

文章编号: 2095-4980(2013)04-0561-06

## 基于隐马尔可夫模型的 Cell-ID 定位跟踪方法

竹 博, 王建辉, 胡捍英, 王大鸣

(解放军信息工程大学 信息工程学院, 河南 郑州 450002)

**摘要:** 针对传统基地台识别码(Cell-ID)定位方法精确度不高的问题, 提出一种新的 Cell-ID 定位跟踪算法。该算法利用移动台当前的服务基站信息和相邻基站信息, 分别加权计算出似然向量, 通过隐马尔可夫模型来获得移动台的运动状态, 最后利用最大后验概率准则求出移动台的位置。实验结果表明, 该算法复杂度低, 在服务基站稳定的区域内定位性能良好, 达到通信委员会(FCC)的定位精确度要求。

**关键词:** 蜂窝网; 定位跟踪; 基地台识别码; 隐马尔可夫模型

**中图分类号:** TN911.72

**文献标识码:** A

**doi:** 10.11805/TKYDA201304.0561

## A location and tracking algorithm using Cell-ID based on Hidden Markov Model

ZHU Bo, WANG Jian-hui, HU Han-ying, WANG Da-ming

(Institute of Information Engineering, PLA Information Engineering University, Zhengzhou Henan 450002, China)

**Abstract:** Considering the low precision of traditional Cell-identity(ID) algorithm, this paper proposes a new algorithm of Cell-ID location and tracking. The proposed algorithm makes use of current Base Station(BS) and the adjacent BS, and calculates the likelihood vectors using different weight coefficients. The motion states of Mobile Station(MS) are obtained through Hidden Markov Model(HMM), and the position of MS is determined. The experimental results show that the algorithm bears good location performance in stable serving BS with low complexity, which meets the location requirements of Federal Communications Commission(FCC).

**Key words:** cellular; location and tracking; Cell-identity; Hidden Markov Model

蜂窝网定位是无线定位技术中重要的一类, 它是通过蜂窝网系统中位置已知的若干测量传感器对移动台(MS)的无线信号相关参数进行测量, 然后采用相应的定位算法获取移动台的位置信息。近几年来, 蜂窝网定位技术引起了各国的广泛关注, 得到了迅速的发展和较深入的研究。

定位往往是通过测量基站(BS)与移动台之间无线信号的相关参数实现的, 如信号强度<sup>[1-4]</sup>、信号到达时间<sup>[5]</sup>、信号到达时间差<sup>[6]</sup>、来波角度<sup>[7-8]</sup>和小区身份等, 另外还有基于以上几种的混合定位方法也在文献[9]中提出。在视距环境下, 这些方法性能不错, 但是实际环境, 尤其是非视距环境下, 上述方法的性能会受到严重影响。针对非视距环境, 文献[10]给出一种指纹匹配方法, 它利用了电磁波在非视距下的多径特征, 但是这种方法要求添加新的硬件, 提高了移动台或基站的成本。文献[11]提出了一种基于隐马尔可夫模型(HMM)的 Cell-ID 定位跟踪方法, 它的思想是: 首先把某一区域均匀划分成若干网格, 在每个网格上均匀取点并实测移动台在每个点接收到的基站信息; 然后建立数据库, 采用 HMM 滤波器来获得当前移动台的位置。但是在实际环境中, 该算法存在两点不足: a) 均匀划分的区域并不能保证数据库中不同网格之间的差异性, 比如划分的网格面积太小会导致数据库上相邻网格的差异性很小, 会带来大量比较, 从而算法变得复杂; 网格面积太大会导致定位的精确度变差。b) 该算法只考虑所有的接收基站, 不区分当前服务基站与相邻基站, 都统一进行相似度比较。而在实际环境的多数情况下, 收到的服务小区基站比相邻小区基站稳定得多, 只有在多个小区边缘时, 服务小区基站较相邻小区基站不稳定。

基于以上分析, 本文在文献[11]的基础上, 提出一种新的 Cell-ID 定位跟踪算法, 该算法考虑到实际环境下服务基站与相邻基站的稳定程度不同, 合理划分区域并建立数据库, 然后分别给定服务基站和相邻基站不同的权

