

一种改进型几何聚类指纹室内定位方法

李石荣,符茂胜,张峰辉,何富贵

(皖西学院电子与信息工程学院,安徽 六安 237012)

摘 要:为解决传统加权 K 最近邻算法(WKNN, Weighting K-Nearest Neighbor)定位方法中选取 K 值存在局限性影响定位精度的问题,提出了一种改进型几何聚类指纹室内定位方法。该方法首先利用网格分布在定位区域构建指纹点几何位置分布,采集指纹点接收信号强度(RSS, Received Signal Strength)和位置信息,建立指纹定位数据库;然后,利用支持向量机分类算法在解决高维度和非线性问题上的优势选取定位点的多个近邻指纹点,根据对定位贡献度的大小筛选近邻指纹点并构建几何聚类定位区域;最后利用 WKNN 算法进行定位。实验结果表明,提出的方法解决了传统 WKNN 方法中多边形定位区域在 K 值选取存在局限性的问题,具有更高的定位精度和工程实用性。

关键词:K 最近邻算法;室内定位;几何聚类;接收信号强度;支持向量机;贡献度

中图分类号:TP391

文献标志码:A

引 言

随着物联网技术的不断应用和发展,基于实时位置服务的需求越来越多^[1]。由于 GPS 系统受到室内障碍物的遮挡使得其不能较好地应用于室内环境,关于室内定位技术的应用和研究成为位置服务的一大热点^[2-3]。目前室内定位技术较多,其中基于 WLAN 的指纹定位方法由于其较低的成本和较高定位精度而得到广泛的研究和应用。基于指纹的定位方法主要分为两个阶段:离线阶段和在线阶段。离线阶段主要通过定位设备采集指纹点的信号强度和位置信息构建数据信息和位置信息之间的映射关系;在线阶段利用映射关系对定位点信号强度信息处理获取定位结果。文献[4]提出一种“动态多雷达搜索策略”自适应选取近邻 K 值,解决 K 值选

取局限性;文献[5]提出了一种动态调整定位字典方法,使模型动态适应环境和接收信号强度(Received Signal Strength, RSS)变化。文献[4-5]解决了室内定位过程中环境变化、位置变化、RSS 变化和 K 值动态选取问题,然而并未考虑指纹分布中几何位置关系。文献[6-10]分别对信号特征性问题提出了相应的解决方法。文献[6]提出了一种利用核判别和混洗蛙跳方法解决信号时变性问题,减小信号冗余和噪声;文献[7]提出了一种利用过滤选取固定的 AP 点的 RSS 进行定位;文献[8]提出了利用 KPCA 算法和 APC 算法进行特征提取和聚类减小冗余并解决非线性问题,进行分类并缩小定位范围;文献[9]提出了利用 RSS 统计特性克服信道干扰问题;文献[10]提出了一种基于相关向量机的非线性偏最小二乘法来解决多径效应和阴影效益问题。文献[6-10]

收稿日期:2019-05-06

基金项目:国家自然科学基金项目(61702375);安徽省自然科学基金项目(1908085MF213);安徽高校自然科学研究重大项目(KJ2015ZD44);皖西学院青年项目(WXZR201806);安徽高校自然科学研究项目(KJ2019A0631)

作者简介:李石荣(1989-),男,安徽六安人,硕士,助教,主要从事信号与信息处理、传感器技术等方面的研究,(E-mail)03000120@wxc.edu.cn

充分考虑信号时变性对定位的影响,提出特征提取、统计分布和聚类等方法解决了信号时变性对定位的影响问题,然而并未考虑到信号时变性是信号传播过程中的固有特性,无法消除由于 AP 设备和复杂的环境影响而产生的部分冗余性和噪声。文献[11-15]利用 SVM(Support Vector Machine)在非线性问题和高维度问题的优势,使用分类或回归算法获取定位结果,然而由于信号时变性的影响可能会产生个别定位点位置误差较大而无法及时发现产生定位误差较大的问题。文献[16-17]分别提出了一种基于几何聚类和多边形的室内定位方法,解决传统加权 K 最近邻算法 WKNN(Weighting K-Nearest Neighbor)存在误差较大点的问题,然而其只考虑到规则的几何聚类多边形,并未考虑复杂条件下(如边界位置附近)几何聚类图形,边界位置附近的定位点误差较大。

