

DOI:10.19431/j.cnki.1673-0062.2018.03.011

WiFi 的室内定位技术的改进研究

何志爽, 万亚平*, 赵庆, 王明艳

(南华大学 计算机学院, 湖南 衡阳 421001)

摘要: 现行传统 WiFi (wireless fidelity) 接收信号强度指示 RSSI (receivd signal strength inication) 的位置指纹室内定位技术存在定位误差大、稳定性差的缺陷. 因此, 我们对原有的 K 最近邻 KNN (K-Nearest neighbor) 算法提出了改进的方案. 同时, 在原有的 KNN 算法的基础上提出了融合朴素贝叶斯概率算法的新算法-BKWNN (Bayes K-Nearest weighted neighbor) 算法. 通过仿真实验的结果表明: 在相同的实验环境下, BKWNN 算法显著地提高了室内定位的精确度, BKWNN 算法相比于原来其它常用的指纹匹配算法具有更高的稳定性.

关键词: 室内定位; 指纹匹配; 朴素贝叶斯概率算法; BKWNN; 比重系数

中图分类号: TP391 **文献标志码:** B **文章编号:** 1673-0062(2018)03-0060-05

Research and Improvement of Indoor Positioning Technology Based on WiFi

HE Zhishuang, WAN Yaping*, ZHAO Qing, WANG Mingyan

(Computer School, University of South China, Hengyang, Hunan 421001, China)

Abstract: The current traditional WiFi (Wireless Fidelity) signal intensity indication RSSI (Receivd Signal Strength Inication) location fingerprint localization technology has the defects of large positioning error and poor stability. In this regard, it proposes an improved scheme for the original K nearest neighbor KNN (K-Nearest Neighbor) algorithm, and on the basis of the original KNN algorithm, a new algorithm, -BKWNN (Bayes K-Nearest Weighted Neighbor), is proposed, which combines the simple Bias probability algorithm (Bayes K-Nearest Weighted Neighbor). The simulation results show that in the same experimental environment, the BKWNN algorithm significantly improves the accuracy of the indoor location, while the BKWNN algorithm has better stability than the other common fingerprint matching algorithms.

key words: Indoor location; fingerprint matching; naive bayes probability algorithm;

收稿日期: 2018-02-23

作者简介: 何志爽 (1989-), 男, 硕士研究生, 主要从事于大数据与舆情分析方面的研究. E-mail: 1290665530@qq.com.

* 通信作者: 万亚平 (1973-), 男, 教授, 博士, 主要从事于分布式计算与网络存储信息检索方面的研究. E-mail: 512828758@qq.com

BKWNN; specific gravity coefficient

0 引言

全球卫星定位系统 GPS^[1] 具有全天候、覆盖广、高效益等特点,对于各类复杂的室内环境, GPS 的定位精度受到很多因素的干扰,定位的性能会受到严重的制约.因此,人们一直在探索新的室内定位技术.近年来,基于 WiFi 的位置指纹定位系统^[2] 具有部署简单,功率低,定位精度高,并且非视距^[3] (Not line of sight, NLOS) 对 WiFi 影响较弱等优点,成为室内定位技术的首选.目前,基于 RSSI^[4] 的位置指纹定位技术实现起来简单,技术成熟,因而它已经成为应用最为广泛的定位技术.

1 基于 RSSI 位置指纹定位模型的原理

基于 RSSI 的位置指纹定位技术主要是分为离线阶段和在线阶段.离线阶段主要是在定位区间部署 AP 信号接入点,选取合适的采样点,通过移动终端设备在采样点 AP (access point) 采集信号.将从各个 AP 信号接入点采集到的信号值和对应点的 MAC(x, y) 地址值保存到后台数据库,建立关于 RSSI 和采样点位置映射的指纹数据库.

在线定位阶段是利用位置指纹匹配算法将移动终端设备在目标定位点采集到的信号强度值与后台数据库中的位置指纹数据进行匹配,找出最相近的一个或者一组数据,利用找出来的这些位置指纹数据所对应的样本点的坐标估算目标定位点的位置.在实时定位阶段常用的指纹匹配算法有 K 最近邻算法、加权 K 最近邻 WKNN (K-Nearest weighted neighbor) 算法^[5].

本文提出对 KNN 指纹匹配算法的改进措施,提出了新的加权 K 近邻算法—BKWNN 算法,通过仿真试验证明,新的指纹匹配算法在指纹定位技术中具有更高的更好的性能.

