

A Design and Implementation of Indoor Positioning Algorithm Based on Wi-Fi

Jiandong Qiu, Ni Ke, Lijian Zhuang, Jiaxian Liang

Shenzhen Urban Transport Planning Center Co., Ltd., Shenzhen Guangdong
Email: qiujd@sutpc.com

Received: May 27th, 2019; accepted: Jun. 11th, 2019; published: Jun. 18th, 2019

Abstract

In order to obtain accurate location information in multi-floor and large area indoor environment, based on the open data set of Tampere University in Finland, a fingerprint algorithm model based on Wi-Fi position is established. The automatic encoder model and neural network model are combined with training to extract the active reduction and principal components, and then the improved floor classification model is constructed by classification training. The accuracy rate of floor classification using the improved floor classification model is 92%, which is better than that of automatic encoder model and neural network model. The location accuracy of neural network model combined with weighted KNN algorithm is about 4 meters, and the efficiency of the algorithm is higher than that of the traditional Wi-Fi fingerprint indoor location algorithm. The improved floor classification model combined with weighted KNN algorithm can be used as the best model for large indoor positioning.

Keywords

Indoor Positioning, Wi-Fi Fingerprinting, KNN Algorithm, Machine Learning, Neural Network

一种基于Wi-Fi的室内定位算法研究与实现

丘建栋, 柯 尼, 庄立坚, 梁嘉贤

深圳市城市交通规划设计研究中心有限公司, 广东 深圳
Email: qiujd@sutpc.com

收稿日期: 2019年5月27日; 录用日期: 2019年6月11日; 发布日期: 2019年6月18日

摘 要

为实现在多楼层大面积的室内环境下获得精准定位信息, 以芬兰Tampere大学的公开数据集为基础, 建

立基于Wi-Fi位置指纹算法模型。使用自动编码器模型和神经网络模型结合训练进行主动降维和主成分提取,进而采用分类训练构建改良后的楼层分类模型。使用改良后楼层分类模型楼层分类精确率达到92%,结果优于自动编码器模型和神经网络模型;神经网络模型结合加权KNN算法实现定位精度在4米左右,算法有效性高于传统Wi-Fi指纹室内定位算法。改良后楼层分类模型结合加权KNN算法,可作为大型室内定位最佳模型。

关键词

室内定位, Wi-Fi指纹, KNN算法, 机器学习, 神经网络

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

伴随着社会的进步和数据业务的快速发展,地理位置信息在人们日常生活中发挥着越来越重要的作用,因此人们对于定位与导航的需求不断增长。近年来,由于互联网和物联网技术的广泛应用,以及智能手机和其他无线设备在市场的广泛普及,各种基于位置的服务应用不断受到关注,其中高精度的定位技术是位置服务应用的核心技术。迄今为止全球定位系统(GPS)是世界上应用最多的导航和定位系统,它的使用让人们能在广阔的室外环境迅速获取可靠的、高精度的位置信息,但是当GPS信号遇到厚的障碍物时,信号会在障碍物表面反射,穿过障碍物的信号强度会大幅的衰减,所以GPS很难满足室内环境下精准定位的需求。特别对于大型室内导航需求已经成为亟待解决的问题。

当前基于Wi-Fi的定位是室内导航研究的热点[1][2],特性是不需要布设其余定位设备,使用成本低。针对基于Wi-Fi的定位算法,Bahl等[3]提出使用KNN算法进行精准定位,定位中使用最小二乘法找到相邻最优位置,取最优结果作为定位结果。但单独使用该算法进行大型室内精准定位有个缺陷,即取前 k 个位置进行位置估算,会累积增加定位误差;Brunato等[4]通过机器学习训练模型的方法解决室内定位,但该方法训练模型使用时间过长以及对样本数据要求较高。基于上述学者的研究思路及存在的问题,本文提出同样基于Wi-Fi信号RSS,首先对实验数据进行预处理,构建基础指纹库,然后对自动编码器和神经网络模型结合形成楼层分类模型进行训练,最后利用RSS相关性,使用加权KNN算法对单个楼层的坐标进行估计,结合楼层分类模型得到精准楼层分类的单个楼层定位。实验结果表明,使用加权KNN算法结合楼层分类模型相对于日常室内定位方法更能够获取到高精度的定位。

