

doi:10.19306/j.cnki.2095-8110.2019.06.017

# 信道状态信息指纹定位算法性能评价方法研究

蒋天润, 尹露, 邓中亮, 王子阳

(北京邮电大学电子工程学院, 北京 100876)

**摘要:**基于WiFi的定位技术大多使用接收信号强度,但该方法受多径和噪声干扰较大,精度有待提高。信道状态信息(channel state information, CSI)能够更加精细地描述信道状态,具有更强的稳定性。将CSI作为格点特征建立指纹定位数据库,利用该指纹库和在线测量数据,比较了多种定位算法在位置指纹法中的定位效果,并提出了评价KNN、wKNN和随机森林算法的一种评价依据和样本容量扩充方法,分析了三种方法随样本容量增加时定位时间和定位精度的稳定性,从包含定位精度在内的多种角度更加全面地评估了三种方法。结果表明,在以上三种定位算法中,随机森林算法的定位时间与定位精度的稳定性最好。

**关键词:**室内定位;信道状态信息;定位稳定性

中图分类号:TN967.2

文献标志码:A

开放科学(资源服务)标识码(OSID):

文章编号:2095-8110(2019)06-0113-06



## Performance Evaluation Method of Fingerprint Localization Algorithm Based on Channel State Information

JIANG Tian-run, YIN Lu, DENG Zhong-liang, WANG Zi-yang

(School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China)

**Abstract:** Most location techniques based on WiFi use received signal strength. However, this method is subject to multipath and noise interference, and its accuracy needs to be improved. CSI (Channel State Information) can describe channel state more finely and has stronger stability. In this paper, CSI is used as a lattice feature to establish a fingerprint location database. Using this fingerprint library and online measurement data, we compared the positioning performance of different location algorithms. We proposed an index for evaluating KNN, wKNN and random forest algorithm and a sample capacity expansion method. Then we analyzed the stability of positioning time and positioning accuracy of three algorithms when the sample size increased. We evaluated these algorithms not only according to positioning accuracy, but also from other aspects. The results show that, in the above three localization algorithms, the stability of positioning time and positioning accuracy of random forest algorithm are the best.

**Key words:** Indoor location; Channel state information; Localization stability

收稿日期:2019-02-01;修订日期:2019-10-10

基金项目:国家重点研发计划(2016YFB0502001)

作者简介:蒋天润(1996-),男,硕士,主要从事室内定位方面的研究。E-mail: jtr@bupt.edu.cn

## 0 引言

室外条件下,用户使用的定位设备主要基于全球定位系统(Global Positioning System, GPS)。由于室内的接收机受到建筑物的遮挡,使得接收到的卫星定位信号衰减很快,甚至完全无法接收到信号。随着移动设备和无线传感器的大量部署,关于室内有源主动定位的研究取得了长足发展,但它在室内定位时需要人携带有关的传感器设备,使得有源定位的广泛应用受到较大阻碍。

随着移动互联网技术的迅速发展,人们对于获得高精度位置信息的需求愈加强烈,未来室内定位将成为一项生活所必须的基本服务需求。目前,室内定位技术种类繁多,主要有基于射频识别(Radio Frequency Identification, RFID)、超声波、红外线和WiFi等的室内定位技术<sup>[1]</sup>。RFID定位需要提前布置硬件设备,实现大范围推广的成本很高<sup>[2]</sup>。红外定位技术的精度会受到室内灯光、热源的影响,另外室内的各种障碍物也会遮挡红外线,导致红外线不能传输很远的距离,以及定位误差增大<sup>[3]</sup>。由于超声波在空气中传播时衰减速度很快,距离稍远超声波定位技术的定位误差就会迅速增大。无线局域网已在人口密集区域广泛普及,各处的室内都大量部署了无线上网设备,基于WiFi的室内定位技术具有定位耗时短、覆盖面积大、成本低等优点,因此该技术成为了室内定位研究领域的热点方向。

