

# 基于改进 WKNN 的 CSI 被动室内 指纹定位方法

邵小强<sup>1,2</sup>, 马 博<sup>1,2</sup>, 韩泽辉<sup>1,2</sup>, 杨永德<sup>1,2</sup>, 原译文<sup>1,2</sup>, 李 鑫<sup>1,2</sup>

(1. 西安科技大学 电气与控制工程学院, 西安 710054; 2. 西安科技大学 西安市电气设备状态检测与供电安全重点实验室, 西安 710054)

**摘 要:** 针对幅值和相位构造包含干扰过多导致定位精度低的问题, 提出了一种基于改进加权  $K$  最近邻算法的信道状态信息被动室内定位方法。离线阶段, 采用隔离森林法, 改进阈值的小波域去噪和线性变换法对采集到的信道状态信息进行预处理, 将处理后的幅相信息共同作为指纹数据, 构造与参考点位置信息相关的稳定指纹数据库。在线阶段, 提出改进的加权  $K$  近邻算法, 对估计坐标进行重复匹配, 该算法在一次匹配中得到位置坐标后, 求该位置坐标在  $K$  个近邻点间的欧氏距离, 并使用高斯变换对  $K$  个距离值进行权重计算, 完成人员的定位。分别在教室和大厅进行实验模拟测试, 实验结果表明: 采用本文算法约 81% 的测试位置误差控制在 1 m 以内, 可以有效提高定位精度。

**关键词:** 室内定位; 信道状态信息; 被动定位; 改进阈值的小波域去噪; 改进的加权  $K$  近邻算法; 高斯变换

中图分类号: TN92 文献标志码: A 文章编号: 1671-5497(2025)07-2444-11

DOI: 10.13229/j.cnki.jdxbgxb.20231101

## CSI Passive indoor fingerprint positioning method based on improved weighted $K$ -nearest neighbor algorithm

SHAO Xiao-qiang<sup>1,2</sup>, MA Bo<sup>1,2</sup>, HAN Ze-hui<sup>1,2</sup>, YANG Yong-de<sup>1,2</sup>, YUAN Ze-wen<sup>1,2</sup>, LI Xin<sup>1,2</sup>

(1. College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, China;  
2. Xi'an Key Laboratory of Electrical Equipment Condition Monitoring and Power Supply Security, Xi'an University of Science and Technology, Xi'an 710054, China)

**Abstract:** A passive indoor positioning method based on improved weighted  $K$ -nearest neighbor algorithm is proposed to address the problem of low positioning accuracy caused by excessive interference in amplitude and phase construction. In the offline stage, the isolation forest method is adopted, and the wavelet domain denoising and linear transformation method with improved threshold are used to preprocess the collected channel state information. The processed amplitude and phase information is used together as fingerprint data to construct a stable fingerprint database related to the reference point position information. In the online stage, an improved weighted  $K$ -nearest neighbor algorithm is proposed to repeatedly match

收稿日期: 2023-10-15.

基金项目: 国家自然科学基金项目(52174198).

作者简介: 邵小强(1976-), 男, 副教授, 博士. 研究方向: 室内人员定位. E-mail: shaoxq@xust.edu.cn

the estimated coordinates. After obtaining the position coordinates in a single match, the algorithm calculates the Euclidean distance of the position coordinates between  $K$ -nearest neighbor points, and uses Gaussian transformation to calculate the weight of the  $K$  distance values, completing personnel positioning. Experimental simulation tests were conducted in classrooms and halls, and it was found that approximately 81% of the proposed algorithm's testing position error was controlled within 1 meter, which can effectively improve positioning accuracy.

**Key words:** indoor positioning; channel state information; passive positioning; improved threshold denoising in wavelet domain; improved weighted  $K$ -nearest neighbor algorithm; Gaussian transformation

## 0 引言

在室内环境中,定位服务<sup>[1]</sup>有许多要求,如大型商场、医院、养老院、地下停车场等。全球卫星导航系统(Global navigation satellite system, GNSS)由于受到密集障碍物的阻塞或破坏,无法在建筑物内提供高精度的导航定位<sup>[2]</sup>。因此,人们开发了许多的室内定位技术,如无线局域网(Wireless fidelity, Wi-Fi)<sup>[3]</sup>、行人航迹推算(Pedestrian dead reckoning, PDR)<sup>[4]</sup>、超宽带(Ultra wide band, UWB)<sup>[5]</sup>、可见光(Visible light, VL)<sup>[6]</sup>、蓝牙(Bluetooth, BT)<sup>[7]</sup>等。其中,Wi-Fi因其成本低、部署简单、无处不在而备受关注。

Wi-Fi定位的方法分为两类:基于测距和基于指纹。基于测距法会受到复杂的多径效应和噪声干扰,很难建立准确的测距模型,而基于指纹的定位方法不需要数学模型,受多径影响较小,可以实现良好的室内定位结果。基于Wi-Fi的室内指纹定位系统是收集接收信号强度(Received signal strength, RSS)或信道状态信息(Channel state information, CSI)作为构建指纹数据库的特征。由于RSS的鲁棒性较差,区分目标位置的能力较弱,所以RSS不是理想的定位功能。因此,准确的室内定位需要更稳定、更细粒度的信息。信道状态信息<sup>[8,9]</sup>是Wi-Fi传输物理层中的细粒度信息,它在子载波的维度上清楚地描述了信号从接入点到参考点的功率损耗和时延。CSI在一定程度上刻画了多径传输,并添加了每个子载波的幅度和相位信息。Yang等<sup>[10]</sup>开发了一种新型高性能室内无设备被动定位方法,该方法可以通过指纹识别实现对单个TX-RX链路的细粒度定位,并利用CSI幅值作为指纹数据,使用概率匹配算法进行位置定位。但是,由于该方法仅使用幅值信息作为指纹,所以定位精度只能达到

1.8~2 m。Wang等<sup>[11]</sup>提出了一种双模态深度学习的室内定位方法,该方法使用平均幅值作为指纹数据,通过将3种典型室内环境作对比,并与3种基准方案进行了比较,平均定位误差为1.55 m。Rao等<sup>[12]</sup>提出了一种具有加权可转移判别降维方法的新型无设备无源无线定位系统,该系统在具有代表性的室内环境中能够有效提高定位精度,同时节省大量的校准成本,但是他们对幅值未进行预处理导致实验定位误差比较大。

针对以上幅值和相位构造指纹包含干扰过多导致定位精度低的问题,本文提出了一种改进加权 $K$ 最近邻算法的CSI被动室内定位系统。系统首先通过硬件设备采集CSI,通过隔离森林消除异常CSI,然后对采集到的数据进行预处理,预处理分为CSI的幅值处理和相位处理,幅值用改进阈值小波域去噪,相位采用线性变换。离线阶段,提取每个参考点的特征信息后,幅值和相位同时作为指纹信息存储在指纹数据库中。在线阶段,对采集到的实时数据进行预处理,使用IWKNN算法对估计的位置信息进行重复匹配,最终完成待测点的位置估计以得到最终的定位结果。

