

一种利用 RFID 和块聚类的室内定位方法

张文杰¹, 董育宁¹, 王新珩²

ZHANG Wenjie¹, DONG Yuning¹, WANG Xinheng²

1.南京邮电大学 通信与信息工程学院, 南京 210003

2.西苏格兰大学 计算机系, 佩斯利, 苏格兰, 英国

1.College of Telecommunications & Information Engineering, Nanjing University of Post and Telecommunications, Nanjing 210003, China

2.School of Computing, University of the West of Scotland, Paisley, Scotland, UK

ZHANG Wenjie, DONG Yuning, WANG Xinheng. Indoor positioning method using RFID and block clustering. Computer Engineering and Applications, 2016, 52(10): 112-117.

Abstract: The indoor positioning technology based on fingerprint attracts the attentions from many researchers. RFID (RadioFrequency Identification) technology is more attractive, due to its high accuracy and adaptability to different environment. However, because of the heavy consumption of the computing power, it limits the applications in practice. A new hybrid Kmeans and Weighted K-nearest Neighbor method is proposed and applied in real-world indoor positioning. The new method divided the mapping area into several classes based on a clustering method. A matching into class is done first and then location is determined. The result shows that the proposed method reduces the accumulated errors and thus reduces the computational power whist maintains reasonable accuracy.

Key words: Radio Frequency Identification(RFID); fingerprinting; indoor_positioning; kmeans; Weighted k-Nearest Neighbor

摘 要:基于指纹的RFID室内定位技术,由于其定位精度高、普适性强等优点受到国内外学者的广泛关注。但因为其计算量较大,在实际应用仍然非常有限。提出基于实际应用场景的Kmeans和Weighted K-Nearest Neighbor(WKNN)联合的定位方法,将指纹地图通过聚类算法分成块,先初步确认待测点所属指纹块,在块的基础上定位,这样可以减小误差累计。仿真结果表明,该方法在保证适当定位精度的同时,也减少了计算量和在线定位时间。

关键词:无线射频定位技术(RFID); 指纹; 室内定位; kmeans; 加权K近邻

文献标志码:A **中图分类号:**TP391 **doi:**10.3778/j.issn.1002-8331.1407-0064

1 引言

随着无线传感网络的发展,无线室内定位技术越来越多地受到国内外的重视。RFID(Radio Frequency Identification)技术由于其成本低廉,系统部署简单而在室内定位和跟踪中得到广泛的应用^[1]。目前,基于RFID的室内定位跟踪方法主要有以下几种:RSSI(Received

Signal Strength Indicator)定位方法依据传播模型中信号强度损耗和距离的关系获得标签与阅读器之间的距离,利用几何方法得到标签的坐标。其中LANDMARC是典型的基于RSSI的系统^[2]。TOA(Time Of Arrival)定位方法和RSSI定位方法原理相同,其标签与阅读器的距离是根据信号传播时间计算得到^[3]。TDOA(Time

基金项目:国家自然科学基金(No.61271233, No.60972038);教育部博士点基金(No.20103223110001)。

作者简介:张文杰(1985—),女,硕士研究生,主要研究领域为网络流分类与识别、室内网络定位, E-mail: wenjie8529@126.com; 董育宁(1955—),男,博士研究生,教授,主要研究领域为下一代泛在异构无线网络中的多媒体通信、QoE/QoS建模、跨层优化、绿色通信、网络流分类与识别、图像和视频信息处理等方面的关键技术及其应用;王新珩(1968—),男,博士研究生,教授,主要研究领域为物联网、无线Mesh网络、传感器网络、智能认知网络、室内网络定位和云计算在无线网络中的应用。

收稿日期:2014-07-06 **修回日期:**2014-09-29 **文章编号:**1002-8331(2016)10-0112-06

CNKI网络优先出版:2015-02-16, <http://www.cnki.net/kcms/detail/11.2127.TP.20150216.1644.010.html>