目前基于WiFi的室内定位研究主要基于接收信号强度指示(Received Signal Strength Indicator, RSSI)。RSSI描述了接收信号功率的大小,由于目前多数终端设备都支持获得RSSI信息,故目前的无线通信技术都使用RSSI作为评价信道质量的标准,从而改进通信系统。基于WiFi的RSSI指纹定位法将处理后的RSSI作为每个位置的指纹特征,从而开展基于位置指纹库的指纹定位。RSSI是多条路径信号的简单叠加<sup>[4]</sup>,因此RSSI并不能充分描述子信道受环境的干扰程度。此外,尽管单个信道的信号波动较小,但多条信道中信号叠加的RSSI波动范围甚至达到了5dB<sup>[5]</sup>。单一使用RSSI作为信号特征进行指纹定位不能满足高精度、高稳定性的室内定位需求。

相较于RSSI,信道状态信息(Channel State Information, CSI)能够更好地描述多径信道,区分来自多个信道的多径信号,在静态环境下具有较好的

稳定性。CSI的一个数据分组中同时包含着多个子载波的幅频响应和相频响应。在室内定位中以子信道的多维矩阵对单一的RSSI进行扩展来构建格点特征,由此可以获得更高的定位精度。近年来,Inter和Atheros网卡供应商对其部分网卡固件程序进行了处理,并有相关组织开源了对应的软件开发包,使得可以对Linux系统以及Windows系统下的网卡开源驱动程序进行修改,利用调试模式来获取某些无线网卡的CSI数据。使用大量部署的基于802.11n标准的WiFi路由设备可以获得CSI<sup>[6]</sup>。基于CSI的指纹定位需要的额外硬件设备较少,实现成本低,具有广泛的应用前景。

## 1 CSI与指纹定位算法

### 1.1 信道状态信息

目前,使用装配Intel 5300网络适配器的普通WiFi设备能够获得一个采样个数为30的CSI信息,利用兼容IEEE802.11.n的无线网卡从接收的数据包中提取一组CSI,每组CSI代表一个正交频分复用子载波的幅度与相位,如式(1)所示

$$H(k) = \|H(k)\| e^{j\angle H(k)} \quad (1)$$

式中: $H(k)$ 是第 $k$ 个子载波的CSI, $\|H(k)\|$ 是第 $k$ 个子载波的幅度, $\angle H(k)$ 是第 $k$ 个子载波的相位。

### 1.2 指纹定位法

指纹定位法主要包括2个阶段。离线训练阶段即数据采集和处理过程,该过程主要工作是采集格点信息,将该信息处理后构建为格点特征,常用的格点特征包括相位和幅度等;实现格点位置与格点特征的映射;建立无线地图和指纹定位数据库。

在在线定位阶段,实时采集待定位点的格点特征,输入至定位服务器,服务器将定位端实时上报的数据与离线阶段建立的位置指纹库的数据进行匹配,计算出实时数据最佳匹配的参考点。在线阶段的主要思路是依赖一定的准则选择与实时数据最为接近的参考点。

### 1.3 研究现状

文献[7]表明WiFi的CSI具有更高的时间稳定性,并具有从多径效应中获益的能力,因此适用于准确的存在检测和定位。文献[8]表明由于室内环境的复杂性对多径信号传播的影响,难以建立精确的信号传播模型,因此基于CSI的指纹定位法优

于基于传统传播模型的室内定位方法。文献[9]基于 CSI 的子信道频率的不同,用合并且求均值后的 5 个子载波的 CSI 代替原来的 30 个子载波的 CSI,并对多根天线的 CSI 求平均值处理,提出了一种基于后验概率的指纹定位算法。Qian K. 等<sup>[10]</sup>利用 CSI 导出运动诱发的多普勒频移,并提取出该频移用于确定运动方向。

## 2 采集 CSI 信息与制作指纹库

在北京邮电大学的新科研楼内选择一间 10m×10m 的实验室,室内放置有大量物品以及人员走动。根据地砖将室内划分成 1m×1m 的方格,在每个参考点测量 CSI 数据,每个格点的 CSI 数据存储在若干个数据包中,如图 1 所示。