收稿日期: 2012-08-26; 修回日期: 2012-09-27

基金项目: 国家自然科学基金资助项目(2011ZX03003-003-02)

值, 求出似然向量, 并通过 HMM 滤波器获得移动台的运动状态, 最后通过最大后验概率准则求出移动台当前所在的区域, 并用区域的几何中心代表移动台的真实位置, 进一步提高了定位精确度。

## 1 HMM 定位跟踪算法

### 1.1 HMM 的参数建立

HMM 是在马尔可夫链基础上发展起来的。实际问题中观察到的事件并不与状态直接对应, 而是通过一组概率分布相联系, 这样的模型就称为隐马尔可夫模型。它由双重随机过程组成: 马尔可夫链和一般随机过程, 见图 1。其中, 马尔可夫链描述系统的状态转移, 受状态转移矩阵  $\boldsymbol{\Pi}$  控制; 一般随机过程描述系统的状态与观测符号之间的统计关系, 受观测符号矩阵  $\boldsymbol{H}$  控制。HMM 模型中, 状态及其转移过程不可观测, 外界只能通过观测符号序列去推测。

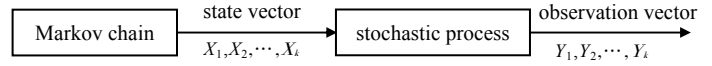


Fig.1 Hidden Markov Model  
图 1 HMM 示意图

将 Cell-ID 定位跟踪刻画成一个 HMM 结构, 首先将定位区域划分为  $N_{\text{spot}}$  个小区, 并依次进行编号。下面对参数选取进行说明。

1) 定义状态向量  $\boldsymbol{X} \in \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_{N_{\text{spot}}}\}$ ,  $\boldsymbol{e}_i$  是第  $i$  个元素为 1, 其他元素为 0 的  $N_{\text{spot}} \times 1$  列向量,  $\boldsymbol{X}_k = \boldsymbol{e}_i$  表示在  $k$  时刻移动台位于第  $i$  个区域。作者认为  $\{\boldsymbol{X}_k\}$  是马尔可夫过程。

2)  $\boldsymbol{\Pi} \triangleq [\pi_{ij}]$  是大小为  $N_{\text{spot}} \times N_{\text{spot}}$  状态转移矩阵。

$$\pi_{ij} \triangleq \boldsymbol{P}(\boldsymbol{X}_{k+1} = \boldsymbol{e}_j | \boldsymbol{X}_k = \boldsymbol{e}_i) \quad (1)$$

$\pi_{ij}$  的选取如下: 设区域  $i$  相邻区域集为  $\boldsymbol{A}_i = \{\lambda_1, \lambda_2, \dots, \lambda_{|\boldsymbol{A}_i|}\}$ , 其中  $|\boldsymbol{A}_i|$  表示与区域  $i$  相邻区域集的个数。选取 2 个数  $0 < p_1, p_2 > 1$  且满足下式:

$$p_1 > |\boldsymbol{A}_i| p_2, \quad p_1 + |\boldsymbol{A}_i| p_2 < 1 \quad (2)$$

则状态转移矩阵如下:

$$\pi_{ij} = \begin{cases} p_1, & j = i \\ p_2, & j \in \boldsymbol{A}_i \\ \frac{1 - p_1 - |\boldsymbol{A}_i| p_2}{N_{\text{spot}} - |\boldsymbol{A}_i| - 1}, & \text{其他} \end{cases} \quad (3)$$

在下面的实际测试中, 考虑到划分区域的大小及移动台的速度, 选取  $p_1 = 0.8, p_2 = 0.02$ 。

3) 定义观测向量  $\boldsymbol{Y} \triangleq \begin{bmatrix} \boldsymbol{Y}_s \\ \boldsymbol{Y}_a \end{bmatrix} \in \mathbf{R}^{2 \times N_{\text{ID}}}$ ,  $N_{\text{ID}}$  是在这一区域上移动台接收到的所有基站个数, 分为服务基站  $\boldsymbol{Y}_s \in \mathbf{R}^{N_{\text{ID}}}$  和相邻基站  $\boldsymbol{Y}_a \in \mathbf{R}^{N_{\text{ID}}}$ 。

