

# 基于多维尺度分析的自适应室内群终端定位方法

付先凯<sup>1,2,3</sup> 蒋鑫龙<sup>1,2,3</sup> 刘军发<sup>1,2,3</sup> 张少博<sup>4</sup> 陈益强<sup>1,2,3</sup>

(中国科学院计算技术研究所泛在计算系统研究中心 北京 100190)<sup>1</sup>

(中国科学院移动计算与新型终端北京市重点实验室 北京 100190)<sup>2</sup>

(中国科学院大学 北京 100049)<sup>3</sup> (长安大学信息工程学院 西安 710064)<sup>4</sup>

**摘要** 室内定位是普适计算领域的热点研究问题。当前,室内定位方法主要分为基于信号传播模型的定位方法和基于无线信号指纹的定位方法。其中,基于指纹的方法由于不需要知道无线信号接入点(Access Point, AP)的位置而得到更加广泛的应用,其需要通过离线阶段采集大量数据来构建丰富的指纹库,满足这一条件需要大量的人工标定工作。对此,文中提出了一种基于指纹空间关系的定位方法,相比于传统的指纹定位方法,该方法无需建立指纹库,只需要通过获取多终端的 Wi-Fi 信号强度,计算所有终端的不相似度并构建不相似矩阵;通过多维尺度分析(Multidimensional Scaling, MDS)算法,构建出所有终端的位置分布图,进而通过确定其中 3 个以上终端的位置来定位所有的终端。采用支持向量回归机(Support Vector Regression, SVR)计算任意终端间的距离,并将距离矩阵作为不相似矩阵。文中在商场场景下选择了约 2500 m<sup>2</sup> 的区域进行实验,所提方法的平均定位误差约为 7 m。

**关键词** 室内定位,指纹定位方法,多维尺度分析(MDS),支持向量回归机(SVR)

中图分类号 TP311 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.10.020

## Adaptive Indoor Location Method for Multiple Terminals Based on Multidimensional Scaling

FU Xian-kai<sup>1,2,3</sup> JIANG Xin-long<sup>1,2,3</sup> LIU Jun-fa<sup>1,2,3</sup> ZHANG Shao-bo<sup>4</sup> CHEN Yi-qiang<sup>1,2,3</sup>

(Center of Pervasive Computing System Research, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)<sup>1</sup>

(Beijing Key Laboratory of Mobile Computing and Pervasive Device, Chinese Academy of Sciences, Beijing 100190, China)<sup>2</sup>

(University of Chinese Academy of Sciences, Beijing 100049, China)<sup>3</sup>

(School of Information Engineering, Chang'an University, Xi'an 710064, China)<sup>4</sup>

**Abstract** Indoor location is a hot research topic in the field of pervasive computing. At present, indoor location methods are mainly divided into the localization method based on signal propagation model and the one based on wireless signal fingerprint. The fingerprint based method is more widely used because it does not need to know the location of the wireless signal AP. But it needs to collect a large amount of data at the offline stage to build a rich fingerprint database, which needs a lot of manual calibration. For this reason, this paper proposed a localization method based on spatial relations of fingerprints. Compared with the traditional fingerprint localization methods, this method does not need to build a fingerprint database. It uses Wi-Fi fingerprint from multiple terminals to extract the similarity of fingerprints and construct a dissimilarity matrix, and finally applies multidimensional scaling (MDS) algorithm to construct the relative location map for all terminals. Then each terminal can be positioned by determining the position of more than 3 terminals. In this paper, support vector regression (SVR) is used to calculate the distance between arbitrary terminals, and the distance matrix is used as the dissimilarity matrix. A shopping mall which is about 2500 square meter is selected as testing environment, and the average positioning error of the proposed method is about 7 meters.

**Keywords** Indoor positioning, Fingerprint location method, Multidimensional scaling, SVR

到稿日期:2018-03-21 返修日期:2018-06-16 本文受国家自然科学基金面上项目:面向可穿戴用户行为识别的增量学习方法研究(61572471),融合多元传播模型和指纹模型的免标定室内定位方法研究(61472399),递归深度神经网络的多极环神经动力学模型及方法研究(61572004),广东省科技计划项目:面向健康监护的新型智能贴片关键技术研发及产业化(2015B010105001)资助。

付先凯(1993—),男,硕士生,主要研究方向为机器学习、室内定位、行为识别等,E-mail: fuxiankai@ict.ac.cn;蒋鑫龙(1989—),男,博士,主要研究方向为室内定位、可穿戴计算等,E-mail: jiangxinlong@ict.ac.cn;刘军发(1973—),男,博士,副研究员,硕士生导师,主要研究方向为普适计算、虚拟现实、数据挖掘等,E-mail: liujunfa@ict.ac.cn(通信作者);张少博(1974—),男,博士,讲师,主要研究方向为智能控制理论、智能计算等,E-mail: shbzhang@chd.edu.cn;陈益强(1973—),男,博士,研究员,博士生导师,主要研究方向为普适计算、人工智能、人机交互等,E-mail: yqchen@ict.ac.cn.