WKNN 方法是指纹定位方法中算法较简单、系统较易实现的一种室内定位方法。文献[16]通过在几何网状结构内自适应选取近邻指纹点,构建最佳多边形区域后缩小定位区域范围来提高定位精度,但该模型只考虑了网状几何结构中矩形或正方形定位结构。文献[17]提出一种对定位区域进行网格区域划分后,利用 SVM 多分类将位置确定在投票数最多的四边形区域内并通过 WKNN 进行定位,该算法并未考虑可进一步优化定位区域,且算法运算时间较长。文献[18]通过判断周围信号强度的大小筛选信号较强的前 M 个定位 AP 参与定位并计算结果,该算法并未考虑 WKNN 算法中 K 的优化,应用上具有一定的局限性。文献[4]考虑了在利用 WKNN 定位过程中固定 K 值可能存在的定位局限性,在确定没有新的不确定参数下利用“多雷达搜索策略”确定 K 值获取定位结果,该算法并未考虑定位点在网格边界线上的临界情况;在实际定位过程中,由于定位点可能出现在网格区域边界位置附近,固定多边形形状和 K 值会降低定位效果。

为解决以上问题,本文提出了一种改进型几何聚类指纹室内定位方法。首先,利用网格分布在定位区域内确定定位指纹点几何位置并采集指纹点 RSS 和位置信息,利用 SVM 算法建立指纹 RSS 数据和位置之间的映射关系,构建指纹分类数据库;然后,采集定位点 RSS 信息并利用映射关系分类获取定位点多个近邻指纹点,计

算近邻指纹点对定位点的贡献度大小,对贡献度大小进行排序,选取最佳几何聚类多边形定位区域;最后利用 WKNN 算法获取定位结果。实验结果表明,本文提出的方法可解决传统 WKNN 定位算法固定 K 值和固定多边形区域存在局限性以及选取近邻指纹点存在误差较大影响定位精度的问题,降低了定位误差。

1 改进型 WKNN 定位方法

1.1 改进型 WKNN 定位方法

传统指纹点分布大多数基于网格分布结构,将定位区域划分成矩形等规则的四边形区域结构中,利用算法将定位点确定在某个固定的区域中,通过 WKNN 算法对定位点进行定位。传统方法忽略了定位点在临界边上的情况,影响了定位精度。同时,传统 WKNN 算法中的 K 值的选取也存在较大的局限性。

为解决以上问题,提出了一种基于贡献度的改进型几何聚类定位结构。如图 1 所示,网格上分布圆圈为指纹点,五角星为某个定位测试点。传统 WKNN 定位过程中通常选取指纹点 1、2 和 7 围成的三角形区域或指纹点 1、2、6 和 7 围成的四边形区域。当定位点位置靠近网格线边界位置时,固定区域和 K 值会引起较大定位误差。由于 RSS 的时变性,定位结构可能存在如图 1 所示几种类型的结构情况,通过计算指纹点贡献度可动态选取定位点指纹定位区域。

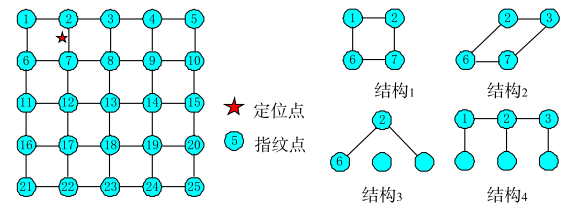


图 1 指纹点位置与定位结构图

1.2 支持向量机

指纹点和定位点接受信号强度数据分别为 $D = \{RSS_1, RSS_2, \dots, RSS_N\}$ 和 RSS , 位置信息 $L = \{L_1, L_2, \dots, L_N\}$, 类别 $y_i \in \{-1, 1\}$; 构造最优线性判别函数:

$$g(RSS) = \omega^T \cdot RSS + b \quad (1)$$

其中: ω 为分类超平面垂直向量, b 为构造最优分类平面常数。

为解决凸函数和非线性问题,引入拉格朗日乘子 $\alpha = \{\alpha_1, \alpha_2, \dots, \alpha_N\}$ 和高斯核函数 $K(\cdot)$ 。构造拉格朗日函数:

$$L(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 +$$

$$\alpha_i \left(\sum_{i=1}^N (1 - y_i \cdot (\omega \cdot RSS_i + b)) \right) \quad (2)$$