2. 数据来源

数据使用的是2017年1月至8月期间在芬兰Tampere大学[5]的一栋4层教学楼中采集,该楼总共部署了992个信号发射源,采集数据包括两组Wi-Fi数据。其中一组数据用于室内定位模型的训练,另一组数据用于模型的测试,总共包含了687个训练指纹和3951个测试指纹数据。因此,我们所用到的数据主要有四种:

- 1) 训练指纹RSS: 一个 697×992 的矩阵;
- 2) 训练指纹坐标: 一个 697×3 的矩阵(x,y,z的3D坐标);
- 3) 测试指纹RSS: 一个 3951×992 的矩阵;
- 4) 测试指纹坐标: 一个 3951×3 的矩阵(x,y,z的3D坐标)。

3. 研究方法

3.1. 三角算法

在基于测距的定位算法中，三角算法[6] [7]是比较简单的算法，算法原理为：平面上有三个不共线的基站 A, B, C, 和一个未知终端 D, 并已测出三个基站到终端 D 的距离分别为 R1, R2, R3, 则以三个基站坐标为圆心，三基站到未知终端距离为半径可以画出三个相交的圆，如图 1 所示，未知节点坐标即为三圆相交点。

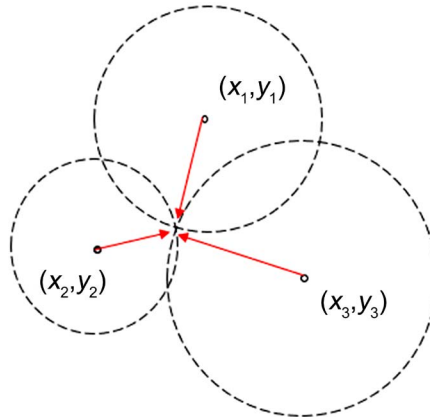


Figure 1. Principle of triangulation algorithm
图 1. 三角算法定位原理

然而，在实际测量中，往往由于测量的误差，使三个圆并不交于一点，而相交于一块区域，如图 2 所示。在此种情况下，便需用其他方法进行求解，如极大似然估计法，最小二乘法进行估计，或者使用三角形质心算法。

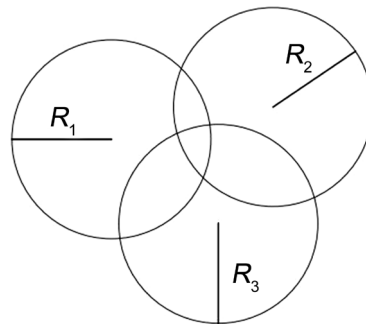


Figure 2. Actual measurement error
图 2. 实际测量误差

本文仅讲述采用最小二乘法求近似解，并针对 n 个基站($n \geq 3$)，已知 n 个基站的坐标分别为 (x_1, y_1) , (x_2, y_2) , ..., (x_n, y_n) , 未知终端坐标为 (X, Y) , 由以下步骤求解:

1) 建立信标节点与未知节点距离方程组

$$\begin{cases} (x_1 - x)^2 + (y_1 - y)^2 = d_1^2 \\ \vdots \\ (x_n - x)^2 + (y_n - y)^2 = d_n^2 \end{cases}$$