## 1 引言

近年来,随着室内无线信号的全面覆盖,基于无线信号的定位方法得到了极大的发展<sup>[1]</sup>。目前室内基于无线信号的定位方法主要分为两类,即基于信号传播模型的定位方法和基于指纹的定位方法。

基于无线信号传播模型的定位方法包括信号到达时间(Time of Arrival, TOA)、信号到达时间差(Time Difference of Arrival, TDOA)、信号到达角度(Angle of Arrival, AOA)、三角定位等<sup>[2-4]</sup>。此类方法需要知道信号源及无线信号热点(Access Point, AP)的真实物理位置,并且对设备的精度有较高要求,往往需要额外的硬件设备,不适合自然环境下的普适定位<sup>[2]</sup>。

由于 Wi-Fi 和蓝牙的基础设施的普及和完善,基于无线信号的定位方法主要包括基于 Wi-Fi 的定位方法和基于蓝牙的定位方法。其中,基于 Wi-Fi 的指纹定位包括两个阶段:在离线阶段,采集 Wi-Fi 信号强度数据,并将其与相应的位置坐标绑定生成指纹数据库;在定位阶段,通过指纹匹配算法或机器学习的方法将当前指纹在数据库中进行匹配并得到定位结果。基于指纹的定位算法往往可以取得较高的定位精度,但是数据采集的成本很大,并且由于 Wi-Fi 信号的高动态性,数据库需要定期更新。

本文的目标是实现一种不需要建立指纹数据库的定位方法,从而避免庞大的数据采集工作。该方法对同一室内环境中的多个终端进行定位。为了构建多个终端的位置分布,文中应用了多维尺度分析(Multidimensional Scaling, MDS)算法。MDS 算法可以使数据在降维后仍然保持在高维特征空间中的相对位置关系,并且通过降维实现数据可视化。在定位领域,MDS 算法已经被成功应用于无线传感网络定位,其通过任意两节点间的距离,计算节点的位置分布。本文通过计算得到任意两终端间的距离,并通过 MDS 算法得到所有终端的位置分布,由于该分布是相对位置分布,因此分布图相对于真实位置可能发生放缩或旋转,但通过确定其中 3 个或 3 个以上终端的位置,可以得到所有终端的绝对位置分布,从而实现定位的目的。该方法可以对多终端进行实时定位,定位成本低,无需建立庞大的数据库,有用户群体即可定位。

本文第 2 节介绍相关的研究工作;第 3 节提出基于 MDS 的 Wi-Fi 终端定位方法;第 4 节给出实验及分析;最后总结全文并提出未来的目标。

## 2 相关研究工作

目前,有较多学者对基于信号传播模型的定位方法和基于无线信号指纹的定位方法展开了相关研究。

在对基于信号传播模型的定位方法的研究中,Cheung 等<sup>[1]</sup>采用信号到达时间(TOA)计算移动端与各基站之间的距离,从而实现定位。但 TOA 需要所有设备的时间绝对同步,定位成本较大<sup>[3-5]</sup>。Han 等<sup>[4,6]</sup>采用信号到达时间差(TDOA)方法实现室内定位,相比于 TOA, TDOA 只需要基站间时间同步即可。Niculescu 等<sup>[7]</sup>利用信号到达角度(AOA)实现室内定位的目的,但 AOA 需要额外的天线来测

量信号角度,增加了定位成本<sup>[3]</sup>。Kim 等<sup>[8]</sup>提出了一种加权的三角定位方法,利用 Wi-Fi 信号实现了室内移动设备的定位。此类方法的定位准确度较高,但在测距和定位时,需要保证设备间的时间同步,对设备的精度要求很高,难以在普通的智能设备上实现;并且,基于信号传播模型的方法需要获得定位环境的基站(Base Station, BS)地图或热点(Access Point, AP)地图作为先验信息<sup>[3-4]</sup>。

基于无线信号指纹的定位方法包括离线训练阶段和在线定位阶段。在离线训练阶段,Bahl 等<sup>[2]</sup>提出了一种名为 RA-DAR 的研究工作,其针对红外定位方法普适性差的缺点提出了基于无线网络射频信号的室内定位方法;Kaemarungsi 等<sup>[9-10]</sup>针对采用 802.11b 协议的 WLAN 网络信号和指纹特征进行研究以提升指纹定位精度;Wang 等<sup>[12]</sup>提出了一种有效的 Wi-Fi 指纹采集方法,并利用此方法提升了 Wi-Fi 指纹定位的精度。在在线定位阶段,Pei 等<sup>[11]</sup>采用 Wi-Fi 指纹定位方法,对手机等智能设备实现了室内定位;Lin 等<sup>[13]</sup>对多种在线定位方法进行了对比研究。

与此同时,一些研究者尝试将降维算法用于位置估计,此类方法挖掘数据在特征空间中的相对位置关系,通过数据可视化降维的方法实现定位,如 MDS 算法等<sup>[14]</sup>。当前,MDS 算法已经被成功应用于室内 AP 位置估算、室内终端定位、传感网络节点定位等方面。其中,Koo 等<sup>[14]</sup>提出了一种基于众包的数据采集方法,采用 MDS 方法对多楼层的 AP 位置进行计算,构建了室内 AP 地图;Cheung 等<sup>[1]</sup>利用 TOA 算法计算移动终端与基站间的距离,并利用 MDS 算法在室内环境中实现了对终端的定位;Section<sup>[16]</sup>将 MDS 算法应用于传感网络节点的定位,并利用 MDS 算法实现了对分布式传感网络节点的三维位置构建。以上研究成果均为本文提出基于 MDS 的自适应室内定位方法提供了参考与思路。

## 3 基于 MDS 的 Wi-Fi 群终端定位方法

MDS 算法在保持变量在高维空间中的原始关系的条件下,降低数据维度,输出维度通常为 2 或 3,以达到数据可视化的目的。该思想为获取二维位置分布图提供了方向。在社会学中,MDS 算法用于挖掘变量间的潜在关系;在定位领域中,MDS 算法曾用于传感网络节点的定位问题<sup>[15]</sup>。在位置估计中,往往使用距离矩阵作为数据间的不相似矩阵。MDS 算法的优势在于:即使距离估计存在误差,算法依然可以得到较好的位置分布结果<sup>[13]</sup>。

