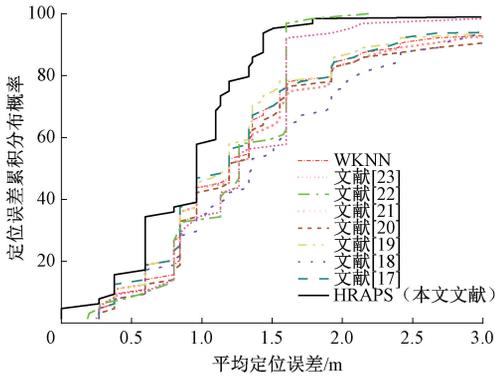


图 12 定位误差累积分布概率对比图

Fig. 12 Cumulative distribution probability chart of positioning error

如表 3 所示为直接使用 WKNN、使用现有其他 AP 选择方法以及使用本文 HRAPS 算法进行定位后的固定精度下的误差累积概率情况。从表 3 可以看出,本文方法定位误差在 0.5、1、1.5 m 内的累积概率分别为 15.62%、57.81%、93.75%。较直接使用 WKNN 进行定位的定位精度和其他 AP 选择法的定位精度有明显提高。

表 3 不同算法误差范围比例

Table 3 Proportion of error ranges of different algorithms

定位算法	误差在	误差在	误差在
	0.5 m 内/%	1 m 内/%	1.5 m 内/%
WKNN	9.38	43.75	68.75
文献[17]	12.50	46.88	68.75
文献[18]	7.81	34.38	53.12
文献[19]	10.94	45.31	73.44
文献[20]	7.81	42.19	65.62
文献[21]	10.94	43.75	62.50
文献[22]	7.81	32.81	59.38
文献[23]	9.38	34.38	56.25
HRAPS(本文算法)	15.62	57.81	93.75

4 结 论

本文为提高定位效率,针对 WiFi 指纹定位中指纹数据库冗余、AP 数量较多无法掌控所有 AP 质量、在线定位匹配时计算量大的问题,提出了一种结合离线阶段 AP 选择和在线阶段 AP 选择的综合算法——HRAPS 算法。离线阶段通过综合信号稳定性指数和互信息以及相关系数评估 AP 的稳定性、相似性,筛选出稳定、相似度低的 AP

子集构建新的轻量化指纹数据库;在线阶段通过信号强度指标评估实时信号质量,筛选出高质量 AP 子集,降低了计算量。实验结果表明,相较于包含环境中所有 AP 的指纹数据库,该方法减少了 35.58% 的规模,有效筛选了冗余 AP,降低了在线定位的计算负担。并且使用 HRAPS 算法的平均定位误差较原始指纹数据库提高了 57.79%,同时定位精度在 1.5 m 以下的概率为 93.75%,较原始指纹数据库提高了 26.67%。

根据稳定性和相似性筛选 AP 时,目前采用了固定阈值,怎样能够根据指纹数据库特征自适应地调整阈值,根据环境因素进行阈值优化从而达到最好的定位效果,未来可以在这方面继续进行研究。

参考文献

- [1] DENG SH H, ZHANG W J, XU L, et al. RRIFLoc: Radio robust image fingerprint indoor localization algorithm based on deep residual networks[J]. IEEE Sensors Journal, 2023, 23(3): 3233-3242.
- [2] ZHANG L, ZHANG M, XU J G, et al. Cluster-based JRPCA algorithm for Wi-Fi fingerprint localization[J]. Electronics, 2023, 12(1): 153.
- [3] 郑安琪, 秦宁宇. 空间指纹测量特征双精简下的室内定位算法[J]. 仪器仪表学报, 2023, 44(10): 80-89. ZHENG AN Q, QIN N N. Indoor positioning algorithm under double simplification of spatial fingerprint measurement features[J]. Chinese Journal of Scientific Instrument, 2023, 44(10): 80-89.
- [4] FEI R, GUO Y F, LI J H, et al. An improved BPNN method based on probability density for indoor location[J]. IEICE Trans on Information and Systems, 2023, 106: 773-785.
- [5] ZHU Y J, LUO X N, GUAN S W, et al. Indoor positioning method based on WiFi/bluetooth and PDR fusion positioning [C]. 2021 13th International Conference on Advanced Computational Intelligence (ICACI). IEEE, 2021: 233-238.
- [6] 刘伟, 王智豪, 李卓, 等. 基于 cGAN-SAE 的室内定位指纹生成方法[J]. 电子测量技术, 2024, 47(14): 57-63. LIU W, WANG ZH H, LI ZH, et al. Indoor positioning fingerprint generation method based on cGAN-SAE[J]. Electronic Measurement Technology, 2024, 47(14): 57-63.
- [7] 马杰, 王健, 李智. 基于改进 PDR 算法的室内定位方法研究[J]. 电子测量与仪器学报, 2024, 38(12): 211-217. MA J, WANG J, LI ZH. Research on indoor positioning method based on the improved PDR algorithm[J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(12): 211-217.
- [8] TAN S, REN Y L, YANG J, et al. Commodity Wi-Fi sensing in ten years: Status, challenges, and opportunities[J]. IEEE Internet of Things Journal,

