

DOI:10.15923/j.cnki.cn22-1382/t.2018.4.09

基于最小值的多映射数据结构 指纹库定位算法

郭昕刚, 胡 朗*

(长春工业大学 计算机科学与工程学院, 吉林 长春 130012)

摘 要: 在离线阶段对指纹库的构建采用多映射的数据结构,使每个采样点坐标对应同一 AP 的一个或多个 RSSI 信息值。在线匹配阶段,通过计算待定位点的 RSSI 信息与离线数据库的指纹信息的欧式距离,欧式距离值越小,说明两者的拟合程度越高,计算出两者拟合程度的最小值,以此找出与待定位点相关的指纹点。最后,使用 WKNN 匹配算法求出待定位点的估计坐标。

关键词: 室内定位; 位置指纹定位; 多映射数据结构; WKNN 算法

中图分类号: TP 393 **文献标志码:** A **文章编号:** 1674-1374(2018)04-0367-06

Location algorithm of fingerprint database based on minimum value for multi-mapping data structure

GUO Xingang, HU Lang*

(School of Computer Science & Engineering, Changchun University of Technology, Changchun 130012, China)

Abstract: Multi-mapping data structure is applied to build fingerprint database at off-line phase to make every sampling point match the RSSI information value of the same AP. While at online matching stage, Euclidean distance between the RSSI information of the undetermined site and the fingerprint information of the off-line database is calculated. The smaller the Euclidean distance value, the higher the fitting degree, which indicates that the minimum value of the fitting degree can be obtained. Fingerprint points associated with undetermined sites can be identified accordingly. Finally, WKNN matching algorithm is used to calculate the estimated coordinates of undetermined sites.

Key words: indoor positioning; location fingerprint positioning; multi-mapping data structure; Weighted K Nearest Neighbor.

收稿日期: 2018-05-17

基金项目: 吉林省重点科技攻关项目(20150204020SF)

作者简介: 郭昕刚(1979-),男,汉族,吉林长春人,长春工业大学副教授,硕士,主要从事嵌入式系统方向研究,E-mail:6889068@qq.com. * 通讯作者: 胡 朗(1995-),男,汉族,河南信阳人,长春工业大学硕士研究生,主要从事嵌入式系统方向研究,E-mail:1377838230@qq.com.

0 引言

近年来,随着移动互联时代的到来,人们对定位导航服务的需求越来越多。基于 GPS 的室外定位已经为人们提供了很大的便利,如汽车的定位跟踪、高德地图、滴滴打车应用等。对于室内环境,由于室内建筑物的遮挡、信号的多径传播等影响,GPS 定位不能满足室内对精度要求较高的定位需求。

针对室内 WiFi 的大量部署,WiFi 室内定位技术已经成为当今室内定位的研究热点。基于 WiFi 的室内定位主流方向通常分为:基于测距(range-based)和测距无关(range-free)两种。基于测距的定位方法一般为通过待定位点与 AP 节点的距离或者角度来计算定位目标的位置,常见的有:接收信号强度指示(RSSI)、信号到达角度(AOA)、信号到达时间(TOA)以及信号到达时间差(TDOA)等^[2]。而与测距无关的定位方法主要是依据对定位环境的场景特征分析,在定位时将待定位点的场景信息与参考点的场景信息进行对比实现定位,典型的应用算法是位置指纹法。该方法的主要优点是无需对 AP 节点位置和距离进行测量计算、无需添加额外的硬件设备。

在研究 WiFi 室内定位的过程中,基于位置指纹的定位方法是比较可行而且精度较高的定位方法,成为室内定位研究的热门方向。位置指纹定位方法主要分为离线阶段(off-line)和在线阶段(on-line)两个过程^[3]。在离线阶段,主要通过智能终端的 WLAN 无线接入点,采集室内各个 AP 的特征信息,把这些信息如 MAC 地址、RSSI 信号强度值以及采集点的位置坐标作为指纹存入指纹库中,遍历室内所有的采集点便构成了完整的位置指纹库。在线阶段通过智能终端实时接收到各个 AP 的 RSSI 信号值,通过与离线数据库的匹配找到与之相似的指纹参考点,进而估算出当前待定位点的位置坐标。