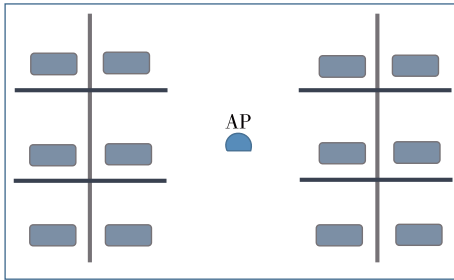


图 1 数据采集环境示意图

Fig. 1 Data collection environment

经过上述操作获得的原始 CSI 数据含有高频噪声,需要对原始 CSI 数据进行低通滤波处理。使用  $h$  表示该巴特沃斯滤波器的传递函数。单个数据包中的 CSI 数据为  $\mathbf{H}(1 \times 30)$ ,滤波后的矩阵记为  $\mathbf{H}'(1 \times 30)$

$$\mathbf{H}' = h * \mathbf{H} \quad (2)$$

滤波处理后每个格点处一个数据包的 CSI 数据对应矩阵  $\mathbf{H}'(1 \times 30)$ 。正交频分复用(Orthogonal Frequency Division Multiplexing, OFDM)系统中所有信道是相互独立的,单个子信道发生深度衰落的概率比多个子信道发生深度衰落的概率小得多。因此选取每个格点处的 50 个数据包中第一根天线的幅度,计算这 50 个数据包对应矩阵  $\mathbf{H}'(1 \times 30)$  的平均值  $\mathbf{H}_{\text{avg}}^i(1 \times 30)$  作为该格点处的特征。该步骤如式(3)所示

$$\mathbf{H}_{\text{avg}}^i = \frac{\sum_{l=1}^{50} \mathbf{H}'_l}{50} \quad (3)$$

## 3 算法关键参数的测试与分析

### 3.1 测试 KNN 与 wKNN 算法

K 最近邻(K-Nearest Neighbor, KNN)算法,  $K$  表示距离自己最近的  $k$  个数据样本<sup>[11]</sup>。KNN 算法使用的相似度量是曼哈顿距离或欧式距离,该算法所选择的邻居都是已经正确分类的对象。该算法在定类决策上只依据最邻近的一个或者几个样本的类别来决定待分样本所属的类别。

在实际的生产实践中, KNN 算法的缺陷是找到的近邻并不是实际意义上的近邻。特征向量之间一般有较大可能性是相关的,最近距离的求解过程中没有考虑特征之间的关系,故距离的计算不够准确,从而影响了定位的准确度。为了解决以上问题,使用加权 K 最近邻(weighted K-Nearest Neighbor, wKNN)算法<sup>[12]</sup>。

将每一邻居样本的坐标与对应权重相乘,结果相加<sup>[13]</sup>,求出总和后,将结果除以全部权重之和。估计坐标的计算方法如式(4)所示

$$f(x) = \frac{\sum_{i=1}^k D_i \omega_i}{\sum_{i=1}^k \omega_i} \quad (4)$$

式中,  $D_i$  表示邻居样本的坐标,  $\omega_i$  表示近邻样本的权重,  $f(x)$  是预测的坐标结果。

$$\omega_i = 1 / (\text{distance} + \text{const}) \quad (5)$$

式中,权重  $\omega_i$  是距离数值加上一个常数的倒数。

在图 2 和图 3 中,蓝色虚线代表了测试所走过的真实路径,\* 号代表在蓝色真实路径上均匀取点所对应的定位结果。在  $K=5$ ,离线指纹库容量为 100,在线数据集容量为 35 时, KNN 定位误差是 2.732m, wKNN 定位误差是 2.651m。

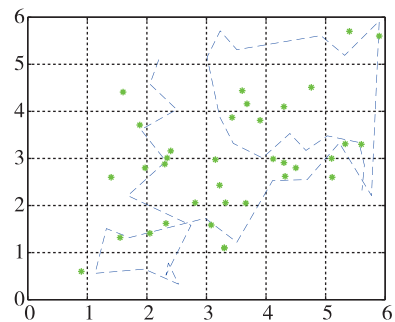


图 2 KNN 定位结果图

Fig. 2 The locating result of KNN algorithm

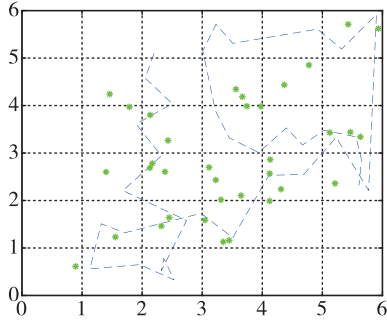


图3 wKNN 定位结果图

Fig. 3 The locating result of wKNN algorithm

在离线数据集容量为 100, 在线数据集容量为 35 时, 改变  $K$  值, 观察其对定位误差的影响。

从图 4 可以看出, 两种算法的定位误差随  $K$  的变化曲线有一定的相似性。在  $K < 9$  时, 两种算法的定位误差随  $K$  的增大而减小; 在  $K > 9$  后, 两种算法的定位误差随  $K$  的增大反而逐渐增大, 但此时的最大定位误差仍小于  $K < 9$  时的最大定位误差。