2 近邻算法(KNN)的原理

KNN 算法是找出与目标定位点最相似的采样点,通过这个采样点的位置坐标,对目标定位点的位置进行估算.通常情况下,使用欧式距离^[6] 来衡量相似度.KNN 算法定位的实现原理如下:

1) 在室内采集 q 个样本点的位置指纹数据以

及它们分别对应的位置坐标 L . 采样点的信号向量 $\mathbf{R}_i = (rss_1^i, rss_2^i, \dots, rss_n^i)$, R_i 表示第 i 个采样点采集到的 n 个 AP 的信号强度.将定位区域中数据 R_i 对应的坐标表示为 $L_i = (x_i, y_i)$.

2) 在定位点采集信号向量为 $S = (s_1, s_2, \dots, s_n)$.

3) 计算 S 与指纹库中的所有样本数据之间的距离为:

$$d(S, R) = \sqrt{\sum_{j=1}^n (s_j - r_j^i)^2} \quad (1)$$

式(1)中 i 为采样点, j 为维度.

4) 通过对 3) 中计算得出结果依据大小进行排序,选取前 k 个最小的对应的位置指纹,最后估计待定位点的位置坐标为:

$$(x, y) = \frac{1}{k} \sum_{i=1}^k (x_i, y_i) \quad (2)$$

3 BWKNN 算法的原理

欧式距离在衡量不同点之间的绝对距离时,各个维度特征值的差异对欧式距离的影响是一样的,即在加权求和的过程中,直接将每个维度的特征差异 $(s_j - r_j^i)^2$ 叠加.对于 AP 接入点部署不均匀的场景,简单的加权求和不能准确的反映不同维度的 rss_i 值对 $d(S, R_i)$ 的影响情况(实际的物理场景不可能是所有的维度之间两两垂直).在实际定位的过程中直接使用欧式距离容易将相似度较高的样本点滤掉而产生误差.为了克服简单欧式距离的缺点,本文提出在传统欧式距离中融合不同维度的比重系数,为采样点的不同维度分配不同的比重系数.比重的大小是依赖于单一维度距离 $\sqrt{(s_j - r_j)^2}$ 的倒数占有所有维度距离倒数总和的比例.

改进后的欧式距离公式为

$$d(S, R) = \sqrt{\sum_{j=1}^n (s_j - r_j^i)^2 \times w_j^i} \quad (3)$$

式(3)中 i 为采样点, j 为维度, w_j^i 为第 i 个采样点的 j 维上对应的比重系数. w_j^i 的大小与采样点和目标定位点在 j 维上的距离相关联.目标定位点与采样点在维度上的距离越大,分配的 w_j^i 越大.同时, w_j^i 的大小也与采样点和目标定位点在所有维度上距离的总和相关.所有维度上距离的总和越大,分配的 w_j^i 越大.通过归纳总结并且对表达式

经过归一化处理,最终得到的 w_j^i 表达式为:

$$w_j^i = \frac{\sum_{k=1}^n \frac{1}{\sqrt{(s_k - rss_k^i)^2 + \varepsilon}}}{\frac{1}{\sqrt{(s_j - rss_j^i)^2 + \varepsilon}}} \quad (4)$$

式(4)中 j 为第 j 个维度, s_j 为第 j 个维度的信号强度值。

通过改进后的欧式距离得到每个采样点与目标定位点之间的距离 $d(S, F_i)$, 选取前 k 个最小的对应的位置坐标估算目标定位点的位置坐标为:

$$(\bar{x}, \bar{y}) = \sum_{i=1}^k \lambda_i (x_i, y_i) \quad (5)$$

在确定目标定位点的位置时, KNN 算法只是简单的对坐标求平均值. 为了区分这 k 个采样点与目标定位点的相似度差异 WKNN 算法是为筛选出来的采样点依据 $d(S, F_i)$ 的大小设置比重 λ .

WKNN 采用的比重计算公式为:

$$\lambda_i = \frac{\frac{1}{\text{dist}(S, F_i) + \varepsilon}}{\sum_{j=1}^k \frac{1}{\text{dist}(S, F_j) + \varepsilon}} \quad (6)$$

WKNN 算法通过为 k 个采样点设置比重系数, 比重系数是通过数学公式计算出来的. 同时, 数学公式并不能够适用于所有的室内场景. 因此, WKNN 算法忽略了真实的室内环境, 通过该方法定位出来的结果必然是存在较大的误差. 所以, 提出在 KNN 算法的基础上融合朴素贝叶斯概率算法进行改进, 将此法称为 BKWNN 算法.