2) 上边方程组为非线性方程组,用方程组中前 $n-1$ 个方程减去第 n 个方程后,得到线性化的方程:

$$A * X = b$$

其中:

$$A = \begin{bmatrix} 2(x_1 - x_n) & \cdots & 2(y_1 - y_n) \\ \vdots & \ddots & \vdots \\ 2(x_{n-1} - x_n) & \cdots & 2(y_{n-1} - y_n) \end{bmatrix}, b = \begin{bmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 + d_n^2 - d_1^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 + d_n^2 - d_{n-1}^2 \end{bmatrix}$$

3) 用最小二乘法求解上边方程得:

$$X = (A^T A)^{-1} A^T b$$

X 便是未知终端的坐标计算值。

3.2. 位置指纹算法

位置指纹法[6][8]将环境的位置与特定的指纹联系起来,指纹中存储一种或多种特征,当实际定位中,通过实际获得的多个特征与指纹中的特征匹配,从而获得适配的指纹坐标从而定位,在室内定位中,用得最多的特征也就是 RSSI 定位。

指纹定位分为两个阶段:离线阶段和在线阶段。离线阶段将每个位置的坐标记录以及保存特征(RSSI),最后收集完所有指纹得到指纹库。在线阶段则通过实际的特征查询指纹库获得定位,由图 3 所示。

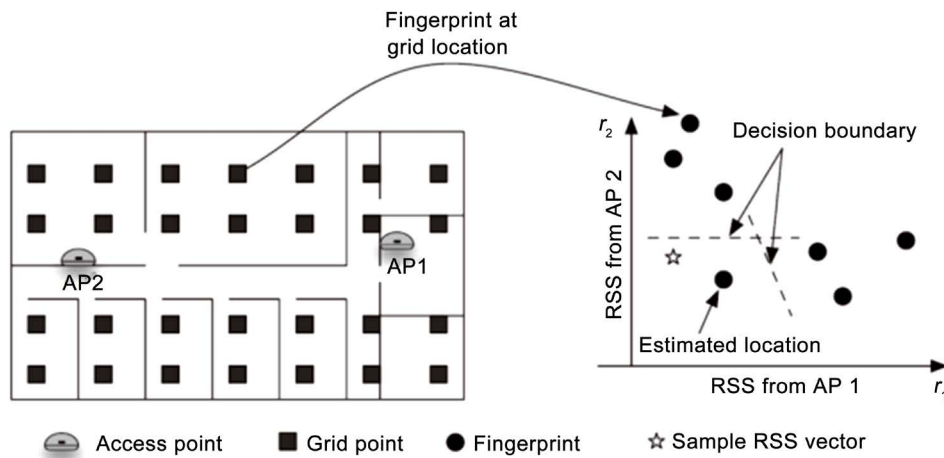


Figure 3. Location principle of position fingerprint algorithm

图 3. 位置指纹算法定位原理

然而,实际的特征中不可能刚好和指纹库中特征相同,这时就需要一定的算法得到最佳适配的指纹,常用得指纹匹配算法有近邻算法(NN)、KNN、WKNN 和神经网络等,这里介绍简单的 KNN 算法。假设 RSSI 向量为 r 、指纹向量为 ρ ,位置指纹是 N 维的,也就是说有 N 个可见的信号发射源, M 个网格点,这样指纹库里面有 M 组指纹数据,每组指纹数据有 N 个 RSSI 数据。 r 和 ρ ($\rho \in M$) 的欧氏距离定义为:

$$D = \sqrt{\sum_{i=1}^N |r_i - \rho_i|^2}$$

在指纹库中的 M 个指纹向量中,通过计算指纹特征距离找到与观测向量差距最小的 K (K 为人为设置的参数)个指纹向量,然后将这些向量对应坐标位置的平均值作为定位点的位置。