离线指纹库的合理构建和在线匹配算法的选择是影响 WiFi 室内定位精度和实时性的两个主要因素。通过规划和测量,按一定间距选取合适的参考点,在每个参考节点上测得各个 AP 的 MAC 地址和信号强度值。WiFi 信号的时变性以及信号传播的复杂性,使得测量的信号强度值并不能直接存储,而是需要滤波之后再使用,避免各种奇异信号值和小概率信号值存储到指纹库中影

响定位的稳定性^[4]。实时在线阶段,需要使用合适的匹配算法去寻找相关的指纹参考点,与待定位点的欧式距离越小,权值就越大,对目标位置的定位贡献程度就越高。

对于上述问题,文中采用了多映射的数据结构来构建离线阶段的位置指纹库,使得每个参考点位置对应同一 AP 的一个或多个 RSSI 值,这样在线匹配阶段通过计算欧式距离最小值,找到与待定位点最匹配的 K 个参考点,用 WKNN 算法实现位置的估算。这种方法有利于提高指纹库的准确性,减小定位误差。

1 多映射数据结构的数据库滤波算法

影响室内信号传播的因素很多,导致在采集信号的过程中同一位置对同一 AP 信号的不同时间 RSSI 值具有差异性^[5],因此需要对每个 AP 节点进行单独的数据采集与分析。移动终端在同一参考点处多次采集所获取的同一 AP 的 RSSI 信息,如图 1 所示。

对同一 AP 距离下的 RSSI 值采样分析发现,RSSI 信号值并不是固定在某个值下不变的,而是随着采样时间的变化,RSSI 值会发生波动^[6]。由于室内环境存在多径效应等影响,这种波动是随机发生的,所以在构建指纹库前,通常需要对信号进行滤波处理。

将采集到的同一 AP 样本值 $\{RSSI_1, RSSI_2, RSSI_3, \dots, RSSI_n\}$ 进行平均滤波:

$$\overline{RSSI} = \frac{1}{n} \sum_{i=1}^n RSSI_i \quad (1)$$

依次遍历所有参考点的各个 AP,并将所得的结果存入数据库,得到整个定位区间的离线指纹数据库^[7]:

$$LF = \begin{bmatrix} x_1 & y_1 & rssi_1^1 & rssi_1^2 & \cdots & rssi_1^n \\ x_2 & y_2 & rssi_2^1 & rssi_2^2 & \cdots & rssi_2^n \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m & y_m & rssi_m^1 & rssi_m^2 & \cdots & rssi_m^n \end{bmatrix} \quad (2)$$

但在匹配的过程中,平均滤波对定位精度的优化效果并不理想,原因主要是平均滤波对所有测得的 RSSI 值进行计算,这将引入一些对定位无关的 RSSI 值,这些值通常出现的概率很小,甚至不会出现^[7]。而在线过程中,采样时间很短,各种匹配算法都很难消除这些误差,导致匹配的结果会偏离实际位置很多。