### 3.1 多维尺度分析(MDS)算法

多维尺度分析算法从数据的相似性或相异性出发,以保持变量在高维空间与低维空间中欧氏距离不变为原则,将变量从高维空间映射到低维空间,挖掘变量的潜在结构。

已知一组具有  $I$  个变量的数据集  $\{x_1, x_2, \dots, x_I\}$ , 并已知变量间的距离函数,由此可得任意两变量之间的距离  $\delta_{i,j}$  ( $i, j$  分别为数据集的第  $i$  个变量和第  $j$  个变量,  $i, j \in 1, \dots, I$ ), 如式(1)所示:

$$\delta_{i,j} \approx \|x_i - x_j\| \quad (1)$$

在度量 MDS(metric-MDS)中,式(1)中的  $\|x_i - x_j\|$  代表一个连续单调的距离函数,在非度量 MDS(nonmetric-

MDS)中,  $\|x_i - x_j\|$  代表一个离散的等级。这些距离可组成变量间的距离矩阵,如式(2)所示:

$$\Delta = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,I} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,I} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{I,1} & \delta_{I,2} & \cdots & \delta_{I,I} \end{pmatrix} \quad (2)$$

将矩阵的第  $i$  行记为  $\delta_i$ , MDS 得到了一组大小为  $I$  的向量集  $\{\delta_1, \delta_2, \dots, \delta_I, \delta_i \in \mathbb{R}^I\}$ , 矩阵的每一行对应  $\mathbb{R}^I$  空间中的一个向量,  $\mathbb{R}^I$  空间的维度代表原始变量与其他变量之间的距离。MDS 算法是在保持变量间相对距离不变的基础上, 寻找一个从数据集  $\mathbb{R}^I$  到  $\mathbb{R}^N$  之间的映射关系。如果选择维度  $N$  为 2 或 3, 则向量  $x_i$  即可在二维平面或三维空间中反映数据集  $I$  中各变量的结构关系。最终, MDS 将转化为计算式(3)的最优化问题:

$$\min_{x_1, \dots, x_I} \sum_{i < j} (\|x_i - x_j\| - \delta_{i,j})^2 \quad (3)$$

求解式(3)时, 构建矩阵  $X, T, D$ , 其中  $X_i$  为  $\mathbb{R}^N$  空间中第  $i$  个坐标点,  $t_{ij}$  为矩阵  $T$  中的元素,  $D$  为距离矩阵,  $d_{ij}$  为  $D$  中的元素。矩阵关系如式(4)所示:

$$\begin{cases} X^T = (X_1^T \cdots X_n^T) \\ d_{ij}^2 = (X_i - X_j)^2 = X_i^2 + X_j^2 - 2X_i X_j \\ T = XX^T \end{cases} \quad (4)$$

由方程组(4)可得:

$$t_{ij} = -\frac{1}{2}(d_{ij}^2 - X_i^2 - X_j^2) \quad (5)$$

同时得到距离矩阵  $D$  中元素与矩阵  $T$  中元素的关系:

$$t_{ij} = -\frac{1}{2}(d_{ij}^2 - \frac{1}{n} \sum_k d_{ik}^2 - \frac{1}{n} \sum_l d_{lj}^2 + \frac{1}{n^2} \sum_{k,l} d_{kl}^2) \quad (6)$$

最后, 对矩阵  $T$  进行特征值分解, 如式(7)所示。其中  $U$  为特征向量,  $\Lambda$  为特征值矩阵。

$$T = U\Lambda U^T = U\Lambda^{\frac{1}{2}}\Lambda^{\frac{1}{2}}U^T \quad (7)$$

令

$$\begin{cases} X = U\Lambda^{\frac{1}{2}} \\ X^T = \Lambda^{\frac{1}{2}}U^T \end{cases} \quad (8)$$

取  $X$  的前  $N$  个最大的特征值与特征向量对矩阵进行近似表示, 即完成了从  $\mathbb{R}^I$  空间到  $\mathbb{R}^N$  空间的降维。

### 3.2 基于 MDS 的定位方法

首先, 本文以一个城市位置估算问题为例, 引出基于 MDS 的定位方法。我们得到英国 6 座城市(伦敦(London)、加的夫(Cardiff)、伯明翰(Birmingham)、曼彻斯特(Manchester)、约克(York)、格拉斯哥(Glasgow))之间的欧氏距离, 试图估算 6 座城市的位置分布, 6 座城市的真实分布如图 1 所示。

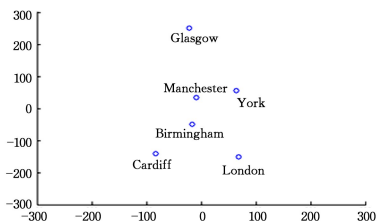


图 1 6 座城市的真实位置图

Fig. 1 Real position map of 6 cities

将每座城市作为变量, 将城市间的欧氏距离作为距离函数, 可得到距离矩阵  $\Delta_{6 \times 6}$ , 矩阵对角线元素是各变量与其自身的距离, 设为 0。设置 MDS 算法的输出维度为 2, 将变量从  $\mathbb{R}^6$  空间映射至  $\mathbb{R}^2$  空间, 即可得到位置估算结果, 如图 2 所示。