## 1 准备工作

### 1.1 信道状态信息

正交频分复用<sup>[13]</sup>是一种通信技术,采用频率选择性和ISL. IN802. 11n标准,将Chan-Nel划分为多个子载波。它可以建模如下:

$$Y = HX + N \quad (1)$$

式中: $Y$ 、 $X$ 分别为接收设备和发射设备的信号矢量; $H$ 为信道矩阵; $N$ 为高斯白噪声。

信道矩阵可由式(2)估计得出:

$$H = \frac{Y}{X} \quad (2)$$

在正交频分复用系统中,信号可以通过多个

子载波进行传输,因此,信道增益矩阵  $H$  还可以表示为:

$$H = [H_1, H_2, \dots, H_n] \quad (3)$$

$$H_i = |H_i| e^{j \sin(\angle H_i)} \quad (4)$$

式中:  $H_i$  为第  $i$  个子载波的信道状态信息;  $|H_i|$  和  $\angle H_i$  分别为第  $i$  个子载波上的 CSI 幅值和相位。

IEEE 802.11n<sup>[14]</sup> 标准中的多输入多输出 (Multi-ple input multiple output, MIMO) 系统支持多输入天线和多发射天线,该系统中使用信道状态矩阵表示信道状态。信道状态矩阵  $H$  可以表示为:

$$H = \begin{pmatrix} h^{11} & h^{12} & \dots & h^{1q} \\ h^{21} & h^{22} & \dots & h^{2q} \\ \vdots & \vdots & & \vdots \\ h^{p1} & h^{p2} & \dots & h^{pq} \end{pmatrix} \quad (5)$$

式中:  $h^{pq}$  为由发射和接收天线之间通信链路的信道状态信息矩阵;矩阵  $H$  的维度为  $t \times r \times s$ ,其中  $t$ 、 $r$  和  $s$  分别为发射天线数、接收天线数及子载波数。

本文采集数据的硬件包括采集端和发射端,发射端为两根天线的路由器,数据采集端为装有 Intel15300 无线网卡且配备 3 根接收天线的联想 X220 计算机,运行系统为 Ubuntu 14.04,装有 CSITool 工具。从 CSI 数据中采集的子载波数为 60,如图 1(a) 所示。

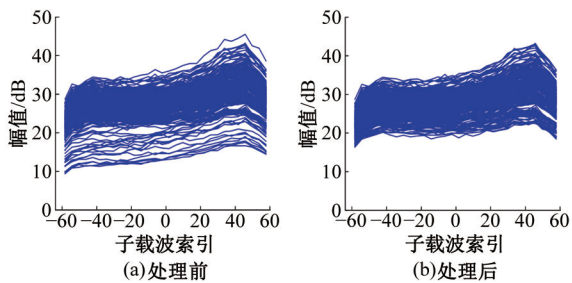


图 1 隔离森林处理前后图

Fig. 1 Isolated forest before and after treatment

### 1.2 异常 CSI 去除

在实验中,发现固定位置的 CSI 数据存在异常样本(见图 1(a)),造成这些不规则的异常样本的原因有很多。传统的异常检测方法都是基于正态分布的,无法消除图中所示的异常样本,因此,本文使用隔离森林法(Isolation forest)<sup>[15]</sup> 去除对指纹图谱有害的异常样本,如图 1(b) 所示。

### 1.3 幅值预处理

去除异常 CSI 数据后,要对幅值进行处理,

幅值的趋势可以表示不同位置,由于采集的数据量大,环境噪声的影响大,因此需要去噪。普通滤波器可以去除一些噪声,但会产生不完全的信号特征。传统小波阈去噪方法(Wavelet domain denoising, WDD)<sup>[16]</sup> 中阈值和阈值函数存在固有缺陷,需要重新设计。因此,本文提出一种改进阈值的小波域去噪方法对采集到的 CSI 幅度进行处理,解决了上述问题。为了避免传统函数的缺点,可以适当降低阈值,并根据每层噪声分量的标准差调整阈值,可以得到阈值公式:

$$\lambda = \sigma_i \sqrt{2 \ln(L)} \quad (6)$$

式中:  $\sigma_i = \text{median}(|w_{j,k}|) / 0.6745$  为第  $i$  层噪声标准差;  $L$  为层数。

根据以下规则对阈值函数进行改进:阈值内的小波系数来自噪声,应加以抑制。阈值函数应在阈值处连续;阈值函数应接近  $y = x$  之外的阈值。因此,以下为改进的阈值函数:

$$w_{j,k,t} = \begin{cases} w_{j,k} - \frac{0.5 \times [\text{sign}(w_{j,k}) \times \lambda]^{m+1}}{w_{j,k}^m}, & |w_{j,k}| \geq \lambda \\ 0.5 \times \text{sign}(w_{j,k}) \times \frac{|w_{j,k}|^{m+1}}{\lambda^m}, & |w_{j,k}| < \lambda \end{cases} \quad (7)$$

式中:  $w_{j,k}$  为阈值处理前的小波系数;  $w_{j,k,t}$  为阈值处理后的小波系数;  $j$  为数据包索引;  $k$  为子载波索引,范围为 0~60;  $\text{sign}(\cdot)$  为符号函数;  $m$  为调整因子,它控制阈值函数的值接近阈值。

为显示阈值函数随  $m$  的变化,假设  $\lambda = 0.05$ ,  $m \in \{1, 3, 5, 8, 13\}$ , 得到图 2 所示的结果。当  $m$  取 1 时,阈值函数趋于软阈值函数,估计值与实际值间总存在固定的偏差,严重影响重构的信号与原始输入信号的逼近程度;当  $m$  取大于 5 时,导致滤波后容易在信号的间断点处产生较大震荡,改进后的阈值满足阈值规则。所以,本文取  $m = 3$ 。

在相对开放的环境下进行比较,要求在随机位置收集 CSI 数据,并在相同的情况下每个位置收集 280 个数据包,从一个比较天线数据中选择数据,如图 3(a) 所示,无线访问点向电脑端发送 280 个连续数据包,并从参考点收集 CSI 数据以获取初始数据集。采集数据集后,将 CSI 数据中的幅值提取出来,对信号进行小波分解,得到每个精确分量和近似分量。在此基础上,采用低阈值对精确分量进行处理,并利用预处理分量进行小

