

# 基于隐马尔科夫模型的LTE室外 指纹定位研究与实践

## Research and Practice of LTE Outdoor Fingerprint-location Based on Hidden Markov Model

赵伟<sup>1</sup>,王勇<sup>2</sup>,刘宏嘉<sup>1</sup>(1. 中国联通浙江分公司,浙江杭州 325800;2. 中国联合网络通信集团有限公司,北京 100033)  
Zhao Wei<sup>1</sup>, Wang Yong<sup>2</sup>, Liu Hongjia<sup>1</sup>(1. China Unicom Zhejiang Branch, Hangzhou 325800, China; 2. China United Network Communications Group Co., Ltd., Beijing 100033, China)

### 摘要:

综合采用全量详单(XDR)数据、互联网实时出行数据、隐马尔可夫模型(HMM)和维特比算法(Viterbi),提出一种新的指纹定位方法。在LTE号码回填和指纹定位基础上,首先在XDR数据特定OTT的URL中提取经纬度,然后结合互联网实时出行数据完成道路用户识别,最后利用隐马尔可夫模型和维特比算法完成用户定位修正和地图匹配,测试验证表明,新算法道路匹配率达到98%以上。

### 关键词:

隐马尔科夫模型;维特比算法;指纹定位;MDT  
doi:10.12045/j.issn.1007-3043.2019.11.005  
中图分类号:TN929.5  
文献标识码:A  
文章编号:1007-3043(2019)11-0020-04

### Abstract:

A new fingerprint localization algorithm is proposed by using XDR data, Real-time Internet travel data, HMM and Viterbi. Based on LTE number backfill and fingerprint location, firstly it extracts latitude and longitude data from URL of specific OTT of XDR data, and then combines the real-time travel data to complete the road user identification. Finally, using hidden Markov model and Viterbi algorithm, it completes user location correction and map matching. Test results show that the road matching rate of the new algorithm is over 98%.

### Keywords:

Hidden markov model; Viterbi; Fingerprint location; Minimization of drive-tests

引用格式:赵伟,王勇,刘宏嘉. 基于隐马尔科夫模型的LTE室外指纹定位研究与实践[J]. 邮电设计技术,2019(11):20-23.

## 0 前言

基于接收的信号强度指示(RSSI)的LTE移动通信用户指纹定位和地图匹配已成为研究热点,然而由于信号瑞利衰落、频繁切换、上下行干扰等因素会造成定位数据误差大、无效点较多、路网匹配难等问题。该方法综合基于深度包检测技术(DPI)的OTT(互联网企业利用电信运营商的网络向用户提供各种应用服务)上报经纬度、互联网路网及拥堵数据,结合隐马尔科夫模型(HMM),较大幅度提升指纹定位地图匹配准确性问题,可用于道路用户分离及用户精准定位。

收稿日期:2019-08-26

## 1 研究背景

如今,用户大量使用基于位置服务(LBS)的APP,并和服务供应商共享他们的位置数据。与此同时,基于用户级感知保障、投诉定界、信息安全的需要,电信运营商在传统IP数据包检测技术上,增加了对应用层数据的应用协议识别、数据包内容检测与深度解码,并采用基于RSSI指纹定位技术完成LTE网络全量MR指纹定位和号码回填。但由于位置数据存在一定偏差,并且这些数据只能与电子地图匹配后才能应用于上述工作。如何能够快速、有效地将轨迹点投影到在线地图中是个难点。

### 1.1 基于RSSI的指纹定位难点

按照使用场景的差异,定位技术主要分为室外定位及室内技术(本文仅研究室内定位技术)。GPS及AGPS定位系统在室外定位误差达到10m以内,但对手机电池的消耗也比其他定位方法高近10倍。基于RSSI指纹定位也是室外定位的主要质量标准之一。