将对偶问题转化为求极值问题:

$$\max_{\alpha} \min_{\omega, b} L(\omega, b, \alpha) \quad (3)$$

对 ω, b 求偏导可得:

$$\begin{cases} \frac{\partial L}{\partial \omega} = 0 \\ \frac{\partial L}{\partial b} = 0 \end{cases} \Rightarrow \begin{cases} \omega = \sum_{i=1}^N \alpha_i y_i RSS_i \\ \sum_{i=1}^N \alpha_i y_i = 0 \end{cases} \quad (4)$$

将对 α 的极大值问题转换成对偶问题:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j K(RSS_i \cdot RSS_j) - \sum_{j=1}^N \alpha_j \\ \text{s. t.} \quad & \begin{cases} \sum_{i=1}^N \alpha_i y_i = 0 \\ \alpha_i \geq 0, i = 1, 2, \dots, N \end{cases} \end{aligned} \quad (5)$$

α^* 为对偶问题的解,根据 KKT 条件可以将分类超平面转化为:

$$f(RSS) = \text{sign} \left(\sum_{i=1}^N \alpha_i^* y_i K(RSS \cdot RSS_i) + b^* \right) \quad (6)$$

SVM 可解决指纹点分类过程中高维度和非线性问题^[19]。

1.3 定位贡献度计算

根据 SVM 分类算法获取定位点近邻指纹点,利用欧氏距离计算定位区域上的近邻指纹点对定位点的贡献度,再通过改进 WKNN 算法重新构建定位区域结构,计算定位结果。具体定位贡献度计算过程如下:

(1) 首先利用 SVM 多分类算法获取定位点 M ($M > 6$) 个近邻指纹点,计算定位点 RSS 到所有近邻指纹点的欧氏距离,记为 $d_1, d_2, \dots, d_M$, $d_i = |RSS - RSS_i|, i = 0, 1, \dots, M$ 。

(2) 其次计算近邻指纹点贡献度大小,表示为 $w_1, w_2, \dots, w_M$, 其中, $w_i = \frac{1/d_i}{\sum_{j=1}^M 1/d_j}, i = 0, 1, \dots, M$ 。

(3) 将近邻指纹点贡献度进行排序,判断选取定位结构。主要分为以下几种情况:

(a) 当前 4 个近邻指纹点位置结构满足结构 1 ~ 结构 3 时,利用结构 1 ~ 结构 3 对定位点进行定位。

(b) 当不满足情况 a 时,增加近邻指纹点的选取,

当选取的近邻指纹点中可组成结构 1 ~ 结构 3 位置时,停止选取,构造结构 1 ~ 结构 3 进行定位。

(c) 当近邻指纹点不低于 6 个且可构造结构 4 时,选取结构 4 定位。

(d) 对选取指纹点贡献度进行二次贡献度分配,重现获取对定位的权值。

(f) 当选取的近邻指纹点不能满足结构 1 ~ 结构 4 几何图形时,说明定位点数据误差较大,重新采集定位点数据。

(4) 利用 WKNN 算法对定位点进行定位。

2 定位流程

定位流程图如图 2 所示:

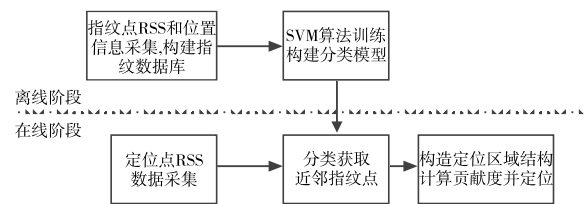


图2 定位流程图

具体定位计算过程如下:

(1) 利用 SVM 获取定位点前 M ($M > 6$) 个近邻指纹点,计算定位点 RSS 到近邻指纹点欧氏距离分别为 $D_1, D_2, \dots, D_M$ 。

$$D_i = |RSS - RSS_i|, i = 0, 1, \dots, M \quad (7)$$

(2) 根据指纹点欧氏距离计算对定位点的贡献度从大到小依次为 $w_1, w_2, \dots, w_M$, 选取前 K ($3 \leq K \leq 6$) 个贡献度较大的指纹点并根据改进 WKNN 方法构建定位结构区域;

$$w_i = \frac{1/D_i}{\sum_{j=1}^M 1/D_j}, i = 0, 1, \dots, M \quad (8)$$

(3) 利用前 K 个指纹点的贡献度和位置信息,计算定位结果 P 。

$$P = \sum_{i=1}^K w_i L_i, i = 0, 1, \dots, K \quad (9)$$

3 实验结果分析

3.1 实验环境

为验证本文提出的改进 WKNN 室内定位方法的效

果,在皖西学院实验实训部一楼实验室进行了实验,实验环境如图 3 所示。实验区域长 36 m、宽 18 m,指纹点之间间隔距离选取为 2 m,测试点位置选取定位区域内,每次数据采集次数为 30 次,取平均值作为 RSS 指纹数据和定位数据。