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_s \\ \boldsymbol{Y}_a \end{bmatrix} = [\boldsymbol{Y}_s(1), \boldsymbol{Y}_s(2), \dots, \boldsymbol{Y}_s(N_{\text{ID}}), \boldsymbol{Y}_a(1), \boldsymbol{Y}_a(2), \dots, \boldsymbol{Y}_a(N_{\text{ID}})]^T \quad (4)$$

$$\boldsymbol{Y}_s(i) = \begin{cases} 1, & \text{当前移动台的服务基站是第 } i \text{ 个基站} \\ 0, & \text{当前移动台的服务基站不是第 } i \text{ 个基站} \end{cases} \quad (5)$$

$$\boldsymbol{Y}_a(i) = \begin{cases} 1, & \text{当前移动台接收的相邻基站包括第 } i \text{ 个基站} \\ 0, & \text{当前移动台接收的相邻基站不包括第 } i \text{ 个基站} \end{cases}$$

4)  $\boldsymbol{H} \triangleq \begin{bmatrix} \boldsymbol{H}_s \\ \boldsymbol{H}_a \end{bmatrix}$  是大小为  $N_{2 \times \text{ID}} \times N_{\text{spot}}$  的观测符号矩阵, 与  $\boldsymbol{Y}$  定义相似,  $\boldsymbol{H}$  分为服务基站  $\boldsymbol{H}_s$  和相邻基站  $\boldsymbol{H}_a$  两部分,  $\boldsymbol{H}$  中的第  $j$  列代表移动台在第  $j$  个区域上所能收到各个基站的概率。

$$\boldsymbol{H}_s(i, j) = \boldsymbol{P}([\boldsymbol{Y}_s]_i = 1 | \boldsymbol{X}_k = \boldsymbol{e}_j), \quad 1 \leq i \leq N_{\text{ID}} \quad (6)$$

$$\boldsymbol{H}_a(i, j) = \boldsymbol{P}([\boldsymbol{Y}_a]_i = 1 | \boldsymbol{X}_k = \boldsymbol{e}_j), \quad N_{\text{ID}} \leq i \leq N_{2 \times \text{ID}}$$

此时, Cell-ID 估计问题可以转换成下式的状态估计问题:

$$\hat{\boldsymbol{X}}_{k|k} = E[\boldsymbol{X}_k | \boldsymbol{Y}_{0:k}] \quad (7)$$

元素  $[\hat{\boldsymbol{X}}_{k|k}]_i$  是  $\boldsymbol{P}(\boldsymbol{X}_k = \boldsymbol{e}_i | \boldsymbol{Y}_{0:k})$  的后验概率, 即已知所有的观测值时移动台位于第  $i$  个区域的概率。



用一个  $N_{2-ID} \times 1$  的列向量  $\mathbf{Y}_{db}^{(i)}$  来表示。

把每一个网格内的  $N$  个  $\mathbf{Y}_{db}^{(i)}$  相加并求平均, 得到矩阵  $\mathbf{P}$ ,  $\mathbf{P}$  的第  $j$  列如下:

$$[\mathbf{P}]_{:,j} = \frac{1}{|\{i | p_{db}^{(i)} \in \mathbf{S}_j\}|} \sum_{\{i | p_{db}^{(i)} \in \mathbf{S}_j\}} \mathbf{Y}_{db}^{(i)} \quad (13)$$

式中  $\mathbf{S}_j$  表示第  $j$  个网格内的所有测试点。可以看出,  $[\mathbf{P}]_{:,j}$  代表在区域  $j$  上所能收到各个基站的概率。

3) 把矩阵  $\mathbf{P}$  中相似的若干列通过相加并求平均, 形成最终的数据库  $\mathbf{H}$ ,  $\mathbf{H}$  中的第  $j$  列代表移动台在第  $j$  个区域上所能收到各个基站的概率。并把相似的网格合并, 见图 4, 黑线围起来的就是最终的区域, 一共有 13 个区域, 大小从  $60 \text{ m} \times 30 \text{ m}$  到  $150 \text{ m} \times 90 \text{ m}$  不等。