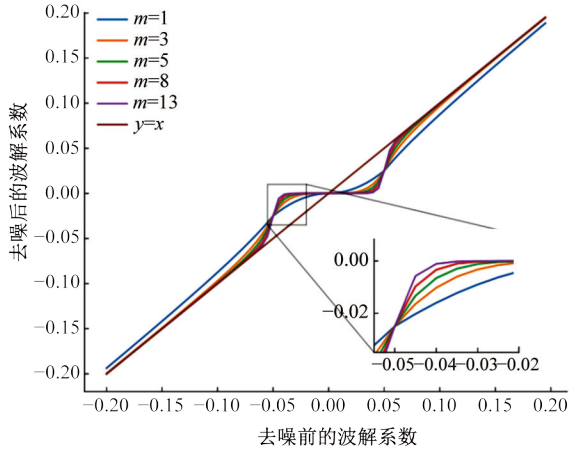


图 2 阈值函数随  $m$  的变化

Fig. 2 Change of threshold function with  $m$

波重构,从而实现去噪信息的获取。在阈值选择准则中,选择极值阈值法,即利用极小极大原理选择阈值,产生最小均方误差极值,然后用 Daubechies(Db6)将信号分解成 5 层,得到最终波形。如图 3(b)所示,对测试环境中的随机位置进行了小波去噪后的比较,进行小波域去噪后,采集到的 CSI 幅度信息更加平滑、直观,能够清晰地反映不同位置的特征,后面和相位结合能更加准确地提高定位精度。

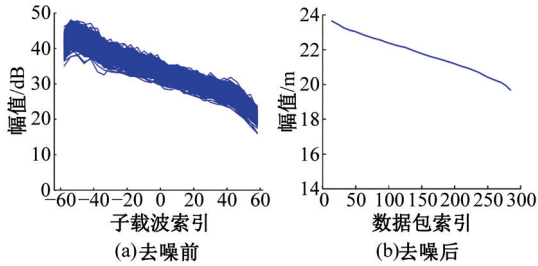


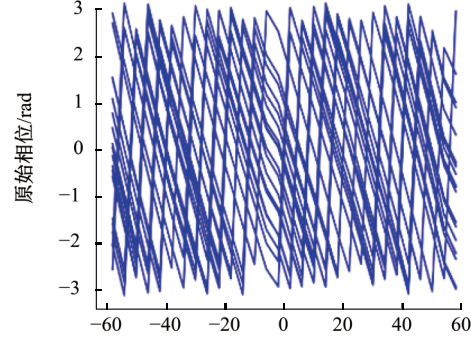
图 3 去噪前后的幅值

Fig. 3 Amplitude before and after denoising

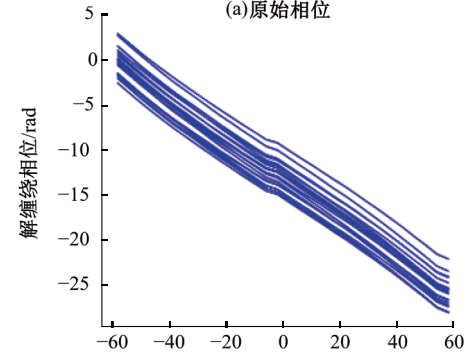
### 1.4 相位预处理

实验采集的 CSI 数据包含幅值信息和相位信息,但原始相位信息不能很好地描述相应的位置,因此,很少将其作为指纹使用。图 4(a)是 60 个数据包的原始相位图,可以看出相位信息布满整个图形,没有明显的规律,理想条件下,两个相邻的数据包应该是相似的,而图中的相位信息并没有规律可言,未经处理的相位角信息不能直接运用到指纹定位中。针对相位角偏移问题,本文采用线性变换<sup>[17]</sup>的方法校正了杂乱的相位信息,得到了较真实相位。假设其线性函数为:

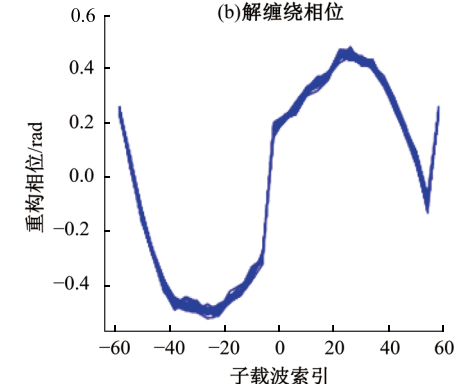
$$y = ax + b \quad (8)$$



(a) 原始相位



(b) 解缠绕相位



(c) 重构相位

图 4 相位经线性变换过程图

Fig. 4 Phase transformation process diagram through linear transformation

已知  $n$  个子载波的信道频率响应相位,假设  $n$  个子载波序列的编号为递增序列  $k_i(i \in [1, n])$ ,那么斜率  $a$  和  $b$  的值计算如下:

$$a = \frac{\dot{\phi}_n - \dot{\phi}_1}{k_n - k_1} = \frac{\phi_n - \phi_1}{k_n - k_1} - \frac{2\pi}{N} \delta \quad (9)$$

式中:  $\phi_1$  和  $\phi_n$  为测量相位;  $\dot{\phi}_1$  和  $\dot{\phi}_n$  为真实相位;  $k_1$  和  $k_n$  为索引号;  $N$  为快速傅里叶的大小;  $\delta$  为信号接收端时间误差。

$$b = \frac{1}{n} \sum_{j=1}^n \dot{\phi}_j = \frac{1}{n} \sum_{j=1}^n \phi_j - \frac{2\pi\delta}{nN} \sum_{j=1}^n k_j + \gamma \quad (10)$$

式中:  $\dot{\phi}_j$  为第  $j$  个子载波提取到的测量相位;  $\phi_j$  为

第  $j$  个子载波的真实相位;  $\gamma$  为相位偏差。

从相位中减去  $\gamma$  就可以得到真实相位角  $\theta_i$ :

$$\theta_i = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^n \phi_j \quad (11)$$

经过线性变换后,图 4(c)展示了相位角经过相位线性变换后呈现明显的规律性变化,相关度增高,规律更加显著。经过适当处理后,这些规律可以应用在室内定位中。

## 2 定位系统实现

### 2.1 系统概述

如图 5 所示,整个系统分为两部分:离线建库阶段和在线定位阶段<sup>[18]</sup>。在离线建库阶段,先进行 CSI 数据采集,然后消除异常 CSI,接着对数据进行预处理,在预处理阶段又分为幅值预处理阶段和相位预处理阶段,然后将经过处理的幅值和相位结合按特征分类存储在指纹数据库中。在线定位过程中,应首先进行 CSI 数据采集,所收集的数据应采用与离线阶段相同的方式进行预处理,并将所处理的数据通过定位算法与指纹数据库进行对比,从而最终获得定位结果。为了减小室内环境变化对 CSI 信号的干扰,本文试图保证测试人员的稳定性和测试环境的稳定性。在离线建库阶段采集数据建立指纹数据库和在线阶段进行定位匹配时,周围环境保持不变。