基于RSSI指纹定位分成基于测距及非测距定位两大类,基于测距定位具体的方式有抵达时间(TOA)、抵达时间差(TDOA)、抵达角(AOA)等。基于非测距定位主要有接收信号强度指示RSSI,一般又分为距离路径衰减模型法、高精度3D地图仿真和位置指纹匹配法等。上述技术有些需要专门设备,成本较大,难以大规模推广应用。有些易受环境、信号等干扰影响定位效果。

近些年,一些学者将K-最近邻、人工神经网络、决策树、贝叶斯等机器学习引入定位中。上述技术仍存在K值选择、大量训练数据获取及训练成本、绝对独立特征选取等难题。

### 1.2 地图匹配的难点

大多数电子地图相匹配算法能划分4个类型:几何图形算法、网络拓扑算法、可能性算法以及先进算法。几何图形遗传算法只充分考虑道路的形状而忽略了路段之间的关联性。概率算法在采样点位置定义置信区域并考虑不同导航传感器传输数据模式的影响。先进算法充分结合拓扑和概率,典型算法有卡尔曼滤波器、模糊逻辑、多元假设和隐马尔可夫模型。然而地图匹配需要经常考虑复杂路网下的双向交通情况。

## 2 道路用户指纹匹配地图的算法研究

### 2.1 LTE 全量MR 号码回填

XDR数据与MR数据的关联主要基于MME\_UE\_S1AP\_ID和MME Group ID将同一时间、同一地点和同一用户的XDR数据和MR数据联系起来,并回填必要的信令字段和业务字段。一般分为MME\_MR关联、HTTP\_MR关联和COMMON\_MR关联3种主要方式。每种关联的不同之处在于XDR数据各不相同,分别是MME数据、HTTP数据以及S1\_COMMON数据。一般分为预处理—关联—回填3个主要过程。

图1示出的是LTE号码回填示意图。

a) 预处理。预处理部分主要为数据清洗和数据处理。先将CELL号异常、开始时间不合理、结束时间不合理等MR无效数据进行去除,只保留有效的MR数据,以此进一步提高数据关联的工作效率以及数据高效性。

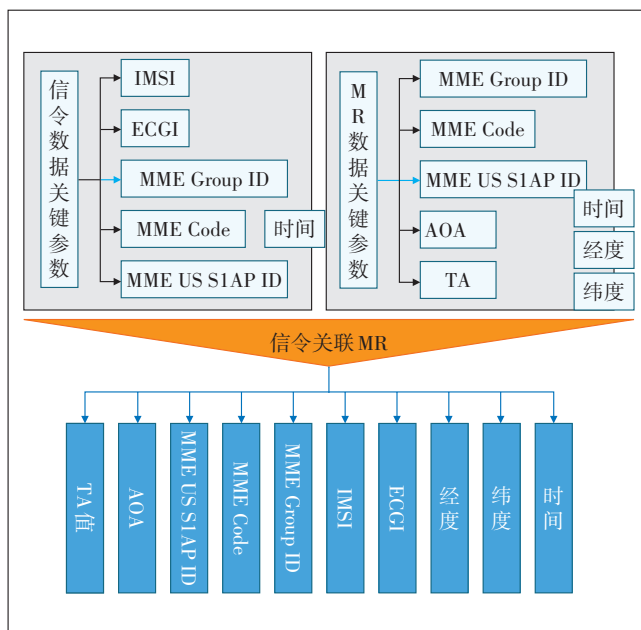


图1 LTE号码回填示意图

b) 关联。关联部分包括2个子过程:第1个过程是使用MME\_UE\_S1AP\_ID和MME Group ID 2个关键字段,将站号相同、MME\_UE\_S1AP\_ID相同、XDR与MR的时间差最小的(一般设置为5min)XDR数据作为可信的关联记录,对MR数据进行回填;第2个过程是针对第1个过程中没有关联的记录,基于MR数据的开始、结束时间,使用时间滑窗(经验值一般为20min)进行向前、向后滑窗搜索,合并,去重,得到XDR数据与MR数据的关联记录。