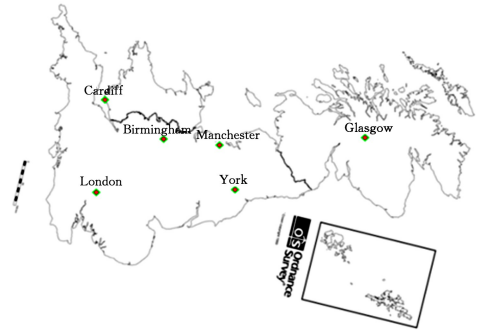


图 2 6 个城市的位置分布估计图

Fig. 2 Position distribution estimation map of 6 cities

通过 MDS 算法, 可得到 6 座城市的相对位置分布, 相对位置较真实位置可能发生放缩或旋转, 需要通过 3 个以上变量的真实位置确定所有变量的真实分布。

基于同样的思路, 本文提出了基于 MDS 的室内定位方法, 假设在室内环境中存在  $M$  个需要定位的用户, 记作  $\{u_1, u_2, \dots, u_M\}$ , 将用户间的欧氏距离  $D_E$  定义为距离函数, 记作  $D_E(i, j) = d(u_i, u_j), i, j \in M$ , 由此可以得到变量间的距离  $\delta_{i,j} = D_E(i, j)$  并构成距离矩阵  $\Delta$ , 如式(9)所示。其中, 距离函数应用机器学习回归的方法对用户所携带的智能终端(如智能手机等)间的欧氏距离进行计算, 该方法将在 3.4 节进行介绍。

$$\Delta = \begin{pmatrix} \delta_{1,1} & \delta_{1,2} & \cdots & \delta_{1,M} \\ \delta_{2,1} & \delta_{2,2} & \cdots & \delta_{2,M} \\ \vdots & \vdots & \ddots & \vdots \\ \delta_{M,1} & \delta_{M,2} & \cdots & \delta_{M,M} \end{pmatrix} \quad (9)$$

使用 MDS 算法, 在保证变量在  $\mathbb{R}^M$  中相对距离不变的基础上, 将向量映射到  $\mathbb{R}^2$  空间, 得到 2 维平面中的位置分布。设  $\mathbb{R}^2$  空间中的变量集合为  $\{X_1, X_2, \dots, X_M\}$ , 采用 3.1 节给出的 MDS 方法对式(10)的最小值问题进行求解。

$$\sum_{i < j} (\|x_i - x_j\| - d_{i,j})^2 \quad (10)$$

通过该方法可以得到原始数据在低维(2 维或 3 维)空间中的分布集合  $\{X_1, X_2, \dots, X_M\}$ , 该分布由相对距离得到, 因此该分布反映的是多个终端间的相对位置分布。当确定多个终端中至少 3 个终端位置时, 即可通过相对位置得到所有终端的绝对位置。该方法将终端间的距离作为变量之间的不相似度, 在计算终端距离时, 算法基于各终端指纹之间的差异特征而非指纹匹配, 因此无需依赖于具体的 AP 环境, 当定位环境中的 AP 发生变化时, 该方法依然可用。

为衡量算法的定位效果, 本文定义误差距离  $ED$  ( $ED$  依次取值 1~15 m, 间隔 1 m), 并统计定位误差在各误差距离内的定位点比例。令  $Y_p$  表示定位结果,  $Y_r$  表示真实位置, 当  $ED = 15$  m 时, 若  $|Y_r - Y_p| \leq ED$ , 则认为定位正确, 否则认为该点定位错误, 如图 3 所示。

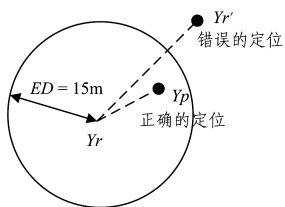


图 3 误差距离示意图

Fig. 3 Sketch map of error distance

### 3.3 Wi-Fi 信号的传播模型

本文选择了用户携带的终端间的欧氏距离作为变量间的距离(不相似度),将距离矩阵作为不相似矩阵。因此,如何准确地获取终端群中任意终端间的距离是本文的关键。对此,首先分析 Wi-Fi 信号的传播模型,式(11)是常用的无线网络信号传播模型,本文将式(11)作为 Wi-Fi 信号传播模型。

$$P_r = P_0 - 10n \log(d/l_0) + X_\delta \quad (11)$$

其中,  $P_r$  代表与发射机距离为  $d$  的信号接收强度指示(Received Signal Strength Indication, RSSI),单位为 dBm;  $P_0$  代表在与发射机(Wi-Fi 热点)距离为  $l_0$  的信号接收强度,通常  $l_0$  取 1 m;  $n$  为路径损耗指数;  $X_\delta$  代表标准差为  $\delta$  的阴影噪声。由该模型可知, Wi-Fi 信号的接收强度(RSSI)与由终端和 Wi-Fi 热点之间的距离呈对数衰减。针对传播模型的对数衰减特性,本文提出了一种分级的特征提取方式。

### 3.4 终端间的不相似(距离)矩阵

基于 MDS 算法的思想,若构建各变量在低维空间的分布,需要得到原始变量之间的距离,以构造其在高维空间中的距离矩阵。将 MDS 算法用于定位时,将欧氏距离作为其距离函数,若对  $M$  个用户(即用户携带的终端)  $\{u_1, u_2, \dots, u_M\}$  进行定位,则需要获取每台终端与所有终端之间的欧氏距离,记作  $D_E(i, j) = d(u_i, u_j)$ ,其中相同终端距离记作 0,即  $D_E(i, i) = d(u_i, u_i) = 0$ 。