文献[11]的数据库中只有移动台当前接收到的各个基站的概率, 而本文根据实际环境中接收到的服务基站和相邻基站稳定程度的不同, 将接收到的基站分为 2 类: 一是当前服务基站, 二是相邻基站, 因此  $\mathbf{H}$  分为接收到的服务基站的概率  $\mathbf{H}_s$  和相邻基站的概率  $\mathbf{H}_a$ 。

结合实测数据, 本文将上述 13 个区域分为 2 部分:

- 1) 服务基站比较稳定的区域, 是指某一个区域能以大于 0.7 的概率稳定收到一个固定的基站作为服务基站;
- 2) 服务基站不稳定的区域, 是指某一区域的服务基站在多个之间来回切换, 概率都小于 0.7, 一般在多个蜂窝网小区的交接处。

这 13 个区域中, 大部分都是类型 1, 只有区域 5, 7 和 10 是类型 2。

## 2.2 实验结果分析

本节利用实测数据对所提算法的性能进行验证, 实验分为 2 部分:

- 1) 在数据库的基础上, 如何得到使平均定位误差最小时的权值  $\theta$ ;
- 2) 在 1) 的基础上, 与文献[11]的算法进行性能比较。

实验 1: 最佳权值  $\theta$  的选取

图 4 中实线代表移动台的运动轨迹, 速度为  $1 \text{ m/s}$ , 运动轨迹基本包括所有区域, 每隔  $1 \text{ s}$  采一个点, 一共有 1490 个采样点。图 5 说明了服务小区的权值  $\theta$  与平均定位误差的关系。可以看出, 当权值为 0.5 时, 此时相当于文献[11]的算法, 平均定位误差大约为  $59.7 \text{ m}$ 。随着权值的增加, 平均定位误差也随之减小, 当权值为 0.87 时, 平均定位误差最小, 为  $45.5 \text{ m}$ , 因此本区域的最佳权值为 0.87。此时, 随着权值的进一步增加, 平均定位误差又开始变大, 这是因为以较大概率接收到同一个服务基站的区域不只一个, 这些区域的差别在于接收到相邻基站概率的差异, 而当给服务基站很大的权值时, 可能就会产生误判, 因此平均定位误差会随着权值的增加而增大。