c) 回填。依照关联部分获得的关联记录,把一一对应的MR数据中的精确位置信息提交到存贮XDR感知数据的表中,顺利完成XDR数据以及MR数据的关联工作。

### 2.2 LTE 全量MR 指纹定位

根据3GPP协议规定,智能手机APP客户端在规定周期类上报当前小区的信号强度样本数值和经纬度,并且在不同位置上报RSRP样本数值也不相同,这种信息的集合,定义为指纹库。手机上报能力与手机终端软硬件版本、运营商网络等因素相关,浙江联通LTE网络统计中约有4.5%的终端上报比例。指纹库中的数据来源还包括CQT拨打测试数据、DT路测数据、最小化路测(MDT)数据以及OTT数据等。由于用户运动随机性和局限性,指纹库中的数据无法覆盖所有的区域,所以需要利用这些准确的数据来训练传播模型,获得最逼近真实无线环境的传播模型,利用该模

型计算所有栅格里面各小区的强度,然后补充到指纹库中,使指纹库中的数据可以覆盖全网区域。

图2示出的是基于KNN的MR指纹定位示意图。

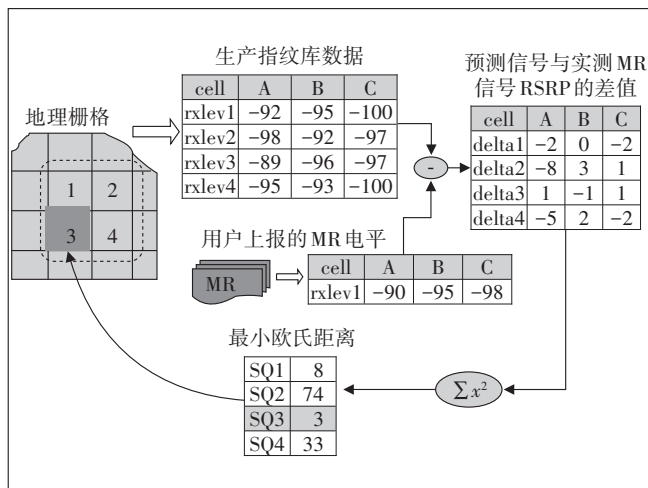


图2 基于KNN的MR指纹定位示意图

### 3 电子地图获取与解析

基于指纹定位的全量MR数据与地图匹配时,需要与路网数据结合,并投影到相应的路段。本文使用OpenGIS标准下的OpenStreetMap(OSM),地图格式为XML(eXtensible Markup Language),地图数据处理流程如下:

a) 解析路网数据,解析从OSM下载的XML格式的地图数据,解析出来的地图数据主要包括该块地图区域的点(node)、路(ways)、关系(relation)以及边界信息。这4种元素构成了整个地图画面。其中,node定义为空间中节点的位置,ways定义为路段或者区域,relation定义为元素之间的关系。

b) 获得路网中的道路级别信息标签的数值,譬如铁路、主干线、匝道等,初始化道路名称、道路等级信息及对应道路限制速度。

### 4 互联网出行数据获取

为进一步修正电子地图中道路等级速度限制,并充分考虑道路实时拥堵对速度的影响。基于高德API接口,按15 min粒度抓取高德地图道路拥堵度,并按时间、道路,拥堵情况存储,作为道路等级信息标签的速度修正依据并录入数据库。例如:绿色路段为道路限制速度的100%,黄色路段为拥堵路段,速度为道路速度限制速度的75%,红色路段或者深红色路段为拥堵路段,速度为道路限制速度的55%。

