

Research on Key Technologies of Fast Retrieval Feedback System

Xuqing Chai

College of Computer and Information Engineering of Henan Normal University, Xinxing Henan
Email: cxq@htu.edu.cn

Received: Nov. 28th, 2019; accepted: Dec. 12th, 2019; published: Dec. 19th, 2019

Abstract

In the field of mobile communication, when users make a phone complaint, especially about the quality of communication, how to give reasonable and acceptable answers to users' problems quickly and accurately, and how to follow up and solve users' problems in the back-end, and how to improve customer satisfaction are the key concerns of each operator. In this paper, a fast retrieval and feedback system for user complaints is proposed. Firstly, using the data based on XDR, MR OTT, MDT and other data, using the location algorithm based on fingerprint database, the grid location is carried out to gather the poor quality areas. Next, the network alarm data and the information of station interruption are aggregated, and the site coverage evaluation algorithm is used to form the quality difference area. The quality difference area calculated by the two algorithms is compared at the grid level, and the overlapping area is marked, and finally a query database is formed. When receiving a complaint from a user, the attendant can query and retrieve it in the query database according to the address description of the complaint, and the system will feed back the label of the complaint area, so as to get the actual situation of the complaint area and explain it to the customer. Finally, the accuracy of the key location algorithm proposed in this paper is verified by performance test experiments, and the practicability of this method is demonstrated.

Keywords

Retrieval, Feedback, Location Algorithm, Aggregation Algorithm

快速检索反馈系统关键技术研究

柴旭清

河南师范大学计算机与信息工程学院, 河南 新乡
Email: cxq@htu.edu.cn

收稿日期: 2019年11月28日; 录用日期: 2019年12月12日; 发布日期: 2019年12月19日

摘要

在移动通讯领域, 在用户进行电话投诉时, 特别是对通信的质量进行投诉时, 如何快速而准确的对用户的问题给予合理的和可接受的解答, 并对用户问题在后端进行跟进和解决, 提高客户满意度是每个运营商重点关注的部分。本文提出一种应用于用户投诉的快速检索和反馈系统, 首先, 利用基于测量报告、MDT最小化路测等数据利用基于指纹库的定位算法, 进行栅格化的定位, 汇聚出质差区域。接下来, 将网络告警数据、断站信息进行聚合, 利用栅格化覆盖评估算法, 形成质差区域, 将两个算法计算出的质差区域在栅格级别进行对比, 重叠区域进行标记, 最终形成查询库。在接到用户投诉时, 话务人员可以根据用户投诉的地址描述, 在查询库中进行查询, 检索, 系统会反馈回投诉区域的标签, 就能够得到投诉区域的实际情况, 给客户以解释。最后, 本文通过性能测试实验, 验证了本文提出的关键定位算法准确性, 展示了该方法的实用性。

关键词

检索, 反馈, 指纹库, 定位算法, 汇聚算法

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

随着移动通讯技术和国家经济的发展, 移动终端的应用已经非常普及了。每天数以亿记终端用户进行通话、上网、看视频等各类业务。由于通信网络不可能十全十美, 总存在这样那样的问题, 当用户遇到通信质量问题时, 第一时间想到的就是向运营商进行投诉, 但是由于接线话务人员非专业的技术人员, 对网络中出现的各种问题并不能给用户解释清楚, 而且对于用户的处理时限等诉求也不能进行有效答复, 就容易给用户造成一种感觉, 客服总是在推脱责任, 并没有真正关心用户的利益等。

对于话务人员而言, 他们对通信网络状态信息的掌握一般仅仅是某些站点拆除了、故障了、断电了等最基本的信息, 而且这些信息也可能由于内部流程的滞后性, 也可能不能随时掌握。所以当用户投诉时, 客服人员一般都会答复用户所在的位置站点故障、没站等原因导致的问题。而实际上网络也可能因为外部干扰、网络内部干扰等因素导致信号强度无问题而信号质量存在问题, 此时用户就会非常的不满, 认为客服纯粹是在进行搪塞, 并没有真正的想为用户解决问题。