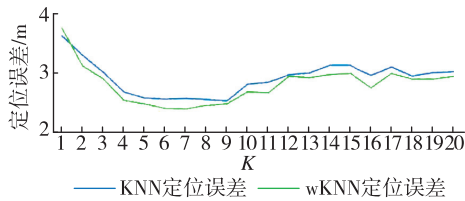


图4 两种算法定位误差随  $K$  值变化图

Fig. 4 The tendency chart of positioning error with  $K$

OFDM 系统中的每个子信道是相互独立的<sup>[14]</sup>, 在 CSI 中, 理论上每个子信道的幅度也是相互独立的, 即在本次测试中理论上各特征向量之间是不相关的。因此, 使用 wKNN 算法对定位精度的改善有限, 这也解释了以上两种算法的定位误差随  $K$  的变化趋势较为相似的原因。

较小的  $K$  值会导致邻居不够, 定位精度受此影响<sup>[15]</sup>, 因此  $K < 9$  时定位误差随  $K$  的变化趋势如上所示; 但是过大的  $K$  值会导致不属于同一类的点也被分入邻居内, 使得定位误差增大, 这是  $K > 9$  后定位误差随  $K$  增大而增大的原因。

### 4 三种算法稳定性测试与分析

#### 4.1 算法评价指标与样本集扩充方法

通常评价一种定位算法性能的依据是该算法的定位精度, 但该评价依据并没有考虑到在线测量数据增加时, 该种算法的稳定性问题。因此本文提

出了一种评价 KNN、wKNN 和随机森林算法优劣的依据: 三种算法定位时间的稳定性和定位精度的稳定性。

某种算法的运行时间在一定程度上反映了该算法的计算量和资源占用的多少, 由于算法的运行时间具有一定的时变性和不稳定性, 选取运行时间的标准差占平均运行时间的百分比作为评价标准。定位精度的稳定性可以直接使用在线样本容量为 100~20000 时定位精度的平均值与标准差进行评价。

如果将  $10m \times 10m$  的办公区域划分为  $0.5m \times 0.5m$  的方格, 则指纹库中最多只有 400 个点的指纹信息。0.5m 的方格长度相较于超过 2m 的定位误差很小, 可以认为在该  $10m \times 10m$  的区域中, 至多只存在 400 个定位结果, 那么, 在线数据集所有可能点的坐标都在这 400 个点中产生。因此, 样本容量的扩充方法如图 5 所示, 将新的在线样本集逐个编号, 每次从以上的原始样本集中随机取样, 重复 400 次, 将 400 个取样结果全部用来扩充原始样本集, 将扩充的结果作为新样本集, 直至样本容量符合实验要求。

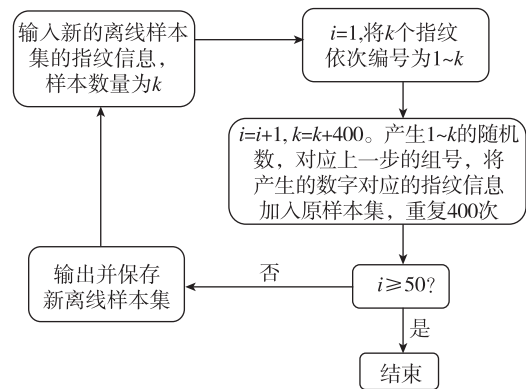


图5 样本集扩充方法

Fig. 5 Sample capacity expansion method

#### 4.2 稳定性测试与分析

测试软件均使用 Pycharm。在线数据集的样本数量为 100~20000。KNN 与 wKNN 算法中  $K = 6$ , 随机森林算法中决策树数量为 110,  $max\_features = 0.45$ ,  $max\_depth = 15$ 。

从图 6 和表 1 中可以看出, 根据定位误差的标准差对比, 随机森林算法的定位误差的波动性最小, KNN 算法最大。在样本容量逐渐增加的过程中, 随机森林算法定位误差最为稳定。同时, 随机