BKWNN 算法的基本原理是在筛选到的 k 个采样点的基础上, 使用朴素贝叶斯概率算法来实现最终的定位. 朴素贝叶斯概率模型是一种基于概率统计的定位算法^[6]. 在实际定位过程中, 通过信号传播模型, 采用贝叶斯准则计算信号强度 S 的条件概率分布 $p(L_i | S)$, 并将信号强度 S 的条件概率分布作为目标定位点的位置分布情况. 通过求出条件概率分布的期望值即可得到目标定位点的位置为:

$$(\hat{x}, \hat{y}) = \sum_{i=1}^k p(L_i | S) \times (x_i, y_i) \quad (7)$$

$p(L_i, S)$ 是与实际的环境和定位区间的指纹信号分布情况直接相关联的, 同时由于这 k 个采样点与目标定位点具有很高的相似度, 贝叶斯概率算法模型结合室内场景实现定位, 定位的结果很接近真实的值.

根据贝叶斯全概率联合估计^[7]的原理可知

$$p(L_i | S) = \frac{p(S | L_i) \times p(L_i)}{\sum_{m \in K} p(S | L_m) \times p(L_m)} \quad (8)$$

在(8)式中 $p(L_i)$ 代表 S 出现在指纹参考点 $L_i = (x_i, y_i)$ 位置上的后验概率^[9]. 对于用户有限次的请求而言, 可以将其看做常量, $p(L_i)$ 与 $p(L_m)$ 相等, 得到表达式为:

$$p(L_i | S) = \frac{p(S | L_i)}{\sum_{m \in K} p(S | L_m)} \quad (9)$$

根据每个信号接收端接收到来自各个指纹参考点的位置指纹信息都是相互独立的, 可以知道 $p(S | L_i)$ 的表达式

$$p(S | L_i) = p(s_1 | L_i) p(s_2 | L_i) \cdots p(s_n | L_i) \quad (10)$$

根据 WiFi 信号在室内分布的情况和统计学的原理可知, 信号接收端在有效的范围内接收到的 WiFi 信号强度在室内近似服从高斯正态分布^[9], 可推出以下式子

$$p(s_i | L_i) = \frac{1}{\sqrt{2\pi} * \delta} \exp\left[-\frac{(s_i - \mu)^2}{2\delta^2}\right] \quad (11)$$

最后, 估计出待定位点的位置坐标为

$$(\hat{x}, \hat{y}) = \sum_{i=1}^k p(L_i | S) * (x_i, y_i) \quad (12)$$

4 实验仿真以及结论

4.1 实验环境

选取会议室作为实验场景, 实验场景轮廓图如图 1 所示.

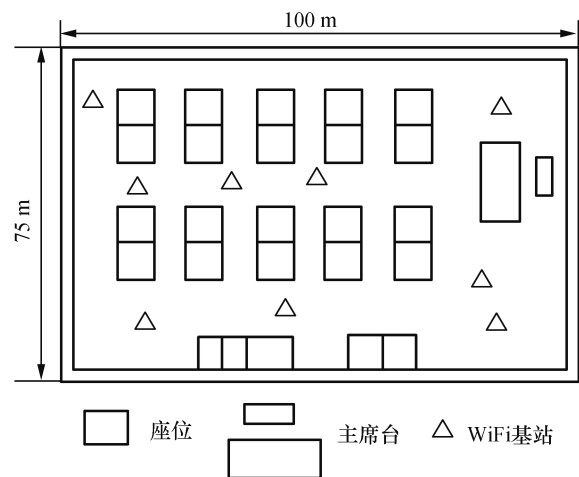


图 1 会议室轮廓

Fig.1 Conference room profile

将会议室的左下方墙角作为坐标原点,同时对会议室区域进行网格化,建立平面直角坐标系.在实验过程中,一共部署了 12 个 WiFi 信号发射点,同时设置了 35 个离线数据采集点.离线阶段对 AP 数据进行采集,建立指纹数据库,在线定位阶段随机选择地点实施定位.

4.2 实验结果分析与总结

为了验证改进后的算法是否会提升定位的精确度,使用了 Matlab 软件进行仿真和分析.分析实验结果,得到误差距离和累积误差累积概率二维分布,如图 2 所示.