2. 概述

随着计算机数据存储能力、计算能力的提升, 为海量的 MR (Measurement Report, 测量报告)、MDT (Minimization Drive Test, 最小化路测)等数据的分析提供了有力的工具[1]。正常情况下 MDT 数据是进行网络分析的最直接的数据(MDT 数据携带用户业务时的信号强度以及定位信息), 但是由于国家对个人隐私的保护, 运营商并不能对 MDT 数据进行全量的采集和分析[2]。因此只能利用 MR 数据进行分析, 通常情况下的数据分析只能进行整网范围的指标分析, 并不具备地理化的能力[3]。

本文提出一种基于指纹库的定位算法, 将 MDT 数据进行整合作为指纹数据, 对 MR 数据进行地理化的栅格定位, 在 MDT 指纹数据比较稀疏的区域, 利用 MDT 数据作为输入数据, 对传播模型进行训练,

计算栅格覆盖指标,作为栅格“指纹”,分别从网络信号强度和网络信号质量两个角度进行分析,定位,对于质差的栅格进行标记[4]。将其他可能影响网络覆盖的网络告警信息、断站信息等进行整合后,按照无线信号传播算法进行影响区域的栅格化估计,打标签。将每个栅格的标签进行对比分析,告警信息的标签栅格与信号强度质差栅格标签进行合并,统一为站点故障标签,其他标签不变。将栅格库作为客服人员进行查询,检索的资源库。当客服人员接到用户投诉时,依据用户描述的投诉地点,在栅格化的地图上检索,查看投诉区域的栅格标签,就可以根据标签内容给用户以答复,并且在栅格上进行标记,后台技术人员看到客服标记后就知道该区域出现了用户投诉,需要进行优化处理。

3. 算法

MR 数据就是测量报告数据,主要携带的是服务小区的 RSRP (Reference Signal Receiving Power, 参考信号接收功率)、RSRQ (Reference Signal Receiving Quality, 参考信号直接质量)、AOA (Angle-of-Arrival, 信号到达角度)、TA (Timing Advance, 时间提前量)、以及邻小区的 RSRP 信息[5]; 这些信息是用户所在无线环境的真实体现,根据这些数据可以反映出真实的用户覆盖情况,是网络覆盖评估的重要数据。

3.1. 定位算法背景

基于 MR 数据的覆盖评估,就是利用 MR 数据所携带的信息,对采样点进行定位计算,对这些采样点进行统计,反映用户所处位置的覆盖情况;在 MR 数据所携带的信息中 RSRP、AOA、TA 等信息均可以用来进行采样点的定位,但不同的无线环境、公共参数的准确度、采样信息的不同均可能对定位精度产生影响。

3.2. 基于场强的定位算法

主要是根据信号强度与信号传输距离呈反比的关系,通过接收信号场强和已知信道衰落模型去估计 UE 的距离,根据多个邻区(≥ 3)算出的距离来定位用户[6]。这种定位算法的缺点在于,对基站天线方位角,高度等信息非常敏感;对传播模型非常敏感,传播模型与真实环境的偏差直接决定定位误差,由于多径效应的影响,定位精度较差,定位准确性还很大程度依赖于地图的准确性。

3.3. 基于角度(AOA)的定位算法

主要是接收天线根据收到信号的入射角度判断用户的方向,进而进行位置坐标的估计[7]。通过接收天线阵利用检测到的多个信号入射角,判断信号的到来方向,通过多个方向值,计算出用户的位置,这种定位算法在信号入射角的微小偏差会导致目标位置估计的较大偏差,定位精度得不到保证[7]。