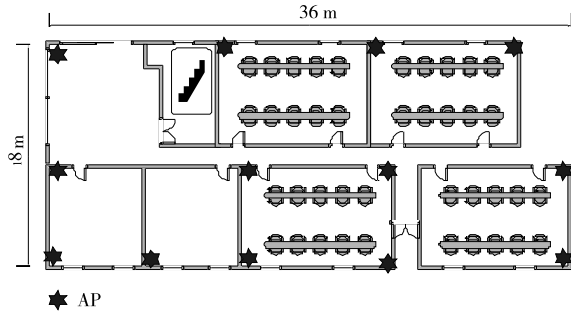


图 3 实验环境

3.2 实验结果分析

3.2.1 实验数据

表 1 和表 2 分别给出的一组指纹点和定位实验数据,分别有 4 个 AP 上的特征信号数据。

表 1 指纹点信号强度数据(单位:dBm)

指纹点	AP1	AP2	AP3	AP4
1	-41	-49	-35	-43
2	-50	-52	-24	-40
3	-47	-43	-24	-38
4	-46	-42	-37	-46
5	-50	-40	-46	-37
6	-44	-40	-42	-18
7	-49	-43	-33	-24
8	-42	-43	-34	-35
9	-39	-40	-33	-42
10	-38	-45	-39	-41
11	-39	-42	-44	-48
12	-39	-43	-41	-42
13	-34	-41	-43	-40
14	-26	-41	-39	-32
15	-21	-40	-45	-26
16	-20	-47	-40	-36
17	-34	-48	-39	-37
18	-40	-42	-33	-25
19	-38	-42	-38	-27
20	-23	-36	-39	-24
21	-28	-39	-41	-38
22	-36	-39	-40	-38
23	-33	-40	-37	-31
24	-41	-41	-44	-22
25	-40	-41	-34	-19

表 2 定位点信号强度数据(单位:dBm)

定位点	AP1	AP2	AP3	AP4
1	-27	-42	-47	-35
2	-34	-38	-40	-32
3	-42	-41	-24	-24
4	-52	-42	-41	-19
5	-42	-44	-39	-40
6	-24	-40	-46	-39
7	-18	-41	-46	-29

3.2.2 不同定位方法结果对比

本文对四种基于 WKNN 算法的不同定位方法的结果进行了对比。定位方法分别为:

- (1) 传统 WKNN 算法:固定 K 值获取定位结果。
- (2) WKNN + 四边形:将定位点固定在某个四边形区域中进行定位。
- (3) WKNN + 几何聚类:利用最佳多边形几何聚类原理动态选取多边形区域和 K 值。

(4) 改进 WKNN 算法:本文提出的利用近邻指纹点对定位的贡献度重构定位区域获取定位结果。

定位结果统计见表 3,对比如图 4 所示。表 3 为一组 7 个定位点数据的多次数据处理定位误差统计结果。图 4 是利用传统 WKNN 算法(固定 K 值为 4)、WKNN + 四边形、WKNN + 几何聚类和改进 WKNN 算法对定位点多次数据处理获取的平均定位误差分别为 1.81 m、1.55 m、1.53 m 和 1.38 m 左右^[16-17]。

表 3 定位误差累计概率统计结果

误差 /m	传统 WKNN /%	WKNN + 四边形/%	WKNN + 几何聚类/%	改进 WKNN /%
0.5	8	13	15	17
1	25	32	33	38
1.5	46	53	54	57
2	65	69	71	74
2.5	78	82	83	87
3	84	93	92	95
3.5	89	95	96	98
4	93	97	98	99
4.5	96	98	99	99
5	98	99	99	99
5.5	99	99	99	99
6	99	99	99	99

3.2.3 K 值对定位结果的影响

本文从不同 K 值对传统 WKNN 算法和改进 WKNN 算法下的定位精度进行了对比。不同的 K 值影响的定位精度的变化如图 5 所示。从图 5 可以看到,当 K 值在 3~6 之间时,改进 WKNN 方法可构建最佳多边形结构