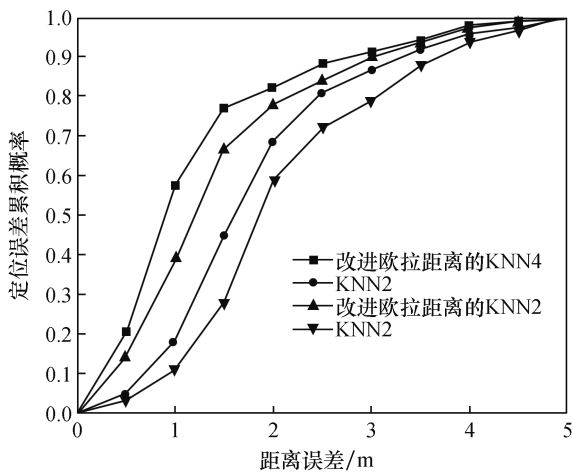


图 2 KNN 算法和改进欧式距离的 KNN 算法的累积误差累积概率分布图

Fig.2 The cumulative error probability distribution diagram of the KNN algorithm and KNN algorithm of improving the euclidean distance

从仿真图 2 可以看出,改进欧式距离的 KNN 算法的平均误差远小于传统的 KNN 算法.在 1.5 m 误差的范围内,改进欧式距离的 KNN 算法的定位精度几乎是传统 KNN 算法的两倍.

取 K 个样本点的基础上,采用朴素贝叶斯概率算法进行定位,估算出待测位置.误差距离和累积误差函数概率二维分布如图 3 所示.

根据不同的算法多次实验获取定位结果,取平均值列于表 1 中.

根据不同的算法多次实验获取定位结果进行误差分析列于表 2.

从误差距离和累积误差累积概率分布图和算法定位误差对比表可以看出,BWKNN 算法的平均误差远小于 WKNN 算法,BWKNN 算法显著的提升了定位的精确度,具有更好的定位性能.

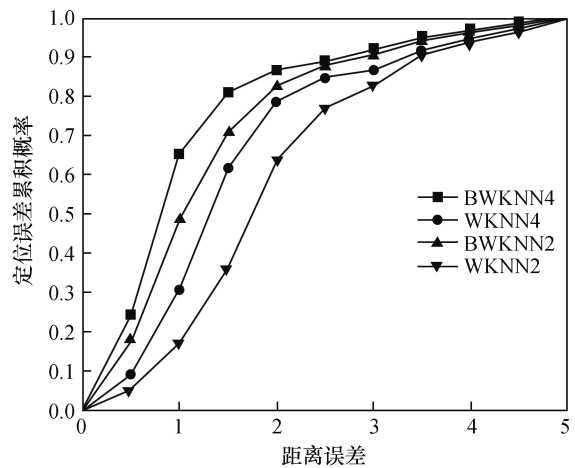


图 3 WKNN 算法和 BWKNN 算法的累积误差累积概率分布图

Fig.3 The cumulative error probability distribution diagram of WKNN algorithm and BWKNN algorithm

表 1 不同定位算法产生的定位结果对比

Table 1 Comparison of location results produced by different location algorithms

定位算法	定位坐标	真实坐标
BWKNN4	(49.2,50.3)	(49.4,50.4)
BWKNN2	(51.0,51.3)	
WKNN4	(48.5,48.5)	
WKNN2	(47.3,48.2)	

表 2 不同定位算法产生的误差对比

Table 2 Error comparison produced by different localization algorithms

定位算法	最大	最小	平均	标准差
BWKNN4	1.14	0.12	0.18	0.69
BWKNN2	1.86	0.89	1.65	0.92
WKNN4	3.18	1.84	2.47	1.26
WKNN2	4.38	1.99	3.19	1.89

5 结 论

文本阐述了基于 WiFi 的室内定位技术的相关理论,分析了当前室内定位技术存在的各种问题,进而提出了一种改进 KNN 算法的方案.同时,本文提出了在改进后的 KNN 算法的基础上融合朴素贝叶斯概率定位算法的 BKWNN 算法.在 KNN 算法的改进方案中,主要是在相似度的匹配阶段对欧式距离进行改进,提高相似度匹配算法

的匹配精度. BKWNN 算法的主要过程是首先运用改进后的 KNN 算法筛选出个采样点, 然后利用朴素贝叶斯概率算法确定最后的位置坐标. 实验表明: 改进后的欧拉公式提高了相似度匹配的准确性; 同时, BKWNN 算法将室内定位的精确度提高了 16.7%, 而且定位的结果中体现出了较好的稳定性.