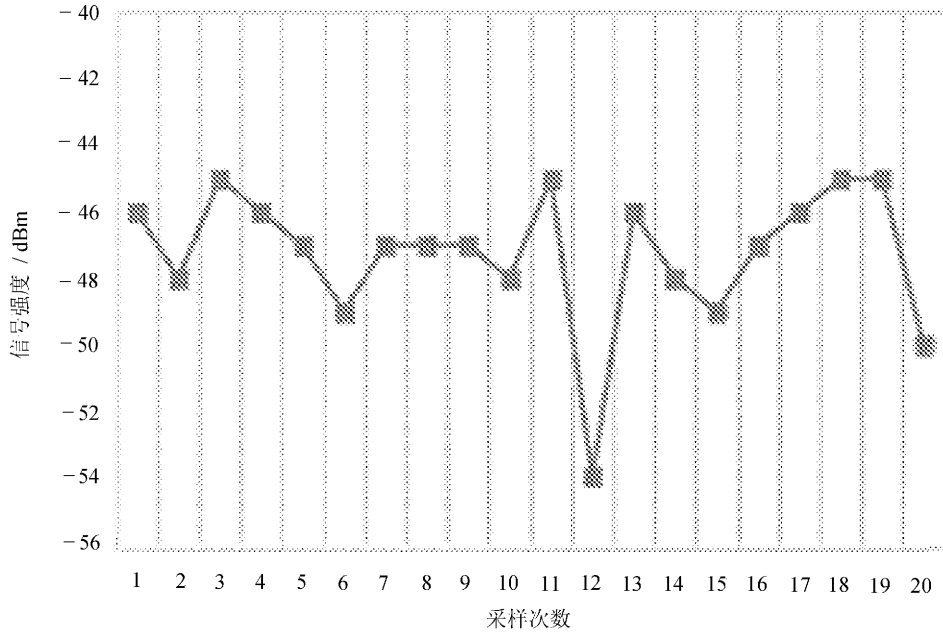
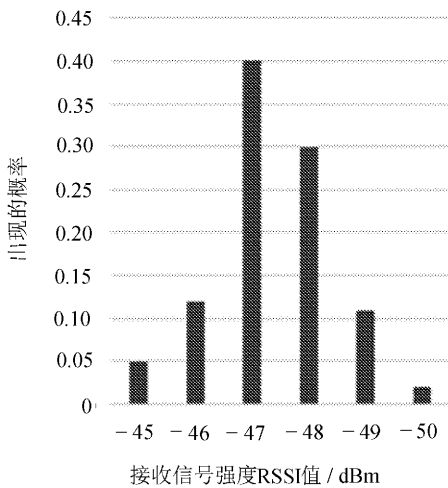


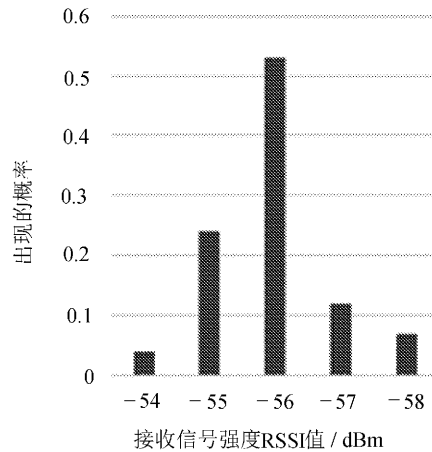
图 1 同一参考点所获取的同一 AP 的 RSSI 变化图

由于室内各个 AP 的独立性, 所以需要对各 AP 进行单独分析, 对两组不同 AP 进行采集所得的 RSSI 值概率分布, 如图 2 所示。

从图中可以看出, RSSI 值的波动呈现一种线性回归状态, 即 RSSI 在一定范围内波动, 且具有上下限。若设置概率阈值为 0.1, 则图 2(a) 中 RSSI 值集中在 $-49 \sim -46$ dBm 之间, 图 2(b) 中 RSSI 值集中在 $-57 \sim -55$ dBm 之间。



(a) 同一参考点接收到的 AP1 信号强度概率分布



(b) 同一参考点接收到的 AP2 信号强度概率分布

图 2 接收信号概率分布图

在室内环境复杂变化以及多径效应影响的情况下, 大多数值都在上下限范围之内, 极少情况会出现在范围之外, 这说明各 AP 的 RSSI 值满足正态分布规律^[8]。针对上述分析情况, 文中提出了一种多映射的数据库滤波算法, 合理构建离线指纹数据库, 以提高在线匹配的精度。

针对各 AP 信号呈现的规律特征, 在构建指纹库之前, 采用高斯滤波滤除 RSSI 信号值中的非常规数值, 保留出现可能性较大的 RSSI 数值。对每组 AP 的 RSSI 值处理则有:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (3)$$

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (4)$$

$$\sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \mu)^2} \quad (5)$$

式中： μ ——样本信号强度 RSSI 的平均值；

σ ——样本信号强度 RSSI 的标准差。

将概率范围在 $0.1 \leq f(x) \leq 1$ 的 RSSI 信号值取出，滤除小概率的异常信号，经过高斯滤波之后的每组 RSSI 样本值的集合为：

$$\text{RSSI}_{\text{gaussian}}^{\mu(x_i)} = \{ \text{rssi}_1, \text{rssi}_2, \text{rssi}_3, \dots, \text{rssi}_k \} \quad (6)$$