3.4. 基于时延(TA)的定位算法

基站利用检测到的多个信号时延,建立定位方程组,进行目标位置的计算。该算法计算精度差,定位精度得不到保证。

3.5. 基于 AOA + TA 的定位算法

通过 TA 计算 UE 与服务小区的距离,再通过 AOA 计算角度,距离 + 角度唯一确定 UE 位置。这种混合算法的优点是实现简单,但是 MR 中规范 TA 最小单位为 16 Ts,单程约 78 米,单位上的精度误差就高达 80 米;加之多径,因此 TA 值本身就不准,因此误差会更大,AOA 在 MR 规范中单位虽然是 0.5 度,但是实际验证情况看,在非视距(NLOS)市区密集场景下,由于存在多径,该值准确性很差[8];因此 TA + AOA 的定位算法比较适用于农村开阔地场景,对于市区等室内场景比较的情况下并不适用[9]。

3.6. 基于指纹库的定位算法

传统的定位方法基于 TA + AOA 或者基于场强定位算法，前者取决于 AOA，TA 的准确性，后者完全取决于传播参数的选择以及仿真的准确性；准确性无法满足 LTE 覆盖评估精度要求[10]。因此本文提出了基于指纹库的定位算法。该算法引入了 MDT 数据变成训练样本指纹库；通过指纹库训练传播模型，从而提高定位的准确度。

3.6.1. 基于指纹库的定位算法流程(见图 1)

定位算法流程主要分为 3 部分：预处理过程、指纹匹配过程、以及结果修正过程；预处理过程重点是将 MDT 数据中的噪声数据进行清洗，以及不同设备厂家的数据格式进行规范化处理。指纹匹配过程重点是 MR 数据与指纹库中的指纹数据进行相似度计算，输出最相似指纹的经纬度，并计算出 MR 数据的定位经纬度。结果修正过程重点是 MR 采样点与其他工参数据进行关联，输出最终的展示数据，并且需要对输出结果进行过滤，对于个别的离群点进行过滤，删除。见图 1。

3.6.2. 预处理过程

MDT 通过数据清洗、处理，加入到指纹库中，再次通过传播模型训练加入到栅格指纹库中，形成综合的指纹库；其他的站点位置、高度、方位角等信息形成工参数据，形成工参数据库。