实验 2: 所提算法与文献[11]算法性能对比

实验 2 分为 2 部分, 在实验 1 确定最佳权值的基础上, 分别在服务基站比较稳定的区域和不稳定的区域下与文献[11]算法进行性能比较。



Fig.3 Area of nonuniform segmentation  
图 3 基站布图

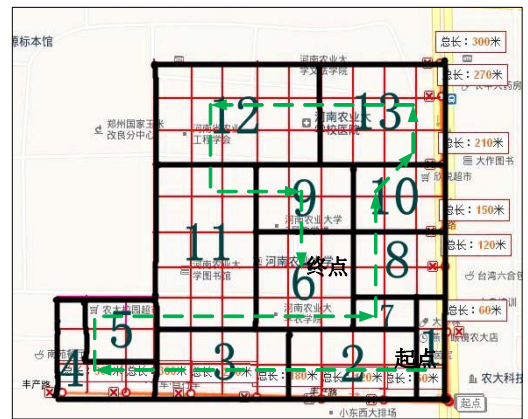


Fig.4 Area after merger  
图 4 合并区域后的图

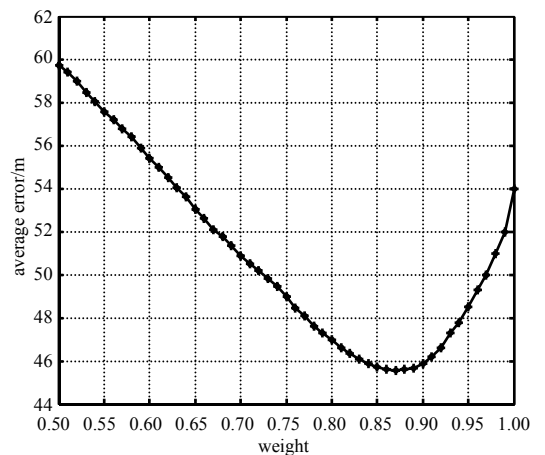


Fig.5 Average error vs. weight  
图 5 平均定位误差和权值的关系图

文献[11]是把区域均匀划分,形成面积相等的网格,然而实际情况中,并不能保证均匀划分形成的网格间数据的差异性,此时就可能产生下面问题:如当移动台位于某一网格内时,因为数据库上相邻的网格差异性不大,定位结果可能就会落在与它相邻的其他网格内。这样既增加了算法的复杂度,且定位精确度没有明显提高。基于以上分析,所提算法在文献[11]的均匀划分区域的基础上,把相似的网格合并,形成不规则的区域,然后用服务基站加权的方法进行定位。

图6分别对比了在服务基站比较稳定的区域下,加权算法、非加权算法在均匀划分和非均匀划分情况下以及传统的 Cell-ID 定位性能。其中 CDF 为累积概率分布函数(Cumulative Distribution Function)。可以看出,所提算法以 67% 的概率定位精确度是 48 m,以 95% 的概率定位精确度是 100 m,而文献[11]中算法以 67% 的概率定位精确度是 60 m,以 95% 的概率定位精确度是 142 m,因此,所提算法性能优于文献[11]算法,且满足 FCC 的定位精确度要求:定位精确度 50 m,准确率 67%;精确度 150 m,准确率 95%。

在同等的加权或不加权情况下,均匀划分区域方法的精确度略高于非均匀划分区域,但是带来了整体计算量的明显上升。按照均匀划分,该区域被分为 99 个均匀的网格,这样在每次定位时,必须把当前的观测向量同 99 组数据依次对比,而本文所提算法只需同 13 组数据进行对比,减少了约 87% 计算量,提高了定位的实时性。为了实现定位精确度与算法复杂度的良好折中,本算法采用非均匀划分区域的方法。

图7是在服务基站不稳定的区域(区域 5,7 和 10)下,分别对比了文献[11]算法、本文所提算法和传统的 Cell-ID 定位性能。可以看出,文献[11]的算法以 87% 的概率定位误差在 60 m 内,以 95% 的概率定位误差在 122 m 内,而所提算法以 67% 的概率定位误差在 73 m 内,以 95% 的概率定位误差在 137 m 内。因此在服务基站不稳定的情况下,所提算法的定位性能略差于文献[11]算法,这是因为在多个小区边缘时,服务基站可能会在多个基站之间频繁切换,而本文的算法提高了观测向量中服务基站相似度的权重,因而可能会影响到其他相邻区域。

与文献[11]的算法相比,本文所提算法大大减少了对比的次数,降低了复杂度,对于服务小区稳定的区域,定位精确度明显提高,代价是在服务小区不稳定的区域内,定位精确度略微下降。

### 3 结论

本文主要研究了蜂窝网定位算法,通过实际测试,提出一种新的 Cell-ID 定位方法。首先合理划分区域并建立数据库,然后分别给服务基站和相邻基站不同的权值求出似然向量,在此基础上,利用 HMM 方法获得移动台的运动状态,最后通过最大后验概率准则求出移动台当前所在的区域,并用区域的几何中心代表移动台的真实位置。实验结果表明,与文献[11]算法比较,所提算法复杂度低,在服务小区稳定的区域内定位精确度明显提高,达到了 FCC 的定位要求。

#### 参考文献:

- [1] WANG W D,ZHU Q X. RSS-based Monte Carlo localization for mobile sensor networks[J]. IET Communications, 2008, 2(5):673-681.
- [2] Fang Shih-Hau, Lin Tsung-Nan. A dynamic system approach for radio location fingerprinting in wireless local area networks[J]. IEEE Trans. on Wireless Commun., 2010,58(4):1020-1025.
- [3] 宋婉甜,李智. 一种新的无线传感网络三维定位方法[J]. 信息与电子工程, 2012,9(3):257-261. (SONG Wantian,LI Zhi. A new three-dimensional localization method for WSN[J]. Information and Electronic Engineering, 2012,9(3):257-261.)