针对采样信号 RSSI 值出现的频数大小从左到右进行降序排序，得到排序后的信号 RSSI 值的集合：

$$\text{RSSI}_{\text{gaussian}}^{\text{Sortp}(x_i)} = \{ \text{rssi}_{a1}^{nk1}, \text{rssi}_{a2}^{nk2}, \text{rssi}_{a3}^{nk3}, \dots, \text{rssi}_{ak}^{nk} \} \quad (7)$$

式中： nk ——每个 RSSI 信号值出现的频数。

为了选取对定位有效的 RSSI 值存入到指纹库，下面将通过对频数排序后的 RSSI 信号值的集合进行筛选。首先选取出现次数最多的 RSSI

值的频数与集合中的总频数进行比较，所得的值如果大于概率阈值，则此次筛选结束，保留当前的 RSSI 值作为该点测得的 RSSI 信号值存入到指纹库。否则，将下一出现次数多的 RSSI 值的频数加到当前的 RSSI 值频数中，累加到一起与集合总频数比较，直到该占比大于概率阈值，并将这些累加的频数对应的 RSSI 值存入到指纹库。具体筛选的准则为：

$$\text{RSSI}_{\text{gaussian}}^{\text{filterp}(x_i)} = \left\{ \frac{\text{RSSI}_{\text{gaussian}}^{\text{sortp}(x_i)} \{nk\}}{\sum_{i=1}^k nk} \geq p_{\text{threshold}} \right\} \quad (8)$$

式中： nk ——对应 RSSI 值出现的频数；

$p_{\text{threshold}}$ ——筛选的概率阈值。

该算法这样筛选的目的是能最大限度地保留对定位有作用的 RSSI 值，在传统的指纹库中，同一位置点对应的同一 AP 仅仅只有一个值，而在该算法中往往可以得到同一位置对应的同一 AP 两个或者多个 RSSI 值。这样在定位匹配的过程中，可以找到与该位置最相近的 RSSI 值做比较。对应的指纹库见表 1。

表 1 多映射指纹库下参考点对应的 RSSI 分布表

参考点位置 Location	AP1	AP2	AP3
RP1	-48, -46	-62, -59, -58	-50, -47
RP2	-51, -49	-60, -58	-48, -45
RP3	-48, -45	-57, -55, -54	-52, -50, -48

室内定位范围内终端搜索到的 AP 节点有 3 个，分别为 AP1、AP2 和 AP3。参考点 RP1 位置处 AP1 值为 -48、-46 dBm，AP2 值为 -62、-59、-58 dBm，AP3 值为 -50、-47 dBm，这样 RP1 参考点存得的 RSSI 向量组合形式就有 12 种，在线匹配阶段，可以找到欧式距离最小的向量组合去参与匹配，降低最后的定位误差。

2 最小值的多映射数据结构在线匹配算法

在线定位阶段，比较常用的定位算法有最近邻算法(NN)、K 近邻算法(KNN)和加权 K 近邻算法(WKNN)^[9]。最近邻算法根据定位终端实时采集到 AP 的 RSSI 值与位置指纹库中的指纹信息进行匹配，找到相似度最大的位置指纹对应的参考点坐标作为待定位点坐标，而 K 近邻算法

是找到位置指纹库中 K 个指纹点并求得对应参考点坐标的平均值。WKNN 算法是在 KNN 算法上的改进，在找到 K 个指纹相似点之后，将每个相似点赋予不同的权重 ω_i ，一般来说，离目标接近的指纹点对待定位点的贡献大，相应的权重值大；反之，离目标稍远的指纹点对带定位点的贡献小，相应的权重值小^[10]。

为了找到对定位贡献最大的指纹参考点，在计算欧式距离过程中，文中引入了最小欧式距离的计算方法，舍弃了传统指纹库同一位置对应的同一 AP 下的单个 RSSI 值对定位造成的误差。具体步骤为：在线阶段我们测得各个 AP 的 RSSI 信息值，通过均值滤波得到每个 AP 的平均值，依次遍历所有 AP 就得到了该定位点的 RSSI 信息值。由于欧式距离的结果是两者所有对应 AP 节点相似度差值的累加，在参与欧式距离计算的过