3.6.3. 指纹匹配过程

MR 数据经过数据清洗、处理，与指纹库中种的指纹数据通过定位引擎的计算，再进行关联、定位匹配，得到每个采样点的定位信息。

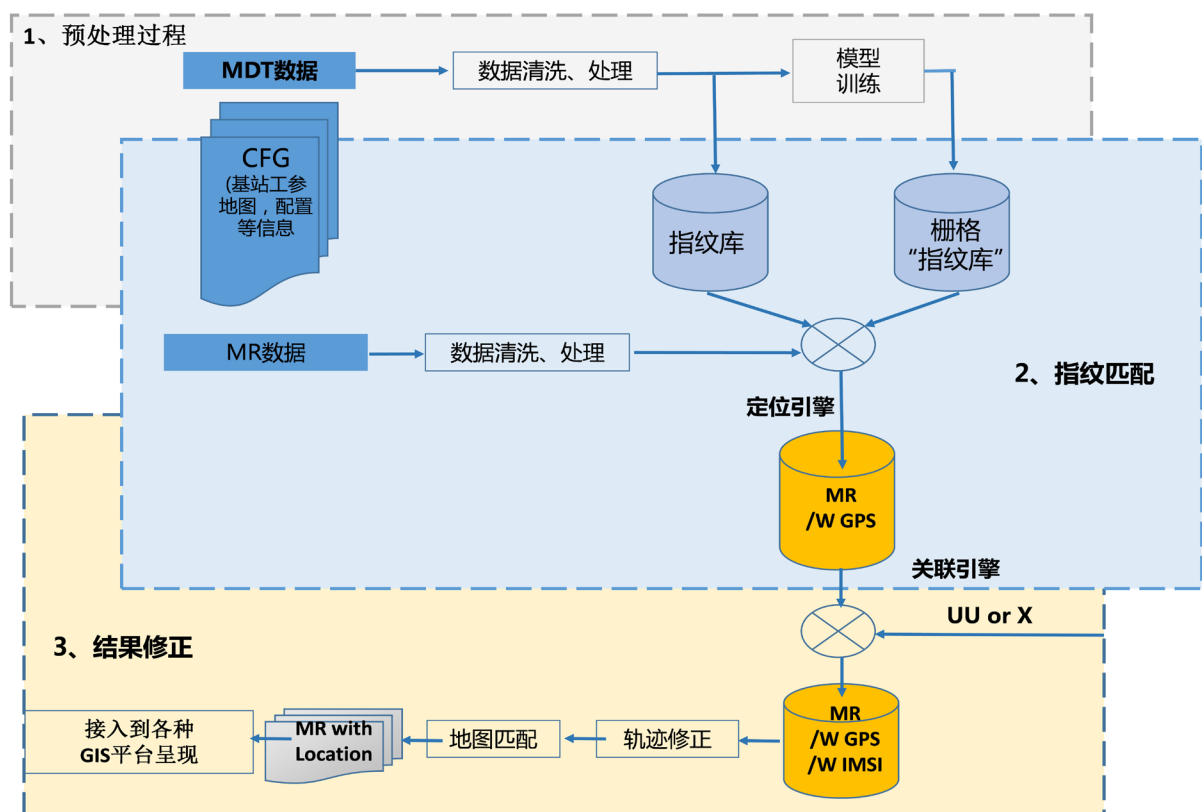


Figure 1. Location algorithm flow based on fingerprint database

图 1. 基于指纹库的定位算法流程

3.6.4. 结果修正

在得到每个采样点的位置信息后,通过轨迹修正、地图匹配,形成最终的 MR 定位数据,在各种 GIS (Geographic Information System, 地理信息系统)平台上进行展现;在 GIS 平台上,可以按照用户的意图,按照区域进行覆盖的统计、评估。

3.6.5. 基于指纹库的定位算法关键技术

关键技术 1: 指纹库构建。利用 MDT 数据以及由训练后的传播模型输出的栅格指纹库作为综合指纹库,将栅格指纹的格式转换为与 MDT 数据一致,作为一个采样点的形式存储,可以保证两类指纹的一致性;由于无线网络的基站是不断变化的(尤其在网络部署阶段),某个区域基站拓扑发生变化,只需要对相应区域的指纹库进行更新,基站拓扑发生变化的依据为工程实施部门给出的站点增补信息、基站工程参数发生变化的登记信息;指纹信息按照样本所在的小区存储,一旦基站工程参数发生变化便将该小区内的指纹进行全量替换,避免指纹数据的冗余;这样能够达到指纹数据与实际站点覆盖的场景一致,也会带来指纹库更新工作量加大的问题。指纹样本基本信息见表 1。

Table 1. Basic information of fingerprint data

表 1. 指纹数据基本信息

样本点 编号	样本点 经度	样本点 纬度	主服务 小区号	主服务小区的 小区 ID	邻区 1 编号	邻区 1 信号强度	邻区 2 编号	邻区 2 信号强度	邻区 3 编号	邻区 3 信号强度
-----------	-----------	-----------	------------	-----------------	------------	--------------	------------	--------------	------------	--------------

关键技术 2: 模型训练技术。尽管指纹库准确性高,但由于采集的局限性(比如农村区域存在 MDT 数据稀疏问题),因此 MDT 采样数据并不能保证所有区域都能够覆盖到;模型训练就利用局部的 MDT 数据对传播模型进行训练见图 2,获得最逼近真实无线环境的传播模型。指纹库数据越完备,模型训练越准确,指纹库模型训练后,再用这个模型计算 50*50 栅格里小区的强度,该输出结果作为栅格中心经纬度的信号强度,归为一个类似 MDT 的采样点,称为栅格“指纹库”,带引号是因为并非真实测量的。LTE 的无线传播模型采用 Okumura-Hata 模型:

$$L_p = 69.55 + 26.16 \lg f_c - 13.82 \lg h_{te} - \alpha(h_{re}) + (44.9 - 6.55 \lg h_{te}) \lg d + C_{cell} + C_{terrain}$$