本文使用机器学习中数据回归的方法实现终端间的距离计算。离线训练阶段可分为数据采集、特征提取、模型训练 3 个部分。

在数据采集部分,我们在训练区域采集 Wi-Fi 指纹数据,数据格式如下:

$$\{(x, y) | \langle mac_1, rssi_1 \rangle, \langle mac_2, rssi_2 \rangle, \dots, \langle mac_n, rssi_n \rangle\}$$

在原始数据中,  $(X, Y)$  为定位点坐标,  $\langle mac_i, rssi_i \rangle$  为第  $i$  个 Wi-Fi 热点的 MAC 地址和 RSSI。

接下来进行数据特征的提取。由于需要计算两终端之间的距离,因此特征提取将从不同指纹间的差异性出发。由于各终端处于不同的位置,其扫描到的 Wi-Fi 热点不尽相同,因此其获取到的 Wi-Fi 指纹数据处于不同的维度且不等长。为了得到等长的可供训练的特征数据,需要进行特征对齐处理。为了提取两条指纹间的差异,首先对任意两条指纹求交集,得到两条指纹间的相同 AP,记作  $I_{a,b} = F_a \cap F_b$ 。例如:

$$F_a = \{(x_a, y_a) | \langle mac_\alpha, rssi_1 \rangle, \langle mac_\beta, rssi_2 \rangle, \langle mac_\gamma, rssi_3 \rangle\}$$

$$F_b = \{(x_b, y_b) | \langle mac_\beta, rssi_4 \rangle, \langle mac_\gamma, rssi_5 \rangle, \langle mac_\theta, rssi_6 \rangle\}$$

则  $F_a \cap F_b = \langle mac_\beta, mac_\gamma \rangle$ 。

经本文实验,得到的 Wi-Fi 信号 RSSI 的范围在  $-20$  dBm 到  $-100$  dBm 之间,由于 RSSI 是对数衰减的,因此 RSSI 的差值在距离上不满足均匀分布,不能准确反映距离差异。因此,本文选择了一种分级的 RSSI 差异特征提取方式<sup>[6]</sup>。首先,将 RSSI 从  $-20$  dBm 到  $-100$  dBm 分为 7 个等级(Rank),阈值为  $-30$  dBm,  $-55$  dBm,  $-70$  dBm,  $-80$  dBm,  $-85$  dBm,  $-90$  dBm,分别记为 1 到 7,将原始指纹中的 RSSI 值转换为 RSSI 等级,记为  $Rank_i$ 。两条指纹的交集  $F_a \cap F_b$  包含了两条指纹搜索到的相同 AP,提取  $I_{a,b}$  中 AP 的 Rank 差的绝对值的最大值( $RankD_{max}$ )、最小值( $RankD_{min}$ )、平均值( $RankD_{mean}$ )以及  $I_{a,b}$  中的 AP 个数与两条指纹扫描到的所有 AP 数量的比值( $Rate_{sameAP}$ )作为特征,从而形成特征向量,形式如下:

$$D = \{RankD_{max}, RankD_{min}, RankD_{mean}, Rate_{sameAP}\}$$

对于少数交集为空的指纹对,采用填入固定值的方式解决。在一个终端附近假想一个 AP,该终端搜索到该 AP 的 RSSI 等级为 1,假设该 AP 功率足够大,另一个终端也可搜索到该 AP,其 RSSI 等级为 7,令两终端间特征向量中的  $Rate_{sameAP} = 0$ ,构成特征向量。

机器学习的回归方法包含逻辑回归(Logical regression)、决策树回归、集成学习(ensemble learning)、随机森林(random forest regression)、支持向量回归机(Support Vector Regression, SVR)、神经网络回归(neural network regression)等。在众多的机器学习方法中,支持向量机(Support Vector Machine, SVM)作为经典的统计学习方法,具有较强的理论支撑,可较好地避免传统机器学习中的“维数灾难”“过学习”等问题,并且由于支持向量机具有核函数特性、对小样本学习具有较高效率的优势以及相比于深度学习算法具有计算速度快的特点,本文选择了支持向量回归机(SVR)作为机器学习方法<sup>[17-18]</sup>,利用 SVR 对终端间距离进行回归预测。SVM 算法由 Vapnic 等提出,对解决小样本问题具有较好的学习效果,在本文的方法中可以满足实时定位的需求。在训练过程中,本文采用 libsvm 工具,选择 e-SVR 回归模型,并使用高斯核(Gaussian RBF)作为核函数对回归模型进行训练。

## 4 实验与分析

本文选择北京市某大型商场 1 层大厅中约 2500 平方米的区域进行实验,实验区域包括店铺、公共区域等,如图 4 所示。

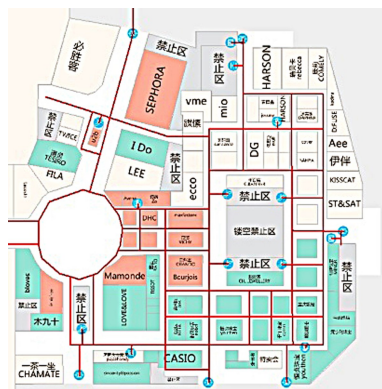


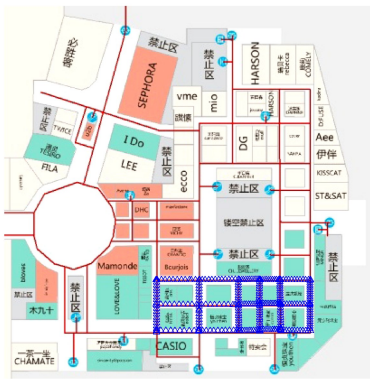
图 4 实验区域平面图(50m \* 50m)