### 2.2 离线建立指纹数据库阶段

离线指纹数据库是通过使用参考点建立的。通常,目标区域被划分为若干个规则网格,网格角被用作参考点。在训练样本点时,除了 CSI 值和位置坐标外,还应该记录参考点的几何特征,以获得参考点指纹数据,如式(12)所示:

$$R_i = \begin{bmatrix} P(AP_1 CSI_{1i} | (Q_i | S_m)) \cdots P(AP_k CSI_{1i} | (Q_i | S_m)) \\ P(AP_1 CSI_{2i} | (Q_i | S_m)) \cdots P(AP_k CSI_{2i} | (Q_i | S_m)) \\ \vdots \\ P(AP_1 CSI_{ji} | (Q_i | S_m)) \cdots P(AP_k CSI_{ji} | (Q_i | S_m)) \end{bmatrix} \quad (12)$$

式中:  $Q_i = (P_{t_i} | SS_i)$  为参考点的坐标和该点的几何特征值,其中  $SS_i$  描述参考点及其相邻参考点的拓扑信息;  $S_m$  表示划分的区域  $m$ , 其中  $M = 1, 2, \dots, N$ , 总共  $N$  个类别;  $AP_k$  为  $S_m$  区域的有效 AP;  $CSI_j$  为与  $S_m$  集群区域相关的参考点修改后的 CSI 值。

建立指纹数据库的具体流程为:首先在待测试区域内选择  $a$  个参考测试点,然后依次采集每个 AP 的特征值  $Z_{acc}$ , 每个点的特征值  $Z_{a1}, Z_{a2}, \dots, Z_{acc}$  对应的坐标为  $(X_a, Y_a)$ 。最后,所有指纹数据都存储在指纹数据库中。目前,指纹定位有一个普遍的限制性:指纹数据库是唯一的。

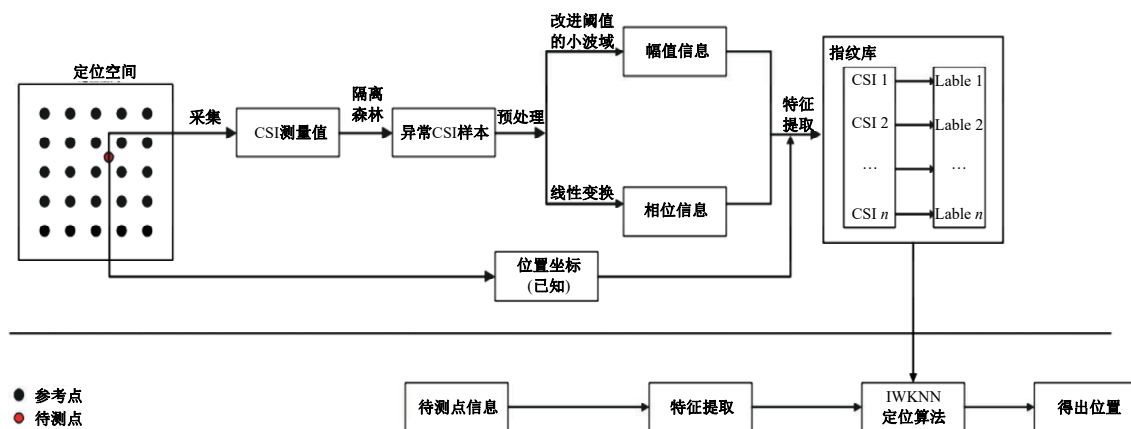


图 5 指纹定位架构图

Fig. 5 Fingerprint positioning architecture diagram

### 2.3 在线定位阶段

K 近邻算法本质是通过计算特征值间的欧氏距离来分类。它的原理是将已知类别的样本作为参考点,计算未知点与所有已知参考点之间的距离,然后,从这些距离中选择与未知点欧氏距离最近的  $K$  个已知参考点,根据少数服从多数的原理,

将未知样本归类为占比最多的类别。不过,该算法的缺点是它未考虑最近参考点之间的距离差别,从而影响定位的准确性<sup>[19]</sup>。针对以上问题,本文提出一种 IWKNN 算法,该算法的基本原理为:传统的 WKNN 算法解算出待定位点的位置坐标后,本文算法会在其基础上进行重复定位,就

可得到最终定位坐标。

传统的 $K$ 近邻算法一般先算待定位点与指纹库参考特征值组间的欧氏距离,公式如下:

$$\text{dis}(S, t) = \sqrt{\sum_{k=1}^n (S_n(i) - t)^2} \quad (13)$$

式中: $\text{dis}(S, t)$ 为待定位点指纹库中第 $i$ 个指纹的距离; $S_n(i)$ 为第 $i$ 个采样点收到的 $n$ 个路由器的CSI; $t_i$ 为待测点指纹特征。

取间隔最近的 $k$ 个参考点的平均坐标来预计最终待测点的位置坐标:

$$(x, y) = \frac{1}{k} \sum_{i=1}^k (x_i, y_i) \quad (14)$$

本文提出的IWKNN算法,它为每个邻近参考对象赋予一个权值。该权值是将待定位的目标点和参考点之间的距离经过高斯变换而形成。

$$W_{ki} = \beta e^{-\frac{(L_{ki} - L_{i\min})^2}{2\alpha^2}} \quad (15)$$

式中: $W_{ki}$ 为第 $i$ 个邻近参考点的权值; $L_{ki}$ 为待定位点到第 $i$ 个邻近参考点的距离; $L_{i\min}$ 为待定位点到邻近参考点的最小距离; $\beta$ 为最大权重值,通常取1; $\alpha$ 为半峰宽度,是 $k$ 个近邻参考点距离的标准差,描述了待定位点对各个近邻参考点的信任度, $\alpha$ 越大,对较远距离的参考点信任度越大,干扰随之变大, $\alpha$ 越小,对近距离的参考点信任度越大,鲁棒性越差。

给每个邻近参考点赋予对应的权重后,WKNN算法即可计算出待定位点的定位坐标。

$$(x, y) = \frac{\sum_{i=1}^k W_{ki}}{\sum_{i=1}^k W_{ki}} \times (x_i, y_i) \quad (16)$$

式中: $(x, y)$ 为待测点的坐标; $k$ 为所选择匹配的参考点的总数量; $(x_i, y_i)$ 为第 $i$ 个参考点的具体坐标。

本文提出了一种重复定位算法,该算法计算第一次得到的定位坐标与邻近点之间的欧氏距离,并通过高斯变换对计算出的距离进行权值计算,从而使最终坐标与距离最近的邻近点产生联系。改进后的计算公式如式(17)所示:

$$(\hat{x}, \hat{y}) = \frac{\sum_{i=1}^k W_{ci}}{\sum_{i=1}^k W_{ci}} \times (x_i, y_i) \quad (17)$$