Difference Of Arrival)定位方法是对TOA的改进,根据同一标签发射的信号到达不同阅读器的时间差计算标签与阅读器的距离^[3]。AOA(Angle Of Arrival)定位方法利用标签天线的方向性来判定标签的位置,不同阅读器收到同一标签的具有方向性的信号,方向直线的交点即标签的位置^[4]。基于指纹(Fingerprinting)定位方法首先建立场景的特征指纹库,将待测位置的信号特征与指纹库特征进行比较匹配来估计待测位置的坐标。由于室内环境复杂多变,多径传播和环境随机噪声的干扰较大,基于指纹定位方法利用多径传播来构建位置信息,较其他方法更适合室内环境,并且硬件成本比较低。

基于指纹的定位方法分为两个阶段:离线训练阶段和在线定位阶段。关于在线定位算法国内外学者已经做了很多研究。比如:概率匹配法^[5],最大似然法^[6],朴素贝叶斯法^[7],以上方法是基于概率的方法,依据WiFi等信号的RSS具有高斯分布特性,计算得到后验概率来定位。另外一类是直接基于RSS来定位。比如:最近邻(NN, Nearest Neighbor)^[8],K近邻法(KNN, K Nearest Neighbor)^[9-10],加权K近邻(WKNN, Weighted K Nearest Neighbor)^[11]。

最近邻是指纹定位最基本的算法,比较待测点RSS特征集与指纹库参考点RSS特征的距离,取距离最近的参考点坐标做为估计结果。K近邻法是取k个距离最近的参考点坐标平均的结果做为定位结果。加权k近邻,是对最近邻的改进,增加了权重因子,来衡量不同近邻点对测试点定位的贡献程度。

为解决KNN和WKNN算法计算量大、定位效率低的问题,本文提出一种利用Active RFID Tag的Kmeans和WKNN联合定位的方法。离线训练阶段对采样数据进行初步处理,利用残差去除粗大误差。利用Kmeans对指纹地图进行块划分,得出块中心。在线定位阶段,首先将待测点进行块级别定位,在块范围内进行最终定位。定位过程采用两次加权,计算距离时采用0、1加权,只取检测到的RFID的RSS,计算定位结果是采用基于距离的加权。最后仿真结果表明,本文提出的方法能够在保证一定定位精度的前提下减小定位计算量。

2 基于指纹定位方法简介

基于位置指纹的定位通常分为两个阶段:离线阶段和在线阶段。

离线阶段的主要工作是建立位置指纹数据库。在定位服务区域内确定参考点(Reference Point, RP),在每个参考点上采集来自各个AP(Access Point)的信号强度样本,存入数据库(Radio Map)。

在线阶段,用户在待测点接收到各AP信号实时采

样,根据一定的搜索匹配算法将待测点信号特征与数据库中的特征进行匹配,得出待测点的位置。图1是文中利用RFID和块聚类的定位系统图。

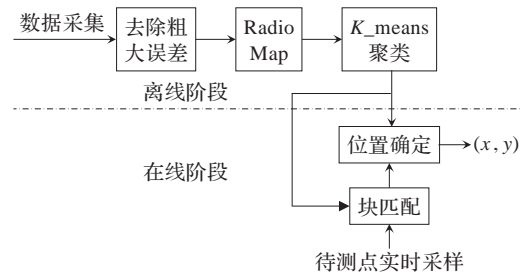


图1 位置指纹定位过程示意图

3 方法

3.1 利用残差去除粗大误差

测量或采样不可避免地存在测量误差。对于粗大误差能依判定准则发现,并对测量或采样数据异常值消除。本文采用 3σ 准则剔除参考点RSS采样的异常值。 3σ 准则定义为:信号采样值为 $X=\{x_1, x_2, \dots, x_n\}$,根据文献[12]残差的定义为:

$$v_i = x_i - \bar{x}, \text{ 其中 } \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

由贝塞尔公式可得采样值的均方根误差:

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n v_i^2} \quad (2)$$

若 $|v_i| > 3\sigma$,则采样值 x_i 视为含有粗大误差而剔除。

本文采用 3σ 准则剔除参考点RSS采样的异常值。

3.2 Kmeans聚类算法

Kmeans聚类算法主要用于处理数据的聚类问题,广泛应用于数据挖掘模式识别,机器学习等领域。Kmeans聚类是一种动态聚类算法^[13],基于聚类准则函数最小化的原则,通过迭代将数据划分到不同的类中,使生成的类尽可能地紧凑和独立。

Kmeans聚类算法的聚类准则是待聚类数据与聚类中心的距离最小准则。假设待聚类数据与第 $j(j=1, 2, \dots, k)$ 个聚类中心的距离最小,则该数据划分到第 j 类。

Kmeans聚类算法迭代步骤:

输入:信号采样数据集 $X=\{X_1, X_2, \dots, X_N\}$

输出: k 个类 $C_j(1 \leq j \leq k, k < N)$

步骤1 初始化各聚类参数

(1)设置聚类数 $k(2 \leq k \leq N)$

(2)设置聚类中心 $m_1(1), m_2(1), \dots, m_k(1)$

(3)设置迭代停止条件 ε

步骤2 对信号采样数据集 $X=\{X_1, X_2, \dots, X_N\}$ 进行

分类

(1) 计算 $X_p(p=1, 2, \dots, N)$ 与各聚类中心的距离 $d_i = \|X_p - m_i(h)\| (i=1, 2, \dots, k)$, $m_i(h)$ 为第 h 次迭代的聚类中心

若满足 $\min(d_i) = \|X_p - m_j(h)\| (1 \leq j \leq k)$, 则判定 $X_p \in C_j(h)$

(2) 计算新的聚类中心 $m_j(h+1) = \frac{1}{N_j} \sum_{X_i \in C_j} X_i$,

其中第 j 个类 C_j 中包含 N_j 个向量。

步骤3 迭代停止条件验证

(1) 若 $\|m_j(h+1) - m_j(h)\| \geq \varepsilon$, 返回步骤2, 用新聚类中心重新分类;

(2) 否则, 停止计算。

本文采用 Kmeans 聚类算法, 使在线定位计算在块级别内完成, 可以减小计算量, 提高定位效率。

3.3 WKNN 算法

WKNN 算法是一种基于距离的匹配算法, 通过权因子来分配 K 个近邻对最终结果的影响比重。计算步骤如下:

步骤1 $D_p = (\sum_{i=1}^n a_i |s_i - x_{ip}|^q)^{\frac{1}{q}}$ (3)

其中, s_i 是待测点的第 i 个 Tag 的信号采样值, x_{ip} 是指纹库里第 p 个点的第 i 个 Tag 的信号采样值。考虑到 RF 信号的时变特性, 为减小误差, 计算距离时只取当前检测到的 Tag 的信号采样值, 因此增加一个取值为 0 或 1 的因子 a_i 。若 $S_i = 0$ 则 a_i 取 0, 若 $S_i \neq 0$, a_i 取 1。

当 $q=1$, D_p 为曼哈顿距离; 当 $q=2$, D_p 为欧式几何距离。

步骤2 $w_i = \frac{1/D_p^2}{\sum_{p=1}^k 1/D_p^2}$ (4)

步骤3 计算定位坐标

$(x, y) = \sum_{i=1}^K w_i(x_i, y_i)$ (5)

其中, w_i 为权因子。本文采用两次加权, 一次在计算距离时利用 a_i 加权, 第二次是在确定定位结果时利用 w_i 加权。这样可以减小由于环境变化引起采样波动造成的误差, 提高定位的精度。