Fig. 4 Experimental area plan (50m \* 50m)

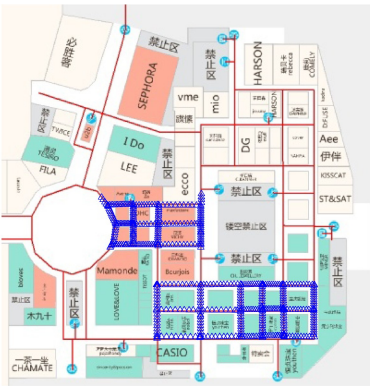
为了获取 Wi-Fi 数据,选择 HUAWEI H60-L03 智能手机作为 Wi-Fi 实验设备,实验设备的运行系统为 Android 4.4.2。在数据采集过程中,实验人员手持设备按商场的步行通道以正常步速行走(速度约 1 m/s)。采集软件以 2 Hz 的频率自动采集 Wi-Fi 数据,数据包含当前位置扫描到的 Wi-Fi 热点的名称、MAC 地址以及 RSSI。在实验区域共扫描到 Wi-Fi 热点 395 个,信号强度最强为 -23 dBm,最弱为 -100 dBm。

#### 4.1 室内终端定位

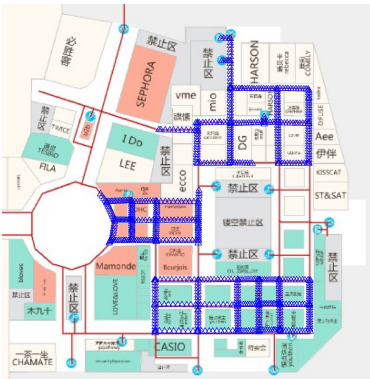
基于本文提出的定位方法,将距离矩阵(不相似矩阵)作为原始的高维数据输入,利用 MDS 算法对数据进行降维,设置输出维度为 2,得到终端的定位结果。在采集到的数据中,选择了 3 个子区域作为数据密集区,以模拟现实中人群集中的情况。本文共设计了 3 次定位实验,实验数据如图 5 所示。3 次实验的数据与结果如表 1 所列,实验区域面积约 2500 m<sup>2</sup>,平均误差为 7.29 m。



(a) 实验 1 终端位置图



(b) 实验 2 终端位置图



(c) 实验 3 终端位置图

图 5 终端真实位置图

Fig. 5 Real position map of devices

表 1 定位实验结果

Table 1 Positioning results

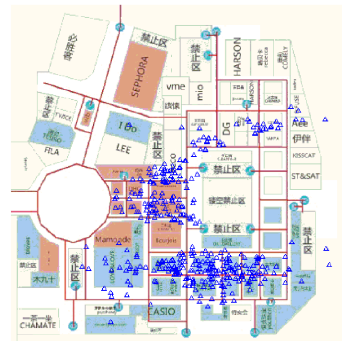
实验编号	终端数量	AP 个数	区域面积/m <sup>2</sup>	定位错误点数	定位准确率	平均误差/m
1	300	218	250	21	0.930	5.17
2	506	297	800	98	0.807	7.67
3	766	395	1600	126	0.836	9.03
平均	—	—	—	55	0.858	7.29

实验 1 在 250 m<sup>2</sup> 的店铺级区域进行,对 300 个终端进行定位,定位准确率达 90% 以上,平均误差约 5 m;实验 2 的定位点分布在 1000 平方米级的公共区域中,实验对 506 个终端进行定位,正确定位终端 408 个,定位准确率超过 80%,平均误差 7.67 m;实验 3 在更大的公共区域内进行,定位点分布面积约为 1600 m<sup>2</sup>,在 766 个定位点中,正确定位终端 640 个,定位准确率超过 83%,平均误差为 9.03 m。3 次定位实验的平均定位准确率超过 85%,平均定位误差约 7 m。

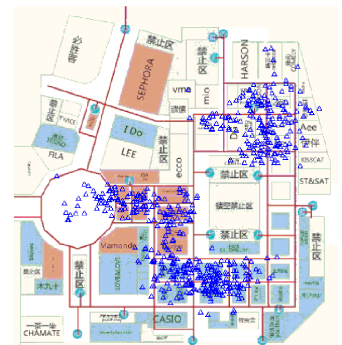
本文选择了面积约 2500 m<sup>2</sup> 的真实商场区域作为实验场地,定位结果如图 6 所示。



(a) 实验 1 的定位结果图



(b) 实验 2 的定位结果图



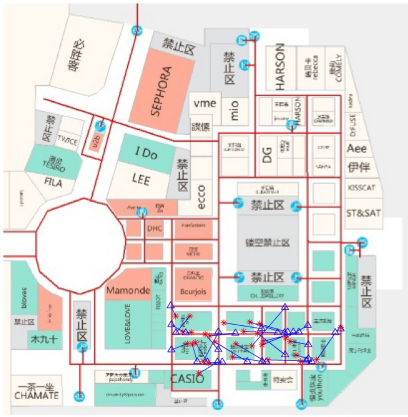
(c) 实验 3 的定位结果图

图 6 终端定位结果图

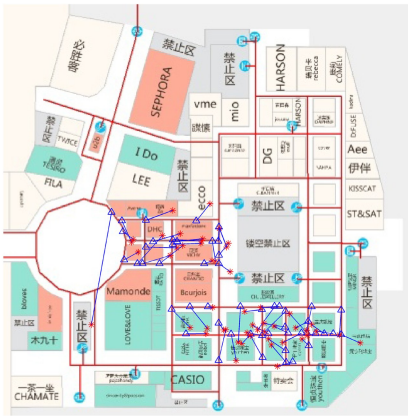
Fig. 6 Map of devices positioning results