式中: $(\hat{x}, \hat{y})$ 为计算的位置坐标; $W_{ci}$ 为由高斯变换而赋予欧氏距离的权值。

### 3 分析与验证

#### 3.1 实验场景

实验分别在西安科技大学教学楼501教室和煤炭科技楼大厅进行,如图6和图7所示。实验采用1个支持802.11协议的TP-Link路由器作为Wi-Fi接入点(AP),AP放在测试区域的左上角,并将其放置在距离地面0.2 m的高度,以减小地面反射对CSI值的影响。数据采集选择晚上或周末进行,可以减小教室内其他人员的干扰,确保数据的准确性。在采集数据时,一个不携带任何具有信息发射功能的移动设备的人可以站立在采样位置点1~2 min,这样做可以减弱移动设备带来的信号干扰。同时,在数据采集过程中还可以采取其他措施,比如使用屏蔽设备来防止外界无关信号的干扰,确保采集到的数据更加准确可靠。

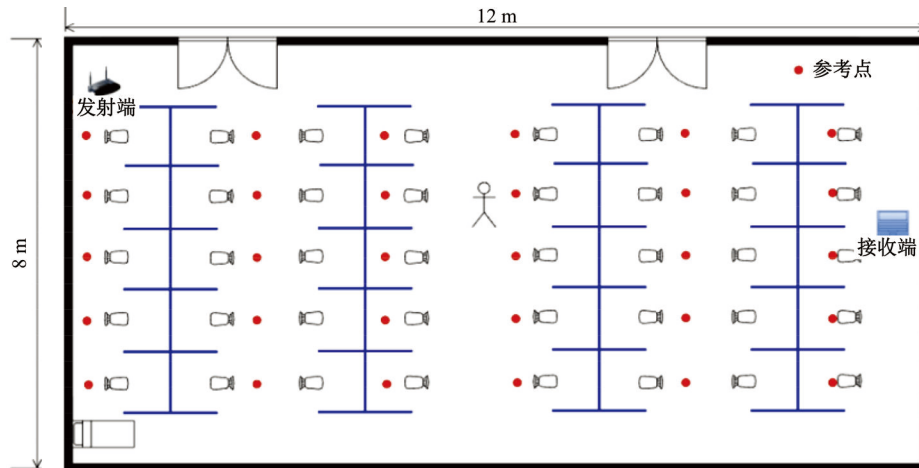


图6 501教室(实验环境示意图)

Fig. 6 Classroom 501(schematic diagram of experimental environment)

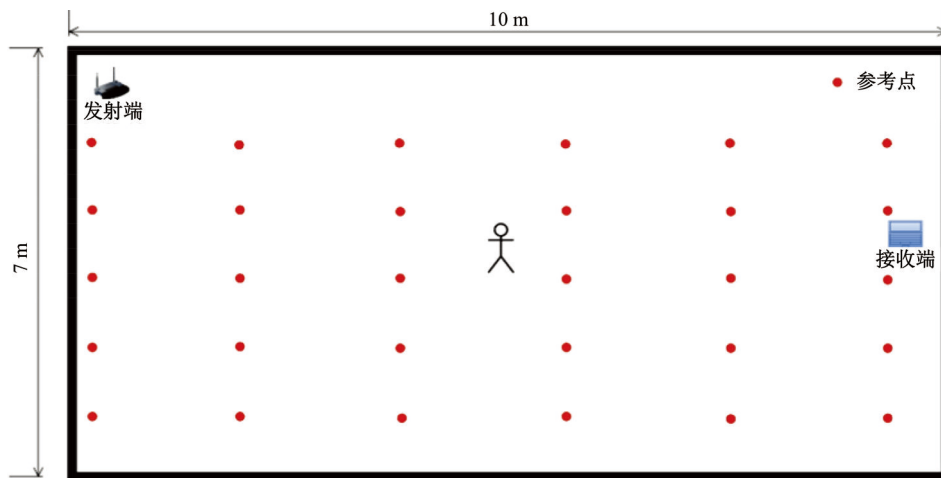


图7 煤炭科技楼大厅(实验环境示意图)

Fig. 7 Coal Science and technology building hall (experimental environment diagram)

### 3.2 实验分析和性能评估

在室内进行定位实验时,选取了5个待测试点,每个待测点进行10组实验,并保证周围环境对实验数据采集不会产生过多的影响。

数据包的影响:在基于CSI的室内定位算法中,样本数量是影响定位精度的关键因素之一。然而,样本量越大,实验的复杂性就越高,需要收集和处理的的数据量也就越大。此外,当样本数量达到一定数量时,对定位精度的影响很小。因此,选择合适数量的样本是确保定位精度和减小数据

大小的关键。

为了分析样本量对实验平均定位误差的影响,在两种不同的环境下进行了对比实验。当分包率固定在50包/s时,测试器收集1000、500、250、100个数据包。同样,当分包率固定在10包/s时,测试器收集1000、500、250、100个数据包。

如图8所示,在相同的环境下,收集了相同数量的数据包,且数据包的平均定位误差较小。在相同的分包速率下,1000、500、250个数据包的平均定位误差并不是很大,但当样本数减少到100