4. 基于 Wi-Fi 指纹的室内定位算法实现

位置指纹算法中的 KNN 算法在机器学习是懒散的学习方法,每次估算新的观测点坐标时都要在原有指纹数据库中找寻最邻近的指纹数据进行匹配从而估算出观测点的坐标值。小量的数据对于算法计算影响不大,但是如果遇到大范围的室内环境需要庞大的数据来支撑算法准确度,那么使用该算法就会产生较大的计算量和过高的运行时间成本,从而无法做到移动终端设备实时定位,因此本文提出了一种精准而快速的基于 Wi-Fi 指纹的室内定位算法。与目前常用的 Wi-Fi 指纹室内定位的算法比较,该研究算法具有如下特点:

1) 先进行楼层分类再进行每个楼层的具体定位,将三维的误差降低到二维误差,部分提高定位精度,大大降低了计算量。

2) 使用神经网络算法训练模型进行楼层定位,模型构成后只需要花费少量时间就可以得到分类的结果,减少了分类计算的时间,具有良好的时效性。

4.1. 楼层分类

4.1.1. 自动编码器模型

自动编码器[9]是深度学习的一种训练方法,它的实质其实是一种数据的压缩降维算法,其中数据的压缩和解压缩函数是数据相关的、有损的、从样本中自动学习并通过神经网络实现的一个过程。

自动编码器模型包括两个过程:编码和解码,编码主要作用是将高纬度的向量矩阵通过相互投影降低到一个低维的矩阵,但是主要的输入数据的典型特征仍然保持,只舍弃掉一些无关紧要的特征信息,相当于一个信息过滤器;解码过程是将低维的向量矩阵通过特征方程反向投影恢复到原始维度的矩阵数据,具体流程如图 4 所示。

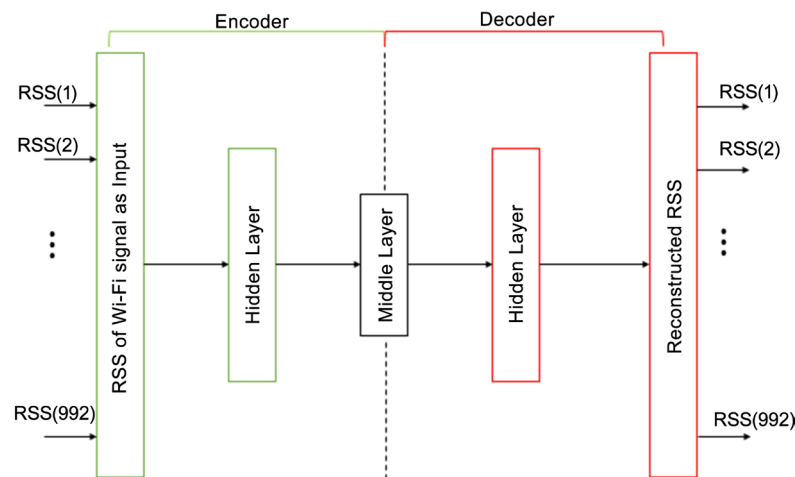


Figure 4. Flow chart of automatic encoder

图 4. 自动编码器流程图

自动编码器的输入是训练指纹中每组测量值的信号强度,解码器的输出是重构的和输入保持相同维度的向量(和输入相比主成分部分的数值相同,其他的数值被置零)。HL 代表隐层,括号中的数字代表层中神经元的数量。

4.1.2. 神经网络模型

神经网络[10]是一个利用多个隐含层和非线性激活函数构造模型,对多个输入数据进行不断加权求

和迭代,从而达到单一输出的机器学习算法。该模型包含一个输入层,多个隐藏神经元层,和一个单一输出层(当神经网络做分类时输出为类别,当神经网络做回归时输出为一个预测的特征值),模型如图 5 所示。