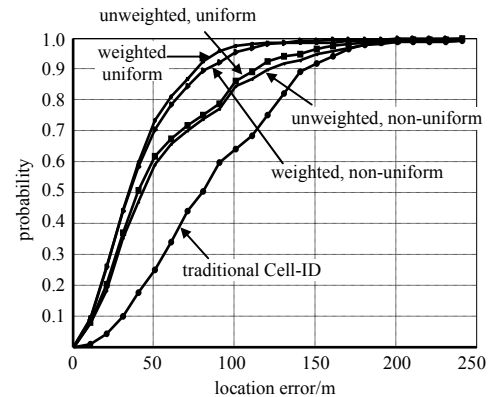


Fig.6 CDF of error when service base is stable  
图6 服务基站稳定区域误差 CDF 图

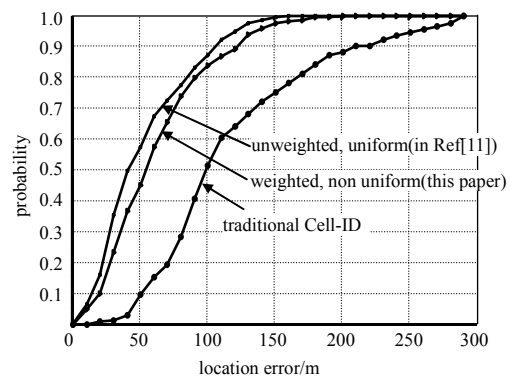


Fig.7 CDF of error when service base is unstable  
图7 服务基站不稳定区域误差 CDF 图

- [4] 杨萍,李智. 一种基于 RSSI 的权重值选择及加权定位算法[J]. 信息与电子工程, 2012,9(2):148–151. (YANG Ping,LI Zhi. Weight-select and weighted location algorithm based on RSSI[J]. Information and Electronic Engineering, 2012,(2): 148–151.)
- [5] Güvenç Ismail,Chong Chiachin. A Survey on TOA Based Wireless Localization and NLOS Mitigation Techniques[J]. Communications Surveys & Tutorials, 2009,11(3):107–124.
- [6] GONG Fuxiang,QING Wang,ZHANG Xiaoguo. A new distance based algorithm for TDOA localization in cellular networks[C]// 2010 IEEE Computer Science and Information Technology. Chengdu:[s.n.], 2010:502–505.
- [7] FAN Dandan,ZHANG Ruyun,JIN Liang. Extended Kalman Filter Location and Tracking Method Using Single Base Station in Cellular Network[C]// 2007 Wireless Communications,Networking and Mobile Computing. WiCom:[s.n.], 2007:740–742.
- [8] Klukas R,Fattouche M. Line-of-sight angle of arrival estimation in the outdoor multipath environment[J]. IEEE Trans. on Veh. Technol., 1998,47(1):342–351.
- [9] CHEN Chienheng, SU Szulin, LU Chyuander. Geometrical positioning approached for mobile location estimation[C]// 2010 IEEE Information Management and Engineering(ICIME). Chengdu:[s.n.], 2010:274–278.
- [10] Taok A,Kandil N,Affes S,et al. Fingerprinting localization using ultra-wideband and neural networks[C]// 2007 ISSSE Signals, Systems and Electronics. Montreal,Que:[s.n.], 2007:529–532.
- [11] Bshara M,OrgunerU,Gustafsson F,et al. Robust Tracking in Cellular Networks Using HMM Filters and Cell-ID Measurement[J]. IEEE Trans. on Veh. Technol, 2011,60(3):1016–1024.

#### 作者简介:



竹 博(1988–), 男, 山西省万荣县人, 在读硕士研究生, 主要研究方向为无线与移动通信、蜂窝网定位与跟踪.email:zhuboapple@163.com.

王建辉(1983–), 男, 河南省南阳市人, 博士, 讲师, 主要研究方向为蜂窝网定位.

胡捍英(1961–), 男, 河南省南阳市人, 博士, 教授, 主要研究方向为移动通信.

王大鸣(1971–), 男, 辽宁省大连市人, 博士, 教授, 主要研究方向为移动通信.