## 5 定位数据预处理

在电子地图匹配前,应当对定位数据进行预处理,以提高匹配效率及有效性。具体步骤如下:

a) 选取疑似道路用户指纹定位数据。在MR全量号码回填及指纹定位基础上,进一步地基于DPI单据S1\_HTTP中HOST链接过滤出特定OTT的download\_http\_content和upload\_http\_content内容及用户标识。例如:Host为am.xiaojukeji.com的用户可以还原为一个滴滴打车司机端或者乘客端真实用户,从关联的MRO数据uelongitude和uelatitude字段,以及S1\_HTTP的Download\_http\_content和Upload\_http\_content字段中提取该用户经纬度。进一步通过经纬度转换可以获得高置信度经纬度。同时,该用户全量MR指纹定位数据也可能包含AGPS等经纬度。

b) 基于互联网出行数据修正的道路速度过滤。在基于Host过滤的道路用户选取基础上,进一步利用定位经纬度值及时间戳,得到任意两点间速度值。通过与基于互联网出行数据修正的道路速度比较,如果速度大于阈值,则该用户不可能为道路用户。

## 6 隐马尔科夫模型构建

### 6.1 观测概率

在基于HMM方法中,地图匹配中每一个定位点与位于预设误差半径内的(基于浙江指纹定位结果置信水平,主城区预设误差半径为50 m)所有候选路段关联。每个定位点被看作是一个观测状态,并且每个候选路段都表示为一个隐藏状态。更具体地说,一个隐藏状态代表一个候选点,即候选路段上与观测位置最接近的点。每一个隐藏状态被赋予一个观测概率,它表示每个隐藏状态表示的候选点是真实位置时所产生的观测状态的条件概率。观测概率取决于自身与定位点之间的距离。直观地看,离定位点较远的候选点观测概率更低。在真实状态下,定位点与候选点之间的距离存在测量误差,一般假设为零均值高斯分布。对于给定的定位点 $O_i$ 与候选点 $C_{i,i}$ ,则观测概率为 $p(O_i | c_{i,i})$ 为

$$p(O_i | c_{i,i}) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{g(O_i, C_{i,i})^2}{2\sigma^2}} \quad (1)$$

式中:

$\sigma$ ——测量距离标准差

$g(O_i, C_{i,i})$ ——定位点 $O_i$ 与候选点 $C_{i,i}$ 大圆距离,

在低纬度较短距离内,可以近似为二维平面上的欧式距离。

## 6.2 转移概率

基于 Newson 和 Krumm 提出的转移概率模型受制于候选点相互之间的路径距离与定位点之间大圆距离的差值,在此基础上,进一步考虑相邻指纹定位候选点的时间间隔,以及候选点之间大圆距离的差值,基于道路网络上2点之间的行驶距离时间最短的 Dijkstra 理论,可以得到新的转移概率。

$$y_m = \frac{\|d(C_{t,i}, C_{t+1,j}) - g(C_{t,i}, C_{t+1,j})\|}{\Delta t} e^{-\beta y_m} \quad (2)$$

$$y_m = \frac{\|d(C_{t,i}, C_{t+1,j}) - g(C_{t,i}, C_{t+1,j})\|}{\Delta t} \quad (3)$$

式中:

$d(C_{t,i}, C_{t+1,j})$ ——候选点之间的路径距离

$g(C_{t,i}, C_{t+1,j})$ ——候选点之间的大圆距离

$i^*$ 和 $j^*$ ——路网道路

$\Delta t$ ——候选点之间的时间间隔

$\beta$ ——描述路径距离和大圆的差值

$\beta$ 评估方法为:

$$\beta = \frac{\ln(2)}{\text{median}_i \{ \| [g(O_t, C_{t+1}) - d(C_t, i^*, C_{t+1}, j^*)] \| \}} \quad (4)$$

## 7 维特比算法轨迹匹配

真实世界中,道路用户通常根据目的地以及始发地优先选择费时尽可能少的路线,则整体概率 $P(O, C)$ 最大的轨迹即为所求移动轨迹。

$$P(O, C) = P(c_0|o_0) \prod_{i=1, \dots, n} [p(c_i|o_i) p(c_i|c_{i-1})] \quad (5)$$