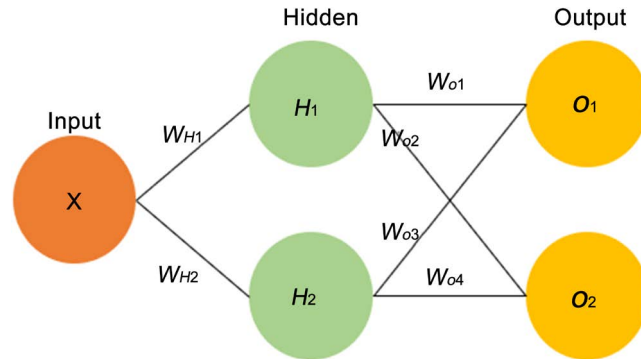


Figure 5. A simple neural network model
图 5. 一个简单的神经网络模型

神经网络使用梯度下降等优化技术进行迭代训练,在每次训练之后,得到预测值和标准值之间的偏差,用该偏差来调整下次迭代的权重参数。每次调整权重之后训练得到的预测和标准偏差都会不断减小,重复该过程直到偏差下降到可接受的阈值以下,神经网络模型训练即完成。

4.1.3. 楼层分类模型

该发明的楼层分类模型组合了自动编码模型和神经网络分类模型,如图 6 所示,已经预先训练的编码器连接到分类器。

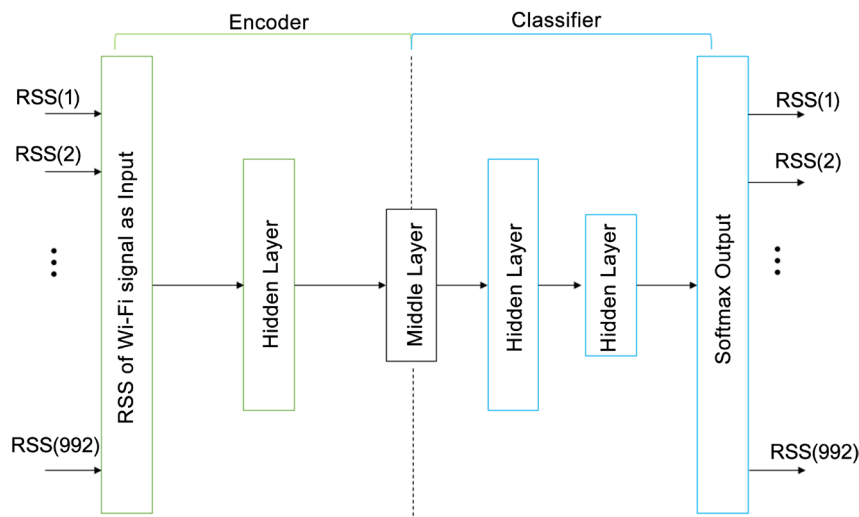


Figure 6. Floor classification model construction
图 6. 楼层分类模型构造

由于得到的每组 Wi-Fi 指纹数据中含有大量的无信号数据,仅含有少量有用信息(检测到 Wi-Fi 信号的数据),因此该模型先将输入数据导入到编码器中进行降维和主成分提取,删除部分无用信息同时减少输入信息量,从而减少后面神经网络分类器模型训练量,然后将编码后的数据当作神经网络模型的输入进行分类训练,提高分类结果精度。

4.2. 单楼层定位

对于常见的 KNN 算法, 根据指纹库中 K 个近似位置的坐标点取平均值的操作会带来一定的误差, 因此该算法使用了加权 KNN 算法, 使得每个距离更近的点能够获得更大的权重。该权重的形式与距离的倒数成正比, 因此距离越小的点获得的权重越大, 距离越大的点获得的权重也就相对小。计算过程如下: 1) 获取经过排序的距离值, 取距离最近的前 K 个元素组成一个向量 D ; 2) 对于每加权系数 $W = 1/d, d \in D$;

3) 根据权值估算测试指纹坐标 $Y = \frac{\sum_{i=1}^K X_i * W_i}{\sum_{i=1}^K W_i}$, X 为训练指纹坐标, Y 为估算坐标。通过使用加权 K 近