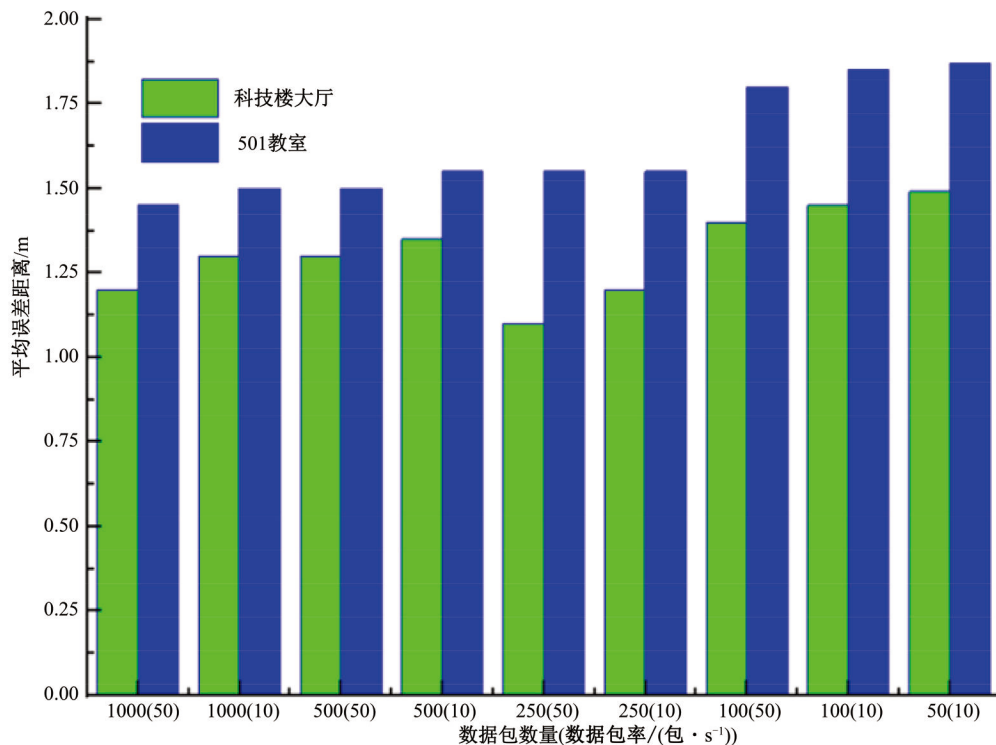


图8 数据包数量对定位精度的影响

Fig. 8 Impact of number of data packets on positioning accuracy

个数据包和 50 个数据包时,平均定位误差明显增大。这表明,小样本不足以提供足够的特征来反映整体环境,同时可以看出,教室的总体定位效果不如大厅,因为大厅的实验环境相对空旷,教室桌子较多,多径效应比较明显。为了保证定位精度和减小数据大小,本文在每个参考点采集了 250 个数据包作为样本数据。

数据质量的影响:室内定位选择幅值信息的方法有很多种,幅值的处理方法也不尽相同。本文使用改进阈值的小波域去噪方法对幅值信息进行处理,对提高定位精度有很大帮助。同时,在大多数室内定位算法中,相位很少被使用,原始相位很难为定位提供相应的特征。本文将线性变换与处理后的幅值信息相结合,并与参考点一一对应,构建了指纹数据库,在一定程度上去除了环境因素引起的噪声,同时保留了每个参考点的特性。

为了分析数据质量对定位性能的影响,本文设计了 4 组对照实验。将原始幅值用作指纹数据进行定位,将改进阈值小波域去噪后的幅值用作指纹信息进行定位,将线性变换后的相位(Phase after linear transformation, PALT)用作指纹信息来进行定位,并将本文中改进阈值小波域去噪后的幅值和线性变换后的相位作为指纹数据进行定位。

如图 9 和图 10 所示,幅值是指将原始幅值直接用作指纹数据,小波域去噪是指将改进阈值小波域去噪声后的幅值用作指纹数据,PALT 是指线性变换后作为指纹数据的相位,幅值和相位(A 和 P)是指将改进阈值小波域去噪后的幅值和线性变换后的相位作为指纹数据。在科技楼大厅环境中,使用幅值的定位误差较大,其次是使用线性变换后相位的定位误差,并且使用改进阈值小波域的定位误差优于两者。然而,定位效果并不理想。同样,在科技楼大厅中,与教室环境中的结果类似,单独使用幅值或相位的 3 种方法的定位效果都不是很理想,而本文方法在两种实验环境中都有很大的优势。

K 值的选取:本文使用不同的 K 值进行实验,然后比较 WKNN 和 IWKNN 算法使用不同 K 值所得到的平均误差。根据图 11 的结果,可以看出不同的 K 值对定位结果有着不同的影响。

图 11 展示了在使用不同 K 值时,改进 WKNN 和 WKNN 算法的平均定位误差。从图 11

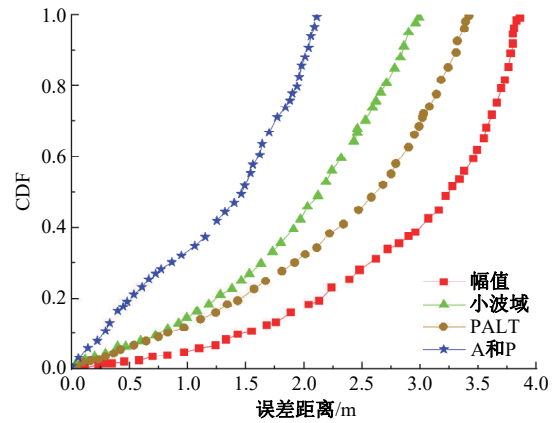


图 9 501 教室(无人)数据质量的 CDF

Fig. 9 Classroom 501 (unmanned) CDF for data quality

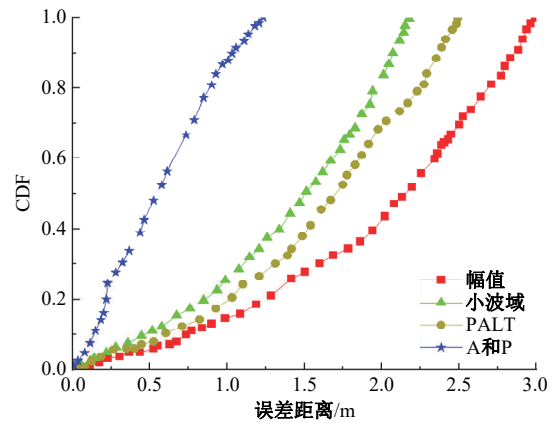


图 10 煤炭科技楼大厅数据质量的 CDF

Fig. 10 Building hall CDF for data quality

中可以观察到以下情况:①当  $K=3$  时, WKNN 算法的平均误差在图中呈现最低值,当  $K$  值缓慢增加时,定位误差也随之增大;②当  $K=3$  时, IWKNN 算法的平均误差在图中呈现最低值,当  $K$  值缓慢增加时,定位误差基本保持不变。综合

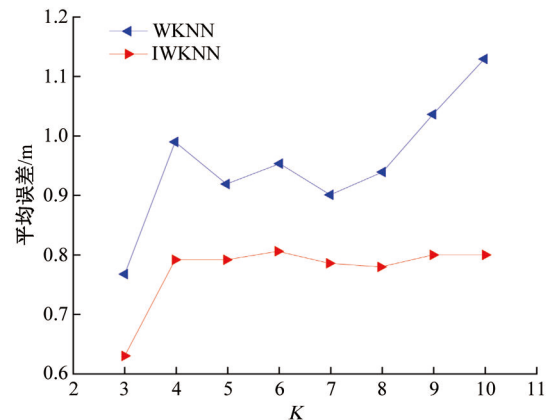


图 11 两种不同算法不同 K 值的平均误差

Fig. 11 Average error of different K values for two different algorithms

以上结果,根据图 11 的分析,确定  $K=3$  可能是一个合适的选择来获得最优的定位结果。这意味着在该研究中,使用改进 WKNN 算法且选择  $K=3$  时可以获得最小的平均定位误差。