4 实验和结果分析

4.1 实验环境搭建和数据采集

本实验数据采集是在无锡知谷网络科技有限公司的实验室完成的, 该实验室模拟候机室环境, 有商店、餐厅、书店、陶瓷展厅等, 真实呈现了室内的复杂环境。如图2所示。



图2 实验环境

实验选取 $10\text{ m} \times 10\text{ m}$ 的区域, 区域里包含有墙壁、桌子、木板等。部署 6 个 RFID Tag, Tag 部署位置如图3所示。考虑到信号在室内传播的特性规律, 为高度逼近真实应用环境。实验没有采用传播模型法建立特征指纹库, 而是以 0.5 m 为间距人工采集 441 个参考点的 RSS, 采样频率为 10 Hz 。

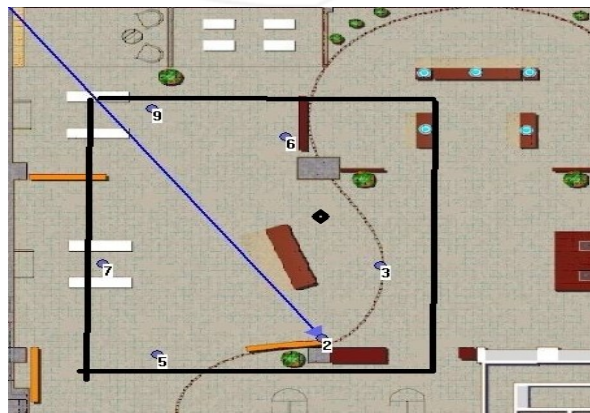


图3 RFID Tag 的部署

对每个参考点采集的数据进行预处理, 首先对于同一 Tag 的 RSS 进行剔除粗大误差处理, 然后求其平均值做为指纹特征。Tag6 的信号谱图如图4所示。

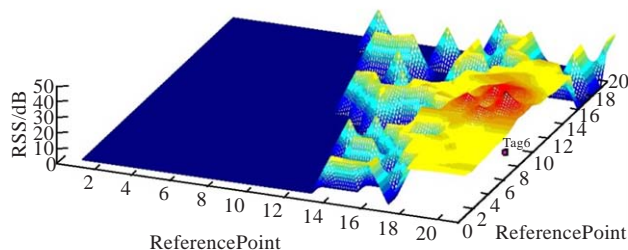


图4 Tag6 的信号强度谱分布图

4.2 离线训练

为使在线定位阶段效率更快, 在数据预处理完成的基础上, 对指纹地图进行块级别分类。采用 Kmeans 聚类算法 ($K=6$), 根据 Tag 的数量, 将指纹地图分为 6 块, 聚类中心为 $C1, C2, C3, C4, C5, C6$ 。实验采取的 Kmeans