经过 3 次实验验证,本文提出的基于多维尺度分析的室内终端定位算法的平均定位准确率可达到 85%以上,平均误差约 7 m,定位效果良好。

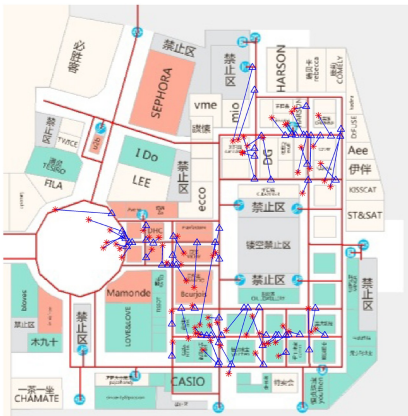
本文选择了实验中 10%的定位点,对比其真实位置与定位结果,如图 7 所示。



(a) 实验 1 定位结果对比图



(b) 实验 2 定位结果对比图



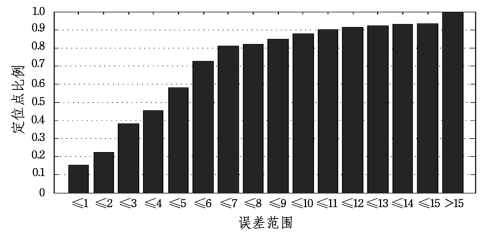
(c) 实验 3 定位结果对比图

注:真实位置用  $\Delta$  符号表示,定位结果用 \* 符号表示

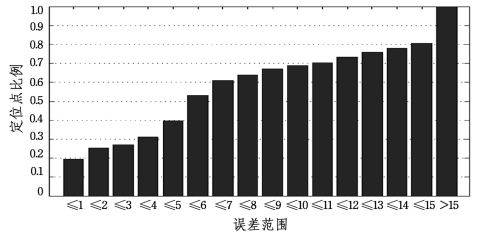
图 7 实际位置与定位结果对比图

Fig. 7 Comparison of real position and positioning results

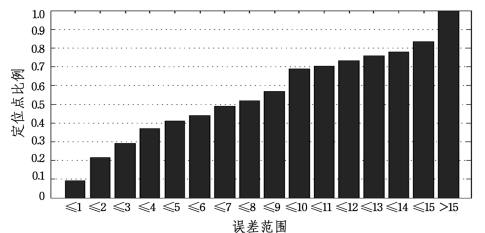
对于定位实验结果,本文统计了误差距离(ED)范围内的定位点比例,ED 取值为 1~15 m,间隔为 1 m,统计结果如图 8 所示。实验结果表明,在 2000 平方米级的实验环境内,上述定位方法的定位准确率为 85.8%,终端平均位置误差为 7.29 m,上述方法的定位效果可以完全满足大型区域中流量统计监控的需求。



(a) 实验 1 定位结果误差分析图



(b) 实验 2 定位结果误差分析图



(c) 实验 3 定位结果误差分析图

图 8 终端定位结果误差分析图

Fig. 8 Analysis of positioning results error

#### 4.2 终端距离计算

本文利用机器学习中的 SVR 算法对分级的 Wi-Fi 特征数据进行回归预测。实验采用 libsvm 工具,从已知的布设了 Wi-Fi 网络的环境中采集训练数据,训练 e-SVR 回归模型。

经实验验证,该模型在未知的室内环境中效果良好,平均距离预测误差为 7.68 m。经对 3 个区域的距离计算实验可知,该回归模型有较好的推广性,在未知环境中具有可用性。实验结果如表 2 所列。

表 2 终端距离计算结果

Table 2 Computed results of distance between devices

实验编号	测试数据量	区域面积/ $m^2$	平均误差/m
实验 1	45150	250	5.25
实验 2	128271	800	8.13
实验 3	293761	1600	9.67
平均	—	—	7.68

#### 4.3 实验分析

本节对处理 Wi-Fi 数据时采用的分级特征提取方法进行对比实验,结果如图 9 所示。

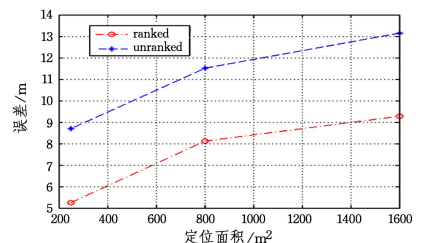


图 9 分级的特征提取方法效果对比图

Fig. 9 Comparison of ranked feature extraction method and unranked method

实验结果表明,依照 Wi-Fi 信号的传播模型对 Wi-Fi 信号 RSSI 进行分级处理的方法相比未分级的方法,可将距离计算误差减小 3~4 m。

下面在不同面积大小的定位环境中对比上述方法与 Wi-Fi 指纹定位方法。随着定位区域面积的增大,两种定位方法的定位效果如图 10 所示。

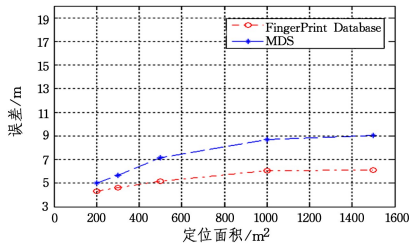


图 10 两种定位方法的效果对比图

Fig. 10 Comparison of two positioning methods

本文提出的基于 MDS 的定位方法在 1000~1500 m<sup>2</sup> 的定位环境中的定位误差约为 9 m,基于指纹数据库的定位方法的误差在 7 m 以内。在 1000 m<sup>2</sup> 以上的中大型定位环境中,本文提出的定位方法的精度和传统的指纹定位方法的精度相差 2 m 左右,但是本文提出的方法对不同定位环境具有自适应性且无需建立指纹数据库,节省了数据采集成本。