森林的定位误差最小,性能也最好。

从三种算法运行时间的标准差占平均运行时间的百分比来看,随机森林是 3.67%,wKNN 是

5.31%,KNN 是 9.87%。随机森林的运行时间稳定性要高于 KNN 和 wKNN 算法。

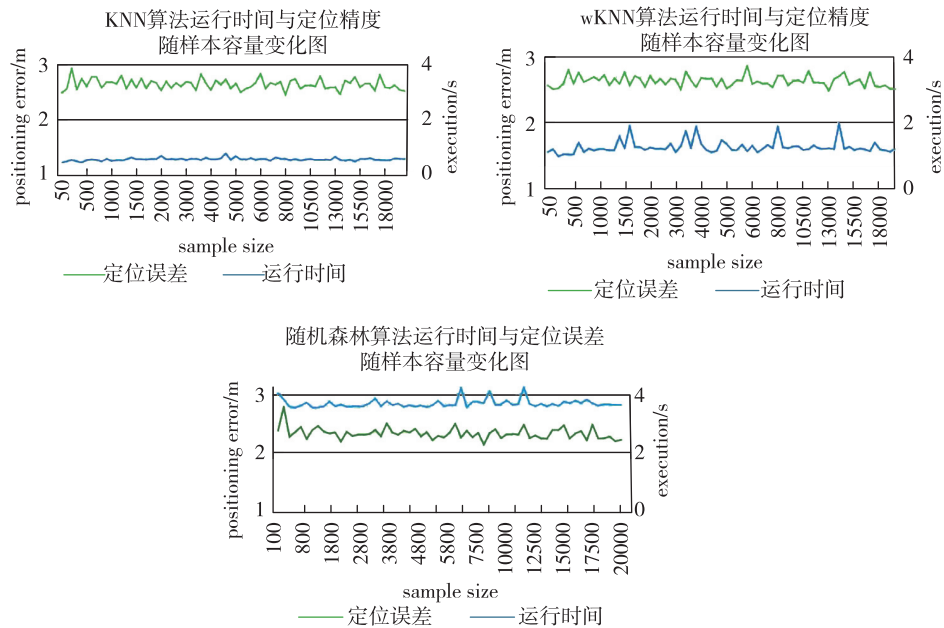


图 6 三种算法运行时间与定位误差随样本容量变化图

Fig. 6 The tendency chart of positioning time and positioning error with sample size of three algorithms

表 1 三种算法评价指标对比

Tab. 1 The comparison of indexes for evaluating three algorithms

	定位时间 平均值/s	定位时间 标准差/s	定位误差 平均值/m	定位误差 标准差/m
KNN	0.496	0.049	2.679	0.152
wKNN	1.261	0.067	2.589	0.101
随机森林	3.867	0.142	2.324	0.079

## 5 结论

1) 本文介绍了 CSI 和指纹定位法的概念,分析了使用 CSI 进行指纹定位的优势,阐述了指纹库的构建过程和定位方法。

2) 分析了 KNN 和 wKNN 算法中关键参数  $K$  对定位误差的影响。结果表明,在  $K$  较小时,增大  $K$  可以使用更多相关邻居来计算定位结果,进而提高定位精度;但过大的  $K$  值会导致原本不相关的邻居加入邻居队列,也会使定位误差增大。

3) 在线样本集容量增加时,测试了 KNN、wKNN 和随机森林算法的运行时间和定位精度的稳定性,结果表明,随机森林算法不仅定位精度最高,而且定位时间与定位精度的稳定性最好。

## 参考文献

- [1] 邓中亮,尹露,唐诗浩,等. 室内定位关键技术综述[J]. 导航定位与授时, 2018, 5(3): 14-23.  
Deng Zhongliang, Yin Lu, Tang Shihao, et al. A survey of key technology for indoor positioning[J]. Navigation Positioning and Timing, 2018, 5(3): 14-23(in Chinese).
- [2] 韩美林,刘有耀. RFID 定位技术和常见算法比较[J]. 产业与科技论坛, 2014, 13(1): 89-91.  
Han Meilin, Liu Youyao. Comparison between RFID positioning technology and common algorithms[J]. Estate and Science Tribune, 2014, 13(1): 89-91(in Chinese).
- [3] 徐驰,邱添,曹广成. 几何估算和 RSSI 相结合的室内定位算法[J]. 计算机应用与软件, 2017, 34(6): 296-302.  
Xu Chi, Qiu Tian, Cao Guangcheng. Indoor location algorithm based on geometrical estimation and RSSI[J]. Computer Applications and Software, 2017, 34(6): 296-302(in Chinese).
- [4] 周启臻,邢建春,杨启亮. 基于信道状态信息相位差的人员入侵检测方法[J]. 传感技术学报, 2018, 31(1): 103-109.