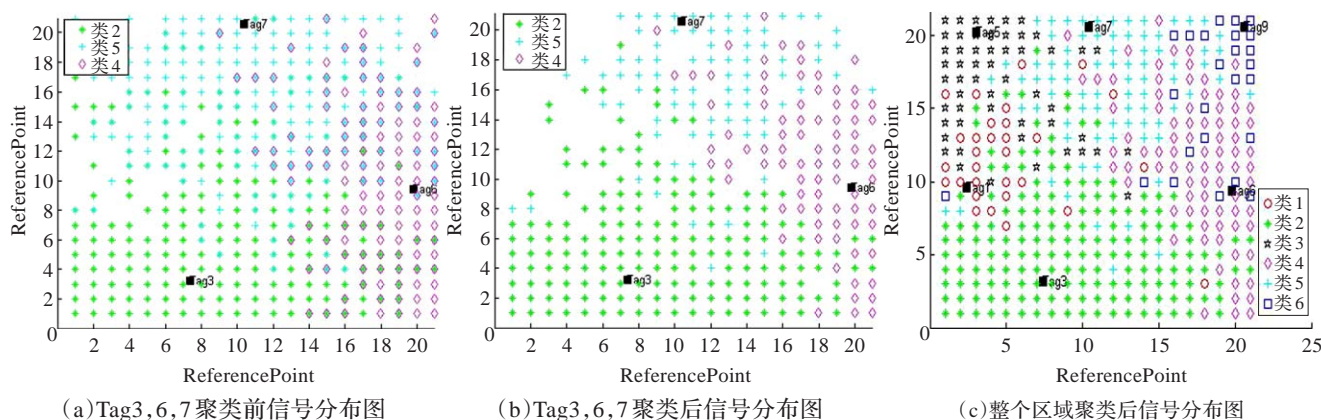


图5 利用 Kmeans($K=6$) 指纹地图聚类

表1 利用不同块地图定位误差和定位时间比较

	1块	2块	3块	4块	5块	6块(全图)
最大定位误差/m	2.33	1.50	1.21	0.70	0.41	0.25
最小定位误差/m	1.14	0.80	0.87	0.45	0.11	0.00
定位时间/s	0.001 8	0.003 5	0.009 4	0.012 4	0.013 2	0.014 4

聚类算法选取初始聚类中心,不是随机选取。根据无线信号强度普分布特点(如图4所示),以发射源为中心,物理位置离发射源越近,对应的RSS信号越强。并且物理空间相近的参考点,其RSS信号特征集也相似。因此选取离RFID Tag最近坐标对应的RSS信号特征向量做为聚类中心,对数据库进行聚类。聚类所需时间平均为0.059 2 s。聚类结果如图5所示。

从图5可以看到,聚类前后信号分布特点是一致的,以Tag为中心呈扇形分布。

4.3 在线定位

为减小计算量并保证定位精度,在线定位阶段做了如下优化。首先分别比较待测点与六块指纹地图中心的距离,根据表1取距离最近的2块指纹地图的参考点做为定位的数据库。从表1中看出利用4块和5块地图定位误差较小,但定位时间和取整个地图一样长。利用3块地图定位最大误差没有比利用2块地图定位有显著减小,且定位时间较长。利用1块地图定位,定位时间较短,但定位误差较大。

其次,计算待测点和参考点距离时采用0、1加权,计算定位坐标时采用基于距离加权的KNN。最近邻的个数 k 取值是影响定位精度和定位计算量的重要因素。在定位前通过实验确定 k 的取值。图6展示了最近邻个数 k 的不同取值对定位误差的影响。从图中可以看到平均定位误差在 $k=1$ 到 $k=5$ 之间快速下降,在 $k=5$ 之后下降缓慢,趋于平稳。因此实验中选取 $k=5$ 。

4.4 结果和分析

图7展示了几种定位算法分别应用的定位结果和误差对比。

从图7中看到KNN定位最大误差2.5 m,最小误差

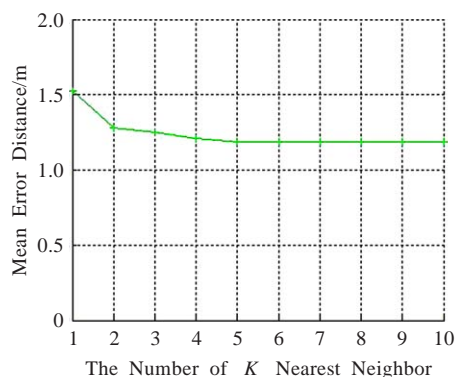


图6 不同近邻个数定位误差的比较

1.3 m;0、1加权KNN定位误差1.9 m,最小误差0.6 m;本文提出的KmeansWKNN定位最大定位误差1.5 m,最小误差0.6 m。KmeansWKNN定位误差在1.5 m以内,WKNN的定位精度最高,最大定位误差为0.25 m,但计算量最大。WKNN定位一次需要与指纹库全部参考点(441个)进行比较计算,平均定位时间为0.014 4 s。KmeansWKNN定位一次只需与指纹库1/3的参考点(约150个),平均定位时间为0.003 5 s。同时,与文献[14]中提出的SVM-C算法和文献[15]中Maximum Likelihood算法进行了性能比较。比较结果如图8所示。