程中,找到接近于实时测得的相应 AP 的 RSSI 值参与计算,这样能够在指纹库中的多种向量组合中找到最接近定位的 RSSI 向量组合。

用 $RSSI = \{RSSI_1, RSSI_2, RSSI_3, \dots, RSSI_n\}$ 表示定位终端在待定位点测量到的 n 个 AP 的

$$d(RSSI, rssi_j) = \sqrt{(RSSI_1 - rssi_{j1})^2 + (RSSI_2 - rssi_{j2})^2 + \dots + (RSSI_n - rssi_{jn})^2} \quad (9)$$

每个参考点的权重表示为:

$$\omega_i = \frac{1}{\frac{d(RSSI, rssi_j) + \epsilon}{\sum_{j=1}^k \frac{1}{d(RSSI, rssi_j) + \epsilon}}} \quad (10)$$

式中: ϵ ——一个无限接近于 0 的整数,目的是防止分母出现 0 的特殊情况。

最终待定位点的估计位置为:

$$(x, y) = \sum_{i=1}^k \omega_i (x_i, y_i) \quad (11)$$

3 实验与分析

文中提出的算法在长春工业大学计算机科学与工程学院实验室走廊区域进行实验验证,事先在定位区域部署 3 个 AP 发射点,以 1.6 m 的间

RSSI 信号强度值,用 $rssi_j = \{rssi_{j1}, rssi_{j2}, rssi_{j3}, \dots, rssi_{jn}\}$ 表示第 j 个指纹点存得的某一组合的信号强度,两点的 RSSI 信号强度的相似度用欧式距离表示为:

隔规划一定的采样点,采用智能手机 APP 采集 WiFi 信号。在每个采样点处采集 AP 节点的 RSSI 信号值,并将每个 AP 采集样本首先进行高斯滤波滤除小概率的噪声值,然后对保留下来的每个 RSSI 依次排序并计算单个 RSSI 值的频数占总体频数的比值,设置概率阈值对其进行有效筛选,得到最终保留下来的单个 AP 的样本值,依次遍历所有 AP 就得到了该采样点的指纹信息,将其保存在数据库中。在线匹配的过程中,在线测得的待定位点的 RSSI 信息值与多映射数据结构的指纹库进行匹配,找到欧式距离最小的 K 个指纹点信息,最后通过 WKNN 求得带定位点的坐标。整个定位的过程具体流程如图 4 和图 5 所示。