式中:

$P(c_0|o_0)$ ——初始状态概率

在隐马尔科夫分析模型中,借助维特比算法测算出概率最大的隐藏序列,即最优化的匹配轨迹。其在每一步的每个选择都储存了每一个历史步骤到当前步骤的最小总代价(或者最大价值)以及现今代价的状况下每一个历史步骤的选择。根据整个定位数据序列中各观测概率及转移概率,利用维特比算法来求解移动轨迹。

## 8 算法验证及效果

在完成4G号码回填和指纹定位的基础上,完成基于隐马尔科夫模型的道路用户指纹定位与在线地图匹配理论研究,并在浙江联通网络大数据平台中完成杭州市全量道路虚拟路测功能搭建。经过理论研究和

DT测试结果匹配,道路匹配率达到98%以上。

## 9 结束语

本文在LTE号码回填和指纹定位的基础上,通过指定HOST筛选、OTT经纬度还原,互联网电子地图、互联网出行数据获取完成LTE道路用户识别、提取,并基于隐马尔科夫模型和维特比算法完成道路用户地图匹配,完成用户指纹定位经纬度二次修正。在此基础上,完成杭州市区道路虚拟路测,经验证,道路匹配率达到98%以上。

## 参考文献:

- [1] ZANDBERGEN P A. Accuracy of iPhone Locations : A Comparison of Assisted GPS, WiFi and Cellular Positioning [J]. Transactions in Gis, 2009, 13(s1):5-25.
- [2] THIAGARAJAN A, RAVINDRANATH L, BALAKRISHNAN H, et al. Accurate, Low-Energy Trajectory Mapping for Mobile Devices [J]. Usenix Association, 2011(1):267-280.
- [3] TSUDA Y, KONG Q, MAEKAWA T. Detecting and correcting WiFi positioning errors [C]//ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM, 2013:777-786.
- [4] 刘春燕,王坚. 基于几何聚类指纹库的约束KNN室内定位模型 [J]. 武汉大学学报:信息科学版, 2014, 39(11):1287-1292.
- [5] 张会清,石晓伟,邓贵华,等. 基于BP神经网络和泰勒级数的室内定位算法研究 [J]. 电子学报, 2012, 40(9):1876-1879.
- [6] 刘侃,张伟,张伟东,等. 一种基于深度神经网络的无线定位方法 [J]. 计算机工程, 2016, 42(7):82-85.
- [7] LIAO J K, CHIANG K W, TSAI G J, et al. A low complexity map-aided Fuzzy Decision Tree for pedestrian indoor/outdoor navigation using smartphone [C]//International Conference on Indoor Positioning and Indoor Navigation, 2016:1-8.
- [8] 刘欢,黄丽,杨晓,等. 一种贝叶斯优化RSSI和ILS的室内定位算法 [J]. 中国科技论文, 2015(20):2377-2381.
- [9] GREENFELD J S. Matching GPS observations to locations on a digital map [C]//Transportation Research Board 81st Annual Meeting, 2002.
- [10] ZHENG Y. Trajectory data mining: an overview [J]. ACM Transactions on Intelligent Systems & Technology, 2015, 6(3):1-41.
- [11] HUANG L, POPIC V, BATZOGLOU S. Short read alignment with populations of genomes [J]. Bioinformatics, 2013, 29(13):i361-i370.
- [12] 马骁骏,李焯. 基于隐马尔科夫模型的基站定位地图匹配算法研究 [J]. 中国水运(下半月), 2018, 18(7):70-72, 202.

### 作者简介:

赵伟,毕业于香港理工大学,工程师,硕士,主要研究方向为网络优化、大数据及人工智能在移动通信中运用;王勇,毕业于北京大学,高级工程师,硕士,主要研究方向为网络优化;刘宏嘉,毕业于哈尔滨工程大学,工程师,学士,主要研究方向为网络优化。