- 2022, 9(18): 17832-17843.
- [9] 陈静, 张晓龙. 基于 MVO-SVR 的室内指纹定位算法[J]. 电子测量与仪器学报, 2024, 38(9): 45-53.
CHEN J, ZHANG X L. Indoor fingerprint positioning algorithm based on MVO-SVR [J]. Journal of Electronic Measurement and Instrumentation, 2024, 38(9): 45-53.
- [10] GUFRA D, PARICHA S. FedHIL: Heterogeneity resilient federated learning for robust indoor localization with mobile devices [J]. ACM Trans Embedded Computing Systems, 2023, 22(5s):125.
- [11] 蒲巧林, 陈有坤, 周牧, 等. 基于改进对比学习和并行融合神经网络的室内 WiFi 定位算法[J]. 仪器仪表学报, 2024, 45(1): 101-110.
PU Q L, CHEN Y K, ZHOU M, et al. Indoor WiFi positioning algorithm based on improved contrastive learning and parallel fusion neural network[J]. Chinese Journal of Scientific Instrument, 2024, 45 (1): 101-110.
- [12] 毕杨, 张杨梅, 刘坤, 等. 基于改进粒子群优化算法的机舱室内指纹定位研究[J]. 航空科学技术, 2024, 35(12): 54-59.
BI Y, ZHANG Y M, LIU K, et al. Research on indoor fingerprint positioning of the fuselage based on the improved particle swarm optimization algorithm [J]. Aviation Science and Technology, 2024, 35 (12): 54-59.
- [13] WEI ZH, CHEN J L, TANG H, et al. RSSI-based location fingerprint method for RFID indoor positioning: A review[J]. Nondestructive Testing and Evaluation, 39(1): 3-31.
- [14] 苏奎, 于曦, 芦思宇. 基于 RSS 指纹的室内定位方法[J]. 智能计算机与应用, 2024, 14(10): 158-163.
SU K, YU X, LU S Y. Indoor positioning method based on RSS fingerprint[J]. Intelligent Computers and Applications, 2024, 14(10): 158-163.
- [15] AYINL S L, AZIZ A A, DRIEBERG M. SALLoc: An accurate target localization in WiFi-enabled indoor environments via SAE-ALSTM [J]. IEEE Access, 2024, 12: 19694-19710.
- [16] 金宏平, 戴玉帆. 基于递阶 ANFIS 树的室内定位算法研究[J]. 国外电子测量技术, 2024, 43(6): 87-93.
JIN H P, DAI Y F. Research on indoor positioning algorithm based on hierarchical ANFIS tree [J]. Foreign Electronic Measurement Technology, 2024, 43(6): 87-93.
- [17] YOUSSEF M A, AGRAWALA A, SHANKAR A U. WLAN location determination via clustering and probability distributions[C]. First IEEE International Conference on Pervasive Computing and Communications. IEEE, 2003:143-150.
- [18] 李意顺. 基于机器学习理论的室内定位算法的改进与研究[D]. 北京: 北京交通大学, 2018.
LI Y SH. Improvement and research on indoor positioning algorithm based on machine learning theory [D]. Beijing: Beijing Jiaotong University, 2018.
- [19] 魏军, 罗恒, 倪启东, 等. 基于稳定 AP 选择的动态室内定位方法[J]. 微电子学与计算机, 2024, 41(1): 37-44.
WEI J, LUO H, NI Q D, et al. Dynamic indoor positioning method based on stable AP selection[J]. Microelectronics and Computers, 2024, 41(1): 37-44.
- [20] HOU Y Y, SUM G Y, FAN B W. The indoor wireless location technology research based on WiFi[C]. 2014 10th International Conference on Natural Computation (ICNC). IEEE, 2014: 1044-1049.
- [21] ZOU H, LUO Y W, LU X X, et al. A mutual information based online access point selection strategy for WiFi indoor localization [C]. 2015 IEEE International Conference on Automation Science and Engineering(CASE). IEEE, 2015:180-185.
- [22] ROY P, CHOWDHURY C. A region-wise indoor localization system based on unsupervised learning and ant colony optimization technique [J]. Applied Soft Computing, 2024, 157:111509.
- [23] 冷泽富, 邓中亮, 杨福兴, 等. 一种室内 WLAN 指纹定位中的 AP 选择策略[C]. 中国卫星导航系统管理办公室学术交流中心第九届中国卫星导航学术年会论文集. 北京: 北京邮电大学, 2018: 6.
LENG Z F, DENG ZH L, YANG F X, et al. An AP selection strategy in indoor WLAN fingerprint positioning [C]. 9th China Satellite Navigation Conference. Beijing: Beijing University of Posts and Telecommunications, 2018: 6.
- [24] EDWIN V D H, ZHAN ZH ZH. Myths about linear and monotonic associations: pearson's r , spearman's ρ , and kendall's τ [J]. The American Statistician, 2022, 76(1): 44-52.
- [25] 杨艳芳, 王伟, 王召巴. 基于 WSN 室内定位的路径损耗模型参数算法研究[J]. 电子测量技术, 2021, 44(13): 54-58.
YANG Y F, WANG W, WANG ZH B. Research on the parameter algorithm of path loss model based on indoor positioning of WSN [J]. Electronic Measurement Technology, 2021, 44(13): 54-58.
- [26] LENG Y H, HUANG F H, TAN W J. A Wi-Fi fingerprint positioning method based on RLWKNN[J]. IEEE Sensors Journal, 2025, 25(1): 1706-1715.
- [27] LIU F, LIU J, YIN Y Q, et al. Survey on WiFi-based indoor positioning techniques[J]. IET Communications, 2020, 14(9): 1372-1383.

作者简介

贺卓悦, 硕士研究生, 主要研究方向为室内应急定位、深度学习。

E-mail: hezhuoyue2001@163.com

王伟(通信作者), 博士, 副教授, 主要研究方向为室内定位、气味识别。

E-mail: 41695559@qq.com