经以上实验验证,在大型商场环境中,本文提出的基于 MDS 的室内群终端定位方法的平均定位正确率约为 85%,平均定位误差约为 7 m。

**结束语** 面向室内定位的极大需求,研究者们提出了很多应用于室内的定位方法,但是未能克服以下问题:基于 Wi-Fi 指纹库的室内定位方法的数据采集成本高;基于信号传播模型的定位方法(如 TOA, TDOA 等)需要额外的定位设备且无法满足普适定位等。

针对以上问题,本文提出了一种对室内定位环境自适应的定位方法,该方法无需建立指纹数据库,也无需布设额外的硬件设备,只需用户携带具有 Wi-Fi 扫描功能的智能设备即可。

经实验验证,本文提出的基于多维尺度分析的群终端定位算法在大型商场环境中,平均可达到 85% 的定位准确度以及约 7 m 的定位精度,无数据采集成本且效果良好。该方法在公共场所的人群流量统计、流量监控、室内布局优化等方面均有重要意义。

后续,我们将面向更大面积和更复杂的室内环境进行研究,改进距离预测模型,进一步提升距离预测精度,并尝试进行多楼层的终端定位实验。

## 参 考 文 献

[1] CHEUNG K W, SO H C. A multidimensional scaling framework for mobile location using time-of-arrival measurements[J]. IEEE Transactions on Signal Processing, 2005, 53(2): 460-470.

[2] BAHL P, PADMANABHAN V N. RADAR: An In-Building RF-based User Location and Tracking System[C]//Proc. IEEE Infocom. 2000: 775-784.

[3] MAI A A, ALHADHRAMI S, ALSALMAN A, et al. Compar-

ative Survey of Indoor Positioning Technologies, Techniques, and Algorithms[C]//International Conference on Cyberworlds. IEEE, 2014: 245-252.

- [4] FARID Z, NORDIN R, ISMAIL M. Recent advances in wireless indoor localization techniques and system[J]. Journal of Computer Networks and Communications, 2013, 2013(42): 15.
- [5] GU Y, LO A, NIEMEGERERS I. A survey of indoor positioning systems for wireless personal networks[J]. IEEE Communications Surveys & Tutorials, 2009, 11(1): 13-32.
- [6] HAN T, LU X, LAN Q. Pattern recognition based Kalman filter for indoor localization using TDOA algorithm [J]. Applied Mathematical Modelling, 2010, 34(10): 2893-2900.
- [7] NICULESCU D, NATH B. VOR base stations for indoor 802.11 positioning[C]//International Conference on Mobile Computing and Networking. DBLP, 2004: 58-69.
- [8] KIM B, BONG W, KIM Y C. Indoor localization for Wi-Fi devices by cross-monitoring AP and weighted triangulation[C]//Consumer Communications and NETWORKING Conference. IEEE, 2011: 933-936.
- [9] KAEMARUNGSI K, KRISHNAMURTHY P. Properties of Indoor Received Signal Strength for WLAN Location Fingerprinting[C]//International Conference on Mobile and Ubiquitous Systems: NETWORKING and Services, 2004. DBLP, 2004: 14-23.
- [10] KAEMARUNGSI K, KRISHNAMURTHY P. Analysis of WLAN's received signal strength indication for indoor location fingerprinting[J]. Pervasive & Mobile Computing, 2012, 8(2): 292-316.
- [11] PEI L, GUINNESS R, CHEN R, et al. Human Behavior Cognition Using Smartphone Sensors[J]. Sensors, 2013, 13(2): 1402.
- [12] WANG F, HUANG Z, YU H, et al. EESM-based fingerprint algorithm for Wi-Fi indoor positioning system[C]//IEEE/CIC International Conference on Communications in China. IEEE, 2013: 674-679.
- [13] LIN T N, LIN P C. Performance comparison of indoor positioning techniques based on location fingerprinting in wireless networks[C]//International Conference on Wireless Networks, Communications and Mobile Computing. IEEE Xplore, 2005: 1569-1574.
- [14] KOO J, CHA H. Autonomous construction of a Wi-Fi access point map using multidimensional scaling[C]//Pervasive Computing, International Conference. DBLP, 2011: 115-132.
- [15] JI X, ZHA H. Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling[C]//Joint Conference of the IEEE Computer and Communications Societies. IEEE Xplore, 2004: 2652-2661.
- [16] SECTION. D3D-MDS: A Distributed 3D Localization Scheme for an Irregular Wireless Sensor Network Using Multidimensional Scaling[J]. International Journal of Distributed Sensor Networks, 2015, 2015(7): 1-10.
- [17] OSUNA E, FREUND R, GIROSI F. An improved training algorithm for support vector machines[C]//Neural Networks for Signal Processing VII-Proceedings of the 1997 IEEE Workshop. 1997: 276-285.