从图8中可以看出,Maximum Likelihood定位误差最大,而且平均定位时间较长,为0.019 2 s;Kmeans-WKNN(本文方法)和SVM-C算法相比定位误差较小,SVM-C的平均定位时间为0.017 3 s,是Kmeans-WKNN的五倍。因此,本文提出的方法相比较于传统KNN和WKNN定位算法,在保证适当定位精度的前提下,平均定位时间可以减少3/4。相比较于其他现有算法,定位时间也有显著减小。

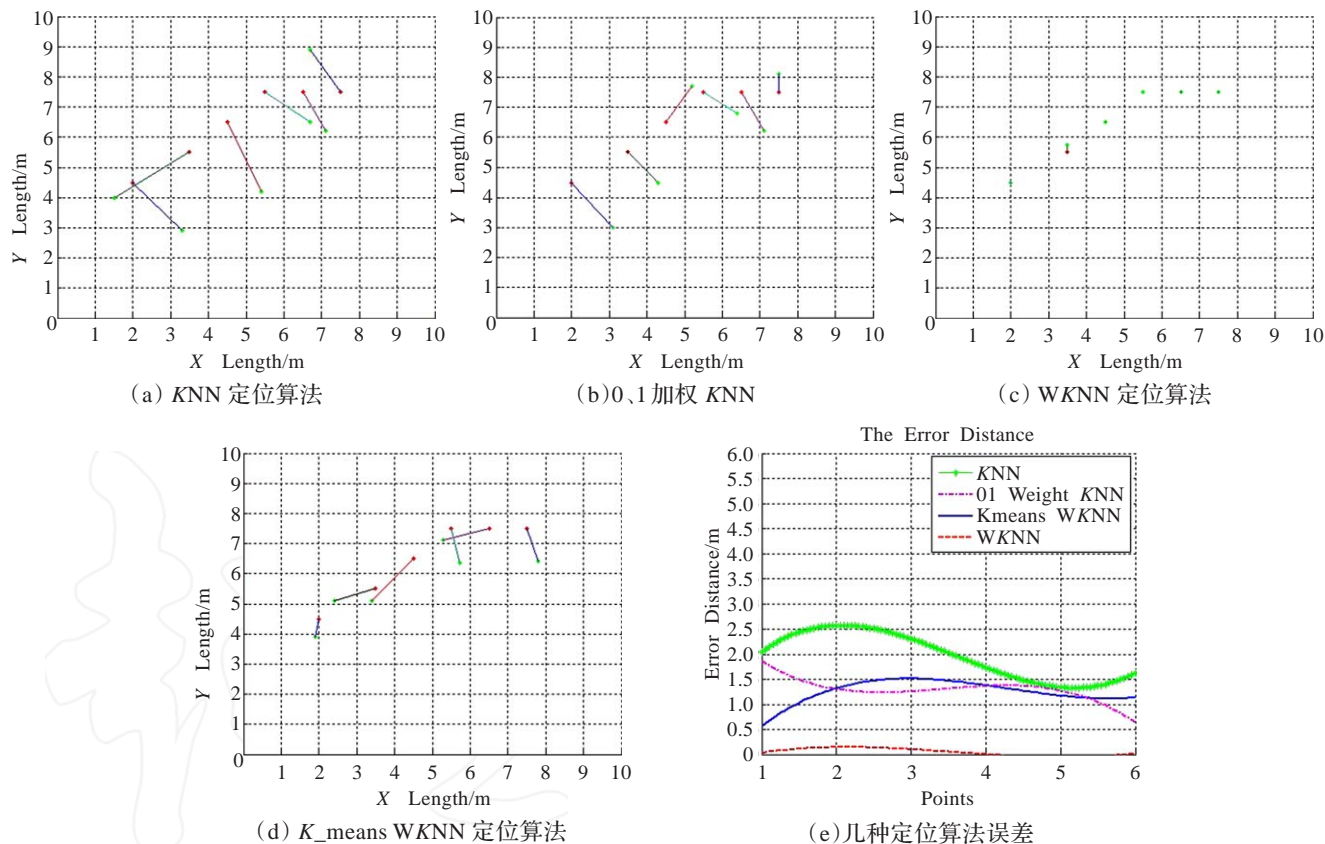


图7 几种定位算法定位结果和误差比较

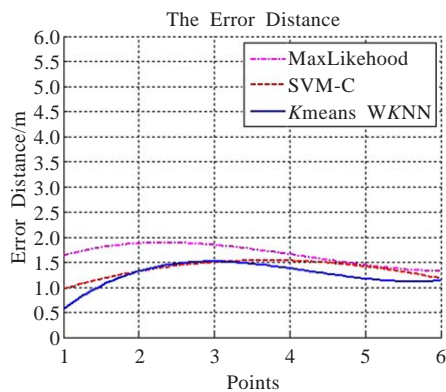


图8 定位算法精度比较

5 结束语

在基于RSS特征指纹的KNN算法基础上,本文提出了Kmeans和WKNN结合的定位方法,实验证明,本文提出的方法在保证一定定位精度的前提下,显著减低了计算量,在一定程度上克服了指纹定位耗时的缺点。对于像机场等室内面积很大的场所,指纹定位需要庞大数据库,本文提出的方法据有一定实际应用意义。未来工作将考虑在小计算量下,通过减小多径效应和信道衰落的影响,进一步提高定位精度。

参考文献:

[1] Rahmati A, Zhong L, Hiltunen M, et al. Reliability tech-

niques for RFID-based object tracking applications[C]// 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks. IEEE, 2007: 113-118.

[2] Yinggang X, Zhiliang W, Shanshan Z. On RFID positioning base on LANDMARC and improved algorithm[C]// Proceedings of the 29th Chinese Control Conference, 2010: 4831-4836.

[3] Liu H, Darabi H, Banerjee P, et al. Survey of wireless indoor positioning techniques and systems[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2007, 37(6): 1067-1080.

[4] Zhou J, Zhang H, Mo L. Two-dimension localization of passive RFID tags using AOA estimation[C]// Instrumentation and Measurement Technology Conference (I2MTC). IEEE, 2011: 1-5.

[5] Jaffre T, Grigis P M, Papanastasiou S, et al. On the efficacy of WiFi indoor positioning in a practical setting[C]// 2013 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2013: 699-704.

[6] Pritt N. Indoor positioning with maximum likelihood classification of Wi-Fi signals[C]// 2013 IEEE Symposium on Sensors. IEEE, 2013: 1-4.