- Zhou Qizhen, Xing Jianchun, Yang Qiliang. Human intrusion detection based on phase difference with channel state information[J]. Chinese Journal of Sensors and Actuators, 2018, 31(1): 103-109 (in Chinese).
- [5] 蒋思琪, 王向勇. 基于信道状态信息的人体行为识别方法[J]. 通信电源技术, 2018, 35(1): 93-96.
- Jiang Siqi, Wang Xiangyong. Human behavior recognition method based on channel state information[J]. Telecom Power Technologies, 2018, 35(1): 93-96(in Chinese).
- [6] Mohammed Abdulaziz Aide Al-qaness, 李方敏, 马小林, 等. 基于信道状态信息的人体行为识别系统[J]. 武汉理工大学学报, 2016, 38(4): 76-80.
- Mohammed Abdulaziz Aide Al-qaness, Li Fangmin, Ma Xiaolin, et al. Human activity recognition system using channel state information of Wi-Fi signals[J]. Journal of Wuhan University of Technology, 2016, 38(4): 76-80(in Chinese).
- [7] Abdel-Nasser H, Samir R, Sabek I, et al. MonoPHY: Mono-stream-based device-free WLAN localization via physical layer information[C]// Proceedings of 2013 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2013: 4546-4551.
- [8] Wu K, Xiao J, Yi Y, et al. CSI-based indoor localization[J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 24(7): 1300-1309.
- [9] Elnahrawy E, Li X, Martin R P. The limits of localization using signal strength: a comparative study [C]// Proceedings of 2004 1<sup>st</sup> Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, IEEE SECON 2004. IEEE, 2004: 406-414.
- [10] Qian K, Wu C, Zhou Z, et al. Inferring motion direction using commodity Wi-Fi for interactive exergames[C]// Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. ACM, 2017: 1961-1972.
- [11] 陈空, 宋春雷, 陈家斌, 等. 基于改进 WKNN 的位置指纹室内定位算法[J]. 导航定位与授时, 2016, 3(4): 58-64.
- Chen Kong, Song Chunlei, Chen Jiabin, et al. An indoor location fingerprint algorithm based on improved WKNN [J]. Navigation Positioning and Timing, 2016, 3(4): 58-64(in Chinese).
- [12] 徐玉滨, 周牧, 刘宁庆, 等. 基于近邻点数优化的 WLAN 室内 KNN 定位方法, CN101883424A [P]. 2010.
- Xu Yubin, Zhou Mu, Liu Ningqing, et al. WLAN indoor KNN positioning method based on neighbor point optimization, CN101883424A [P]. 2010(in Chinese).
- [13] 汤丽, 徐玉滨, 周牧, 等. 基于 K 近邻算法的 WLAN 室内定位技术研究[C]// 2009 国际信息技与应用论坛, 2009.
- Tang Li, Xu Yubin, Zhou Mu, et al. Research on WLAN indoor positioning technology based on K nearest neighbor algorithm [C]// 2009 International Information Technology and Applications Forum, 2009 (in Chinese).
- [14] 陈良明, 韩泽耀. OFDM——第四代移动通信的主流技术[J]. 计算机技术与发展, 2008, 18(3): 184-187.
- Chen Liangming, Han Zeyao. OFDM—main technology of 4G mobile communication system [J]. Computer Technology and Development, 2008, 18(3): 184-187(in Chinese).
- [15] 张晓亮, 赵平, 徐冠青, 等. 基于一种优化的 KNN 算法在室内定位中的应用研究[J]. 电子设计工程, 2013, 21(7): 44-46.
- Zhang Xiaoliang, Zhao Ping, Xu Guanqing, et al. Research of indoor positioning based on a optimization KNN algorithm [J]. Electronic Design Engineering, 2013, 21(7): 44-46(in Chinese).