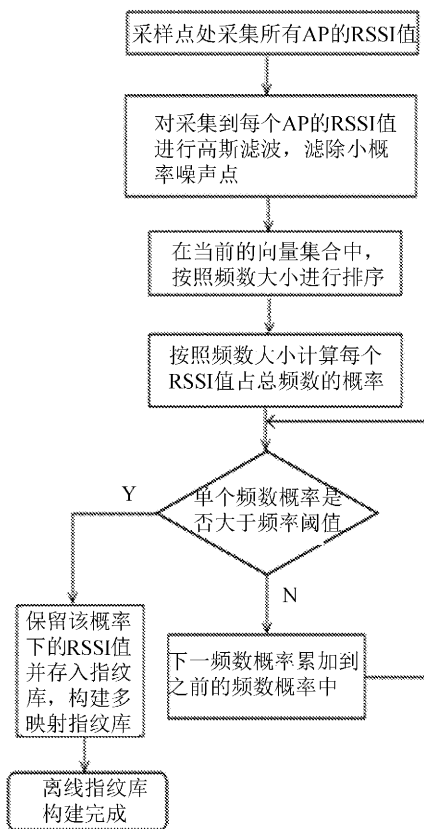


图 4 多映射数据结构指纹库构建流程

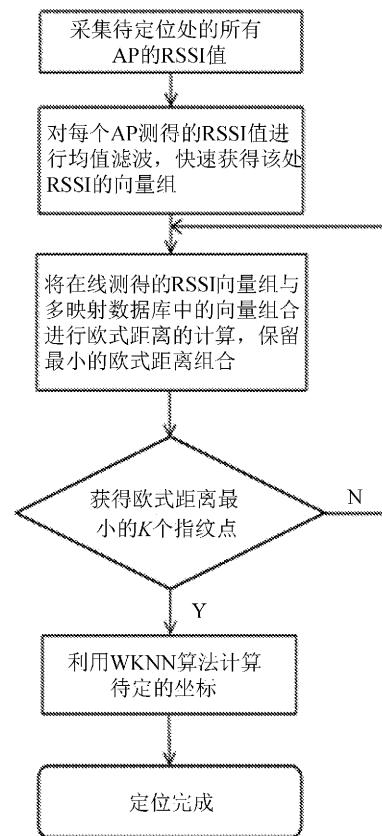


图 5 基于最小值的多映射数据结构在线定位流程

实验结果采用误差累计分布函数(Cumulated Denstiy Function, CDF)来衡量本定位算法相对于传统指纹库以及匹配算法的精度和稳定性,结果如图 6 所示。

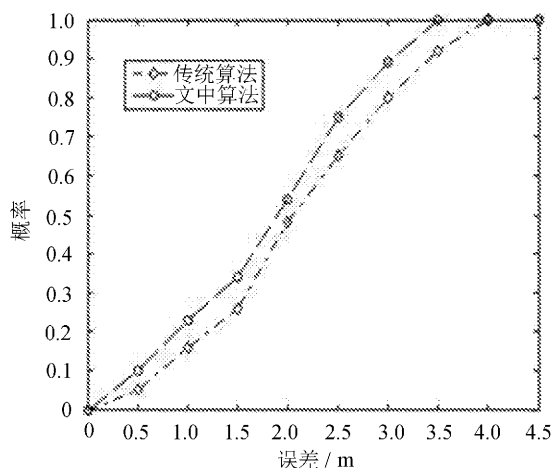


图 6 传统指纹库对应匹配算法和文中算法累计误差图

文中提出的算法能够有效提高定位的精度,减小了 WiFi 信号的变化以及传统指纹库的局限性对定位造成的误差。从图中能够看出,文中算法的误差在 3.5 m 范围内,同等误差距离下,该算法累计误差概率提高了将近 10%,验证了该算法的有效性和可行性。

4 结 语

针对传统指纹库单向量的构建形式造成的精度不高问题进行了深入分析,提出了一种多映射的数据结构指纹库以及最小值在线匹配算法,改变了传统指纹库的采样点对应每个 AP 的单个 RSSI 值的情况,以多个 RSSI 值映射同一个位置。在线匹配阶段,通过最小欧式距离的选择,能够找到拟合度最高的 RSSI 向量组合,得到对待定位点贡献大的参考点。实验结果表明,文中提

出的定位算法在精度以及稳定性方面具有一定的提升,对基于位置指纹的室内定位研究提供了一定的参考作用。

参考文献:

- [1] Xia Mingzhe, Chen Jiabin, Song Chunlei, et al. The Indoor Positioning Algorithm Research Based On Improved Location Fingerprinting [A]. Chinese Control and Decision Conference (CCDC) [C]//, 2015.
- [2] 张文学.基于 WiFi 的 RSSI 指纹定位算法研究[D].成都:电子科技大学,2015.
- [3] 高健.基于 WLAN 的室内定位指纹算法研究及应用[D].杭州:杭州电子科技大学,2017.
- [4] 陈戈珩,谢泽亮,李寻.多映射数据结构位置指纹库滤波算法[J].长春工业大学学报,2017,38(6):521-525.
- [5] Leu J S, Yu M C, Tzeng H J. Improving indoor positioning precision by using received signal strength fingerprint and footprint based on weighted ambient Wi-Fi signals[J]. Computer Networks, 2015, 91: 329-340.
- [6] 毕京学,汪云甲,曹鸿基,等.一种基于全向指纹库的 WiFi 定位方法[J].测绘通报,2018(2):25-29.
- [7] Wang F, Huang Z, Yu H, et al. EESM-Based fingerprint algorithm for Wi-Fi indoor positioning system [C]// IEEE International Conference on Communications in China, 2013:674-679.
- [8] 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):1067-1072.
- [9] 李林.基于位置指纹的 WiFi 室内定位技术研究与实现[D].哈尔滨:哈尔滨工业大学,2016.
- [10] 唐洋,白勇,马跃,等.基于 WiFi 的指纹匹配算法在室内定位中的应用研究[J].计算机科学,2016,43(5):73-75.