参考文献:

- [1] LI X, GE M, DAI X, et al. Accuracy and reliability of multi-GNSS real-time precise positioning: GPS, GLONASS, BeiDou, and Galileo [J]. *Journal of geodesy*, 2015, 89(6): 607-635.
- [2] WU Z, JEDARI E, MUSCEDERE R, et al. Improved particle filter based on wlan RSSI fingerprinting and smart sensors for indoor localization [J]. *Computer communications*, 2016, 83: 64-71.
- [3] 李石荣, 李飞腾. 基于 RSSI 概率统计分布的室内定位方法 [J]. *计算机工程与应用*, 2016, 52(11): 119-124.
- [4] WU D, XU Y, MA L. Research on RSS based indoor location method [C]// *Knowledge engineering and software engineering*, 2009. KESE'09. Perth; Pacific-asia conference on IEEE, 2009: 205-208.
- [5] LI J, FU J Q, LI A, et al. An improved WKNN indoor fingerprinting positioning algorithm based on adaptive hierarchical clustering [C]// *Advanced computational methods in life system modeling and simulation*. Singapore: ICSEE 2017, LSMS 2017. *Communications in computer and information science*, 2017: 253-262.
- [6] WANG L, ZHANG Y. On the euclidean distance of images [J]. *IEEE transactions on pattern analysis & machine intelligence*, 2005, 27(8): 1334-1339.
- [7] 李拥军, 曾标, 徐克付, 等. 复杂背景下基于贝叶斯-全概率联合估计的前景检测 [J]. *电子与信息学报*, 2012, 34(2): 388-392.
- [8] 李滔, 王俊普, 吴秀清, 等. 后验概率估计及其应用: 基于核 Logistic 回归的方法 [J]. *模式识别与人工智能*, 2006, 19(6): 689-695.
- [9] 靳超, 邱冬炜. 基于 WiFi 信号室内定位技术的研究 [J]. *测绘通报*, 2017(5): 21-25.
- (责任编辑: 龙威)
-
- (上接第 26 页)
- [21] PIALLA D, TENCHINE D, LI S, et al. Overview of the system alone and system/cfd coupled calculations of the phenix natural circulation test within the thins project [J]. *Nuclear engineering and design*, 2015, 290: 78-86.
- [22] LIU X J, CHENG X. Sub-channel/system coupled code development and its application to scwr-fqt loop [J]. *Nuclear engineering and design*, 2015, 285: 39-47.
- [23] BUCHHOLZ S, CRON D V D, SCHAFFRATH A. System code improvements for modelling passive safety systems and their validation [J]. *Kerntechnik*, 2016, 81(5): 535-542.
- [24] WANG G, GU Z, WANG Z, et al. Verification of neutronics and thermal-hydraulics coupled simulation program ntc by the PDS-XADS transient simulation [J]. *Progress in nuclear energy*, 2015, 85: 659-667.
- [25] CHOI C, HA K, JEONG H Y. Analyses of UTOP events for the design of control rod stop system in PGSFR using MARS-LMR [J]. *Annals of nuclear energy*, 2016, 96: 422-431.
- [26] LORENZI S, CAMMI A, BORTOT S, et al. Analytical models for a small lfr core dynamics studies [J]. *Nuclear engineering and design*, 2013, 254(702): 67-88.
- [27] GU Z, WANG G, WANG Z, et al. Safety analyses of 10 mwth forced circulation lbe-cooled fast reactor during unprotected transients [J]. *Annals of nuclear energy*, 2016, 88: 112-117.
- [28] TIAN Y H, SU G H, WANG J, et al. Code development and safety analyses for pb-bi-cooled direct contact boiling water fast reactor (pbwfr) [J]. *Progress in nuclear energy*, 2013, 68: 177-187.
- [29] ERIKSSON M, CAHALAN J E. Inherent shutdown capabilities in accelerator-driven systems [J]. *Annals of nuclear energy*, 2002, 29(14): 1689-1706.
- [30] SATO I, TOBITA Y, KONISHI K, et al. Safety strategy of jsfr eliminating severe recriticality events and establishing in-vessel retention in the core disruptive accident [J]. *Journal of nuclear science and technology*, 2011, 48(4): 556-566.
- (责任编辑: 龙威)