邻算法得到较 K 近邻算法更精确的定位结果。

5. 实验结果

本文采用四种不同的算法对数据进行训练和预测, 包括 KNN 算法、加权 KNN 算法、神经网络算法、神经网络和加权 KNN 的组合算法。从表 1 可以看出, KNN 算法的楼层分类准确率在 4 种算法中最低, 为 87%, 单楼层定位误差为 8.09 m; 加权 KNN 算法的准确率达到 91%, 单楼层定位较纯 KNN 算法误差要减少, 为 5.3 m; 神经网络算法模型、神经网络和加权 KNN 的组合算法模型的楼层分类准确率都达到了 92%, 但是神经网络和加权 KNN 的组合算法在单楼层的定位误差最小, 为 4.4 m。这与刘喜梅[10]等研究表明的无线网络定位算法中加权 KNN 算法在室内定位精确度较高的结论具有相似性, 可以作为一种高精度的组合算法进行应用。

Table 1. Comparisons of algorithms
表 1. 算法比较

算法	楼层分类准确率	单楼层定位误差(m)
KNN	87%	8.09
加权 KNN	91%	5.3
神经网络	92%	8.7
神经网络 + 加权 KNN	92%	4.4

6. 结论

本文基于传统 Wi-Fi 指纹室内定位算法, 研究并实现了一种能够在多楼层大范围的室内环境快速定位的算法, 结论如下:

1) 基于神经网络的楼层分类模型加上加权 KNN 算法在时效性上要优于各个传统 Wi-Fi 指纹室内定位算法。

2) 实验结果表明基于神经网络的楼层分类模型可以实现坐标与对应楼层的快速匹配, 使用加权 KNN 算法对单个楼层坐标可以进行更高精度的具体定位, 楼层分类准确率在四个算法中最高, 达到 92%, 同时多次定位的实验精度在 4 米左右, 效果较好, 往后可以应用于大型室内导航。

致 谢

感谢深圳市科技计划项目(项目编号 JSGG20170413170917828, 项目名称“室内停车位导航的关键技术研发”)的资助。

参考文献

- [1] 吴楠, 王旭东, 胡晴晴, 等. 基于多 LED 的高精度室内可见光定位方法[J]. 电子与信息学报, 2015, 37(3): 727-732.
- [2] 周宝定, 李清泉, 毛庆洲, 等. 用户行为感知辅助的室内行人定位[J]. 武汉大学学报(信息科学版), 2014, 39(6): 719-723.
- [3] Kolodziej, K.W. and Hjelm, J. (2006) Local Positioning Systems: LBS Applications and Services. CRC Press, Boca Raton, 101-158. <https://doi.org/10.1201/9780849333491>
- [4] Brunato, M. and Battiti, R. (2005) Statistical Learning Theory for Location Fingerprinting in Wireless LANs. *Computer Networks*, 47, 825-845. <https://doi.org/10.1016/j.comnet.2004.09.004>
- [5] 刘喜梅, 张超, 胡继珍. 一种复杂环境下无线传感器网络定位算法[J]. 青岛科技大学学报, 2009, 30(2): 173-175.
- [6] 卢恒惠, 刘兴川, 张超, 林孝康. 基于三角形与位置指纹识别算法的 WIFI 定位比较[J]. 移动通信, 2010, 34(10): 72-76.
- [7] 黄华晋, 秦建息. 基于三角定位算法的 WIFI 室内定位系统设计[J]. 广西科学院学报, 2016, 32(1): 59-61.
- [8] 田家英. 基于 WIFI 位置指纹算法室内定位技术研究[D]: [硕士学位论文]. 兰州: 兰州交通大学, 2018.
- [9] 司阳, 肖秦琨, 李兴. 基于自动编码器和神经网络的人体运动识别[J]. 国外电子测量技术, 2018, 37(1): 78-84.
- [10] 费扬. 基于 BP 神经网络的室内定位指纹算法研究[D]: [硕士学位论文]. 昆明: 昆明理工大学, 2018.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2163-3983, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: hjwc@hanspub.org