[7] Castro P, Chiu P, Kremenek T, et al. A probabilistic room location service for wireless networked environments[C]// Ubicomp 2001: Ubiquitous Computing. Berlin Heidelberg:

- Springer, 2001:18-34.
- [8] Tran Q, Tantra J W, Foh C H, et al. Wireless indoor positioning system with enhanced nearest neighbors in signal space algorithm[C]//Vehicular Technology Conference, IEEE 64th 2006.IEEE, 2006:1-5.
- [9] Lin T N, Lin P C. Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks[C]//2005 International Conference on Wireless Networks, Communications and Mobile Computing. IEEE, 2005, 2:1569-1574.
- [10] Ma J, Li X, Tao X, et al. Cluster filtered KNN: A WLAN-based indoor positioning scheme[C]//2008 International Symposium on World of Wireless, Mobile and Multimedia Networks. IEEE, 2008:1-8.
- [11] Ni L M, Liu Y, Lau Y C, et al. LANDMARC: indoor location sensing using active RFID[J]. Wireless Networks, 2004, 10(6):701-710.
- [12] 费业泰.误差理论与数据处理[M].6版.北京:机械工业出版社, 1987:44-51.
- [13] Li M J, Ng M K, Cheung Y, et al. Agglomerative fuzzy k-means clustering algorithm with selection of number of clusters[J]. IEEE Transactions on Knowledge and Data Engineering, 2008, 20(11):1519-1534.
- [14] Lee C W, Lin T N, Fang S H, et al. A novel clustering-based approach of indoor location fingerprinting[C]//2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC). IEEE, 2013:3191-3196.
- [15] Pritt N. Indoor positioning with maximum likelihood classification of WiFi signals[C]//2013 IEEE on Sensors. IEEE, 2013:1-4.