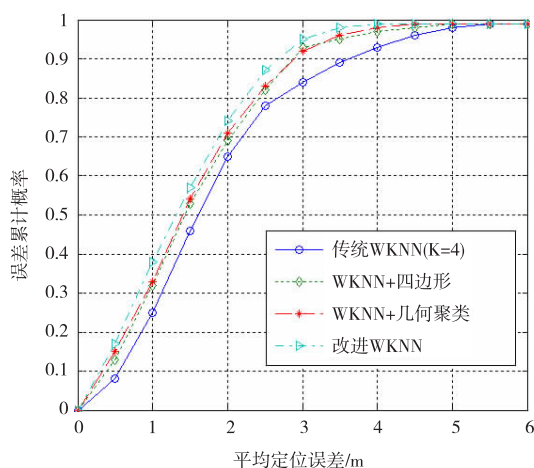


图4 定位结果对比

定位区域,当 K 值超过这个范围时,定位误差较大。传统WKNN算法由于在选择 K 值过程中未对定位结构区域作优化处理,存在误差较大的近邻指纹点使得定位误差较大。当 K 值取3、4、5和6时,与传统WKNN方法相比下平均定位误差分别减小了0.57 m、0.45 m、0.56 m和0.17 m。

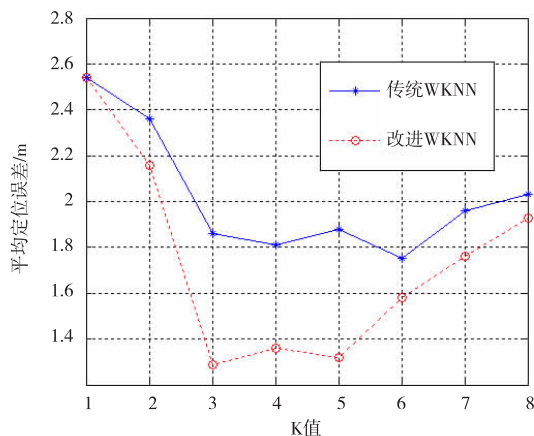


图5 不同K值对定位误差影响

3.2.4 边界近邻定位点测试

在网格分布临界边上选取7个定位点进行测试,经过多次反复的采集数据、处理数据和计算并获取定位结果。7个不同位置的定位点在传统WKNN算法和改进WKNN算法下定位误差的对比如图6所示。从图6可以看出,当定位点在临界边附近时,传统WKNN算法定位误差较大,而本文提出的改进WKNN方法可较好地适应边界附近定位点。

4 结束语

本文针对传统WKNN算法在定位过程中由于 K 值

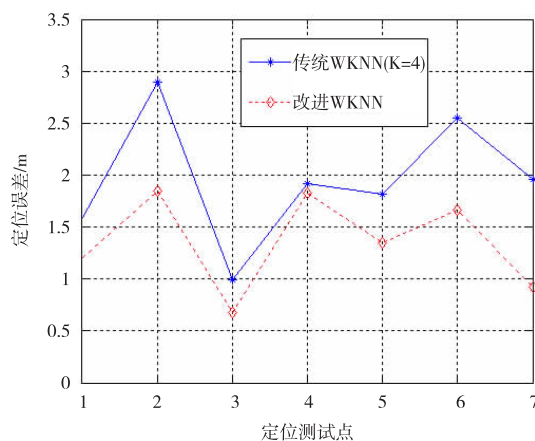


图6 临界定位点定位结果对比

和定位结构区域选取过程中存在误差的问题,提出了一种改进型WKNN室内定位方法。该方法首先构建最佳多边形定位结构区域,接着利用SVM分类获取近邻指纹点并根据指纹点的贡献度大小构建最佳定位区域,最后利用WKNN算法获取定位结果。实验结果表明,本文提出的方法优于传统基于WKNN的定位方法。本文在研究过程中并未考虑定位点移动的情况,关于定位点移动条件下如何解决信号时变性的问题有待下一步研究。

参考文献:

- [1] FENG C, AU W S A, VALAEE S, et al. Received-Signal-Strength based indoor positioning using compressive sensing[J]. IEEE Transactions on Mobile Computing, 2012, 11(12): 1983-1993.
- [2] GAO Y B, LIU S F, ATIA M M, et al. INS/GPS/LiDAR integrated navigation system for urban and indoor environments using hybrid scan matching algorithm[J]. Sensors, 2015, 15(9): 23286-23302.
- [3] ZHANG K, SHEN C, ZHOU Q, et al. A combined GPS UWB and MARG locationing algorithm for indoor and outdoor mixed scenario[J]. Cluster Computing, 2018(4): 1-10.
- [4] 胡久松, 刘宏立, 徐琨. 自适应动态K的WKNN室内定位方法[J]. 重庆邮电大学学报: 自然科学版, 2018, 30(4): 431-438.
- [5] 王婷婷, 柯炜, 孙超. 自适应环境变化的RSS室内定位方法[J]. 通信学报, 2014, 35(10): 210-217.
- [6] 张勇, 李飞腾, 王昱洁. 基于KDDA和SFLA-LSSVR

- 算法的 WLAN 室内定位算法[J].计算机研究与发展,2017,54(5):979-985.
- [7] 徐潇潇,谢林柏,彭力.基于 WiFi 信号强度特征的室内定位系统设计[J].计算机工程,2015,41(4):87-91.
- [8] LUO J H, FU L. A smartphone indoor localization algorithm based on WLAN location fingerprinting with feature extraction and clustering [J]. sensors, 2017, 17 (6): 1339-1350.
- [9] HUSEN M N, LEE S. Indoor location sensing with invariant Wi-Fi received signal strength fingerprinting[J]. Sensors, 2016, 16(11):1898-1908.
- [10] CHEN C, WANG Y J, ZHANG Y, et al. Indoor positioning algorithm based on nonlinear PLS integrated with RVM[J]. IEEE Sensors Journal, 2017, 18(2):660-668.
- [11] ZHANG S, GUO J M, WANG W, et al. Indoor 2.5D positioning of WiFi based on SVM[C]//Proceedings of 2018 Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS), Wuhan, China, March 22-23, 2018:1-7.
- [12] 韦燕华,周彦,王冬丽.基于 LS-SVM 的位置指纹室内定位[J].计算机工程与应用,2016,52(9):122-125.
- [13] PEI L, LIU J B, GUINNESS R, et al. Using LS-SVM based motion recognition for smartphone indoor wireless positioning[J]. Sensors, 2012, 12(5):6155-6175.
- [14] 姚英彪,毛伟勇,姚瑞丽,等.基于改进支持向量回归的室内定位算法[J].仪器仪表学报,2017,38(9):2112-2119.
- [15] 朱宇佳,邓中亮,刘文龙,等.基于支持向量机多分类的室内定位系统[J].计算机科学,2012,39(4):32-35.
- [16] 刘春燕,王坚.基于几何聚类指纹库的约束 KNN 室内定位模型[J].武汉大学学报:信息科学版,2014,39(11):1287-1292.
- [17] 张勇,徐小龙,徐科宇.基于加权质心法的 WLAN 室内定位系统[J].电子测量与仪器学报,2015,29(7):1036-1041.
- [18] 王磊,周慧,蒋国平,等.基于 WiFi 的自适应匹配预处理 WKNN 算法[J].信号处理,2015,31(9):1067-1074.
- [19] 邓乃扬,田英杰.支持向量机:理论、算法与拓展[M].北京:科学出版社,2009.

An Improved Indoor Positioning Method Based on Geometric Clustering Fingerprinting

LI Shirong, FU Maosheng, ZHANG Fenghui, HE Fugui

(School of Electronics and Information Engineering, West Anhui University, Lu'an 237012, China)

Abstract: In order to solve the problem of limitation about how to choose the value K of Weighting K-Nearest Neighbor algorithm (WKNN, Weighting K-Nearest Neighbor) in indoor positioning, an improved indoor positioning method based on geometric clustering fingerprinting is proposed. Firstly, the grid distribution is used to establish the geometric distribution of fingerprint points, and the data of the Received Signal Strength (RSS) and location can be collected, then the positioning database can be built. Secondly, the support vector machine algorithm is used to obtain the neighboring fingerprint points about positioning point due to the advantages in solving the problems of nonlinearity and high dimensions, and the geometric clustering positioning area can be constructed by the neighboring fingerprint points according to the contribution to positioning point. Finally, the results can be got by WKNN algorithm. The experimental results show that the proposed method solved the problem of limitation in selecting the value K of WKNN algorithm and polygonal positioning area, it has higher precision and better engineering practicality than traditional methods.

Key words: K-Nearest Neighbor algorithm; indoor positioning; geometric clustering; Received Signal Strength (RSS); Support Vector Machine (SVM); contribution