图 12 展示了 WKNN 和 IWKNN 算法在 501 教室(无人)环境、501 教室(有人)环境与煤炭科技楼大厅环境的定位误差对比。在 501 教室(无人)环境,传统 WKNN 算法的定位误差大约为

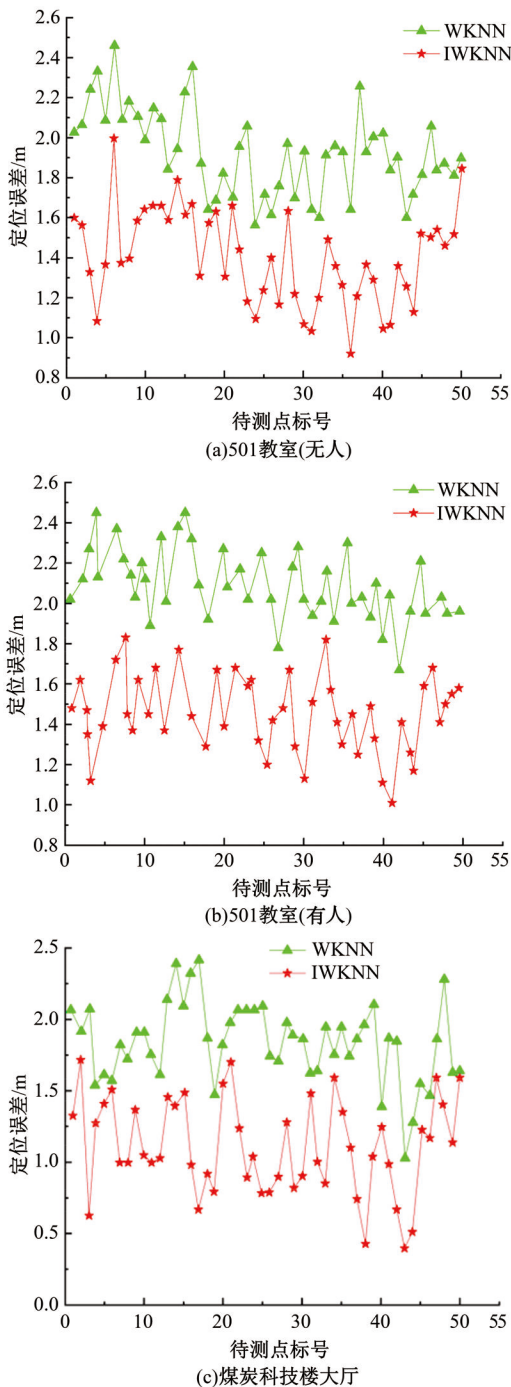


图 12 两种算法误差对比图

Fig. 12 Comparison of error between two algorithms

1.89 m,改进 WKNN 算法的定位误差约为 1.40 m。在 501 教室(有人)环境,传统 WKNN 算法的定位误差约为 2.15 m,改进 WKNN 算法的定位误差约为 1.52 m。在煤炭科技楼大厅中,传统 WKNN 算法的定位误差约为 1.83 m,而改进 WKNN 算法的定位误差约为 1.10 m。

为了验证教室有学生对本文算法定位精度的影响,选择场景为上午教室有人且学生随机活动。图 13 展示了本文算法在有人情况下和无人情况下的定位精度对比,由于教室内人员随机活动存在复杂的 NLOS 时延和多径效应,数据含有更多的随机噪声,有人情况下定位误差为 1.52 m,无人情况下定位误差为 1.4 m,两者之间存在 0.12 m 的定位误差,说明本文所改进的去噪方法能很好去除环境所引起的噪声,使本文算法在干扰下具有一定的鲁棒性。

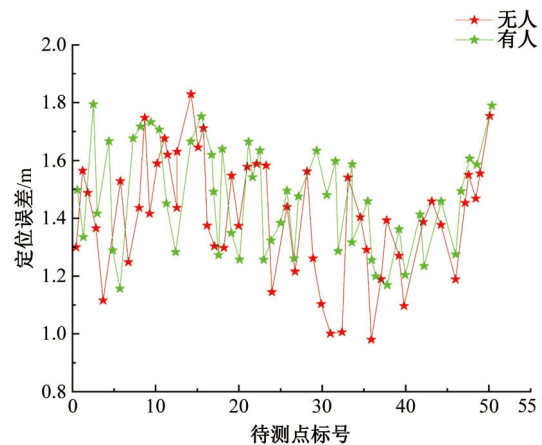


图 13 本文算法在不同干扰下的误差对比图

Fig. 13 Comparison chart of error of the algorithm in this article under different interferences

图 14 展示了支持向量机(Support vector machine, SVM)、贝叶斯(Bayes)、随机森林(Random forest, RF)、KNN、WKNN 和 IWKNN 算法在教室环境与煤炭科技楼大厅环境的定位精度对比。由于教室中桌子数量大多存在很强的多径效应,定位精度会降低。然而,由于本文提出的方法是稳定的,在强多径效应下仍然可以保持一定的定位精度,约 74% 的测试位置误差控制在 1 m 以内,平均定位误差为 1.4 m;在多径效应影响不大的情况下,约 81% 的测试位置误差控制在 1 m 以内,平均定位误差为 1.1 m,展示出明显的定位误差减小效果,提高了定位系统的稳定性,可满足室内定位需求。

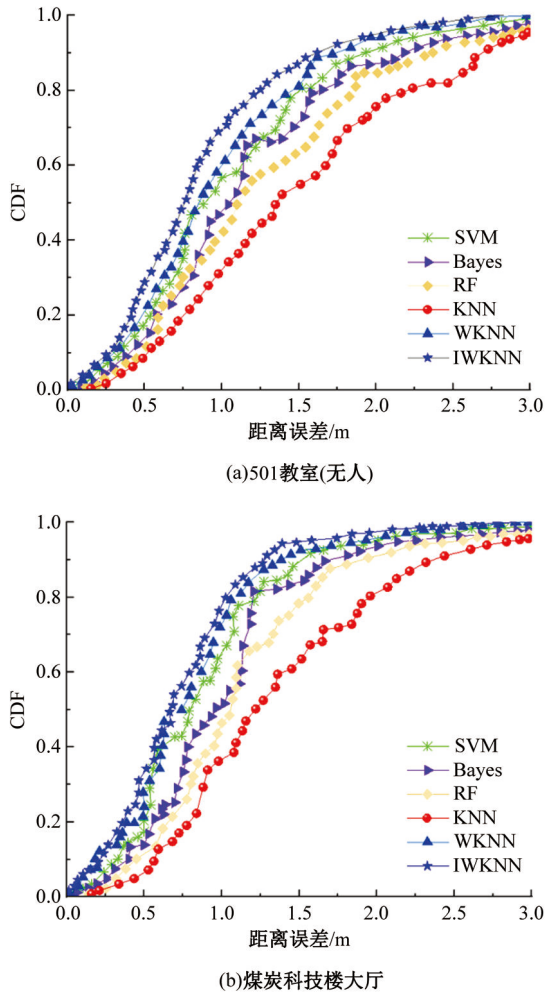


图14 6种定位算法误差累计分布函数图

Fig. 14 Cumulative distribution function diagram of six positioning algorithm errors