(上接58页)

- [5] Pawlak Z. Rough sets[J]. International Journal of Computer & Information Sciences, 1982, 11(5):341-356.
- [6] Lingras P, West C. Interval set clustering of web users with rough k-means[J]. Journal of Intelligent Information Systems, 2004, 23(1):5-16.
- [7] Mitra S, Banka H, Pedrycz W. Rough-fuzzy collaborative clustering[J]. IEEE Transactions on Systems, Man, and Cybernetics: Part B, 2006, 36(4):795-805.
- [8] Maji P, Pal S K. Rough set based generalized fuzzy c-means algorithm and quantitative indices[J]. IEEE Transactions on Systems, Man, and Cybernetics: Part B, 2007, 37(6):1529-1540.
- [9] 楼晓俊, 李隽颖, 刘海涛. 距离修正的模糊C均值聚类算法[J]. 计算机应用, 2012, 32(3):646-648.
- [10] 王建锋, 金健, 王晶晶. 一种具有影响力因子的硬聚类算法[J]. 计算机工程与应用, 2009, 45(19):177-180.
- [11] Cover T, Hart P. Nearest neighbor pattern classification[J]. IEEE Transactions on Information Theory, 1967, 13(1):21-27.
- [12] 朱剑英. 应用模糊聚类法应注意的若干关键问题[J]. 模糊系统与数学, 1987, 1(1):104-111.
- [13] 王学恩, 韩德强, 韩崇昭. 采用不确定性度量的粗糙模糊C均值聚类参数获取方法[J]. 西安交通大学学报, 2013, 47(6):55-60.
- [14] 郭晋华, 苗夺谦, 周杰. 基于阴影集的粗糙聚类阈值选择[J]. 计算机科学, 2011, 38(10):209-210.
- [15] Davies D L, Bouldin D W. A cluster separation measure[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1979, 1(2):224-227.

(上接103页)

- [8] Yao Yuan, Rao Lei, Liu Xue. Performance and reliability analysis of IEEE 802.11p safety communication in a highway environment[J]. IEEE Transactions on Vehicular Technology, 2013, 62(9):4198-4212.
- [9] Sheu T L, Wu Y J. Jamming-based medium access control with dynamic priority adjustment in wireless Ad-hoc networks[C]//IET Communications, 2007, 1(1):34-40.
- [10] Sahebgharani S S M. A scheduling algorithm for downloading data from RSU using multicast technique[C]//Ninth International Conference on Information Technology: New Generations (ITNG), 2012:809-814.
- [11] 孙健. 车辆自组织网络中安全消息的优先级调度和拥塞控制[D]. 西安:西安电子科技大学, 2012.
- [12] Wang C D, Thompson J P. Apparatus and method for motion detection and tracking of objects in a region for collision avoidance utilizing a real-time adaptive probabilistic neural network[P]. US Patent, 1997:3-18.
- [13] Willke T L, Tientrakool P, Maxemchuk N F. A survey of inter-vehicle communication protocols and their applications[J]. Communications Surveys & Tutorials, 2009, 11(2):3-20.
- [14] 朱钧宇, 黄传河, 徐利亚, 等. VANET中一种链路稳定度的计算方法[J]. 计算机工程与科学, 2013, 35(1):57-60.
- [15] Liu Jianhang, Ge Yuming, Bi Jingping, et al. Dynamic optimization model for cooperative downloading strategy of VANET[J]. Journal of Internet Technology, 2013, 14(6):963-972.