#### 4 结束语

本文提出了一种基于改进加权的 $K$ 最邻近算法的CSI被动室内指纹定位系统。为了消除室内环境对信号的干扰,获得更准确的指纹数据,使用隔离森林对离线阶段采集的CSI数据进行了处理,对幅值信息使用改进阈值的小波域去噪,使用线性变换方法对相位信息进行相位角处理,最后,通过与参考点一一对应,构建了较稳定的离线指纹数据库,在一定程度上提高了定位精度。在线阶段提出了对加权 $K$ 近邻法对待定位点进行实时定位。本文在501教室和煤炭科技楼大厅这两个不同的实验环境中验证了系统的性能,与其他基于信道状态信息的指纹室内定位方法相比,本文提出的方法具有更高的定位精度,可满足养老院老人监护和学校学生签到打卡等需求。目前,数据的收集和处理非常耗时,数据量越大,定位精度

就越高。然而,数据量越大,处理难度就越大。随着深度学习方法的应用,将在未来的工作中重点提高定位精度,减少数据量。

#### 参考文献:

- [1] 陈锐志,郭光毅,陈亮,等.室内高精度定位技术研究应用现状与发展趋势[J].武汉大学学报:信息科学版,2023,48(10):1591-1600.  
Chen Rui-zhi, Guo Guang-yi, Chen Liang, et al. Research and application status and development trends of indoor high precision positioning technology [J]. Journal of Wuhan University(Information Science Edition), 2023, 48(10):1591-1600.
- [2] 张楠,史建华,亿吉,等.基于加权质心定位的井下移动目标实时跟踪方法[J].吉林大学学报:工学版,2023,53(5):1458-1464.  
Zhang Nan, Shi Jian-hua, Yi Ji, et al. Real-time tracking method of underground moving target based on weighted centroid location[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(5):1458-1464.
- [3] Liu J, Wang L, Guo L, et al. A research on CSI based human motion detection in complex scenarios[C]// Proceedings of the IEEE 19th International Conference on E-Health Networking, Applications and Services, Dalian, China, 2017:1-6.
- [4] 王玫,宋志远.基于TrAdaBoost算法为内核的行人航迹推算技术[J].吉林大学学报:工学版,2023,53(8):2364-2370.  
Wang Mei, Song Zhi-yuan. Pedestrian dead reckoning technology based on TrAdaBoost algorithm[J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53(8):2364-2370.
- [5] 张海军,孙学成,赵小虎,等.煤矿井下UWB人员定位系统研究[J].工矿自动化,2022,48(2):29-34.  
Zhang Hai-jun, Sun Xue-cheng, Zhao Xiao-hu, et al. Research on UWB personnel positioning system in coal mine[J]. Industrial and Mining Automation, 2022, 48(2): 29-34.
- [6] Wei F, Wu Y, Xu S W, et al. Accurate visible light positioning technique using extreme learning machine and meta-heuristic algorithm[J]. Optics Communications, 2023, 532:No. 129245.
- [7] 郭苏,邱继芸,余江涛,等.变电站室内导航系统蓝牙定位技术与算法研究[J].信息技术,2020,44(8):141-145.  
Guo Su, Qiu Ji-yun, Yu Jiang-tao, et al. Research

- on Bluetooth localization technology and algorithm of substation indoor navigation system[J]. Information technology, 2020, 44(8): 141-145.
- [8] Wang F, Feng J, Zhao Y, et al. Joint activity recognition and indoor localization with WiFi fingerprints[J]. IEEE Access 2019, 7:80058-80068.
- [9] Zhuo Y, Zhu H, Xue H, et al. Perceiving accurate CSI phases with commodity Wi-Fi devices[C]//Proceedings of the IEEE INFOCOM—IEEE Conference on Computer Communications, Atlanta, USA, 2017: 1-9.
- [10] Yang W, Gong L, Man D, et al. Enhancing the performance of indoor device-free passive localization[J]. International Journal of Distributed Sensor Networks, 2015, 11(11): 79-90.
- [11] Wang X Y, Gao L J, Mao S W, et al. DeepFi:deep learning for indoor fingerprinting using channel state information[C]//IEEE Wireless Communications and Networking Conference, New Orleans, USA, 2015: 1666-1671.
- [12] Rao X P, Li Z, Yang Y B, et al. Device-free passivewireless localization system with weighted transferable discriminative dimensionality reduction method[J]. Telecommunication Systems: Modeling, Analysis, Design and Management, 2020, 75(1):15-29.
- [13] 王博远, 贾瑞才, 贾浩男. 高精度位置服务终端技术与应用[J]. 无线电工程, 2016, 46(11): 5-8.  
Wang Bo-yuan, Jia Rui-cai, Jia Hao-nan. High precision location service terminal technology and application[J]. Radio Engineering, 2016, 46(11): 5-8.
- [14] Wu C S, Zheng Y, Zhou Z M, et al. Phase U:real-time LOS Identification with Wi-Fi[C]//Computer Communications, Hong Kong, China, 2015: 2038-2046.
- [15] Fang N, Fang X W, Lu K. Anomalous Behavior detection based on the isolation forest model with multiple perspective business processes[J]. Electronics, 2022, 11(21): No. 3640.
- [16] Wu Y, Gao G, Cui C. Improved wavelet denoising by non-convex sparse regularization under double wavelet domains[J]. IEEE Access 2019, 7: 30659-30671.
- [17] Zhuo Y, Zhu H, Xue H. Identifying a new non-linear CSI phase measurement error with commodity WiFi devices[C]//Proceedings of the IEEE 22nd International Conference on Parallel and Distributed Systems(ICPADS), Wuhan, China, 2016:72-79.
- [18] 邓中亮, 尹露, 唐诗浩. 室内定位关键技术综述[J]. 导航定位与授时, 2018, 5(3): 14-23.  
Deng Zhong-liang, Yin Lu, Tang Shi-hao. Overview of key techniques of indoor positioning[J]. Navigation Positioning and Timing, 2018, 5(3): 14-23.
- [19] 史林军, 戴滔, 劳文洁, 等. 基于改进KNN算法的新能源发电单元运行状态识别[J]. 电力自动化设备, 2024(5): 65-72.  
Shi Lin-jun, Dai Tao, Lao Wen-jie, et al. New energy power generation unit operating state recognition based on improved KNN algorithm[J]. Power automation equipment, 2024(5): 65-72.