各参数含义如下: f_c ——工作频率,单位为 MHz; h_{te} ——基站天线有效高度,单位 m,定义为基站天线实际海拔高度与基站沿传播方向向实际距离内的平均地面海拔高度之差,也就是 $h_{te} = h_{BS} - h_{ga}$; h_{re} ——移动台天线有效高度,单位 m,定义为移动台天线海拔高度; d ——基站天线与移动台天线之间的水平距离,单位 Km [10]。

关键技术 3: 指纹匹配技术。指纹匹配采用模式匹配的标准算法 KNN (K Nearest Neighborhood),选择 MR 与指纹库最“相似”的栅格,将待定位采样点的主服务小区以及各个邻区的信号强度作为特征向量,计算与指纹库中指纹的欧式距离,将计算得到的欧式距离按照从小到大排列,找到 K 个值最小的指纹,这 K 个指纹的中心坐标就定义为 MR 的位置(如果是 KNN 算法, $K = 3$),见图 2。

图中待定位的 MR 采样点主服务小区为 Cell A,邻区分别为 Cell B, Cell C, Cell D; 该采样点首先按照邻区号进行排序,与指纹 A、B、C、D 的邻区相符,然后分别与 4 个指纹点计算欧式距离,挑选最小的 3 个指纹,那么 3 个指纹的采样点的经纬度的中心即为待定位采样点的经纬度。

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

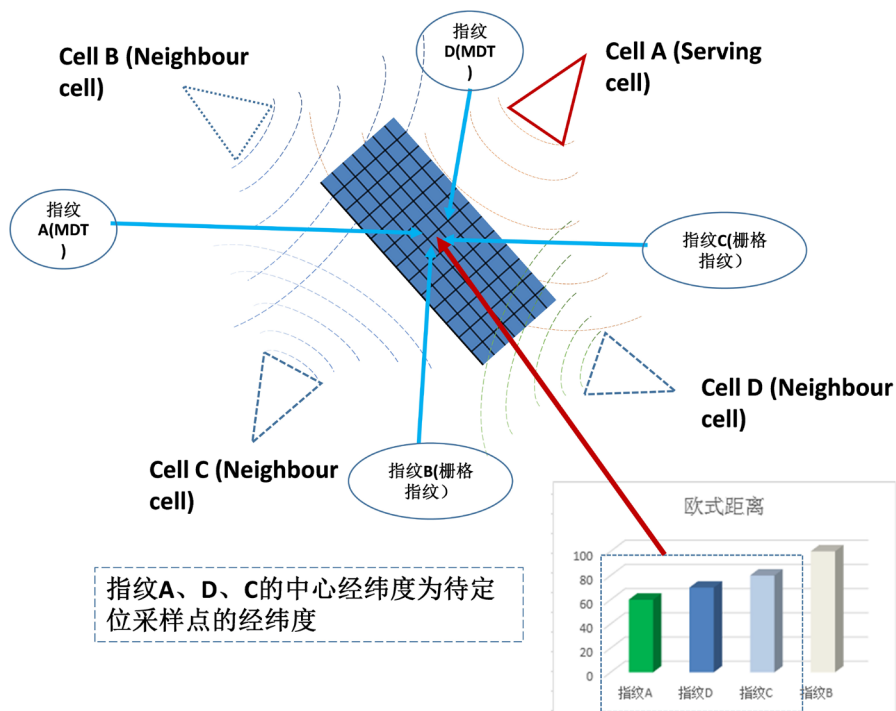


Figure 2. Fingerprint matching process
图 2. 指纹匹配过程

x 为待定位样本点的信号强度特征, y 为指纹样本的信号强度特征。 i 的取值可以是一个阈值, 也可以自动匹配至两个采样点的邻区号不一致时即终止。

4. 算法结果验证

在验证区域选择 64 个固定的点(随机分布在测试区域内)做测试验证, 在 0 到 400 米的区间内, 64 个固定的点的误差各不相同, 总的平均误差在 100 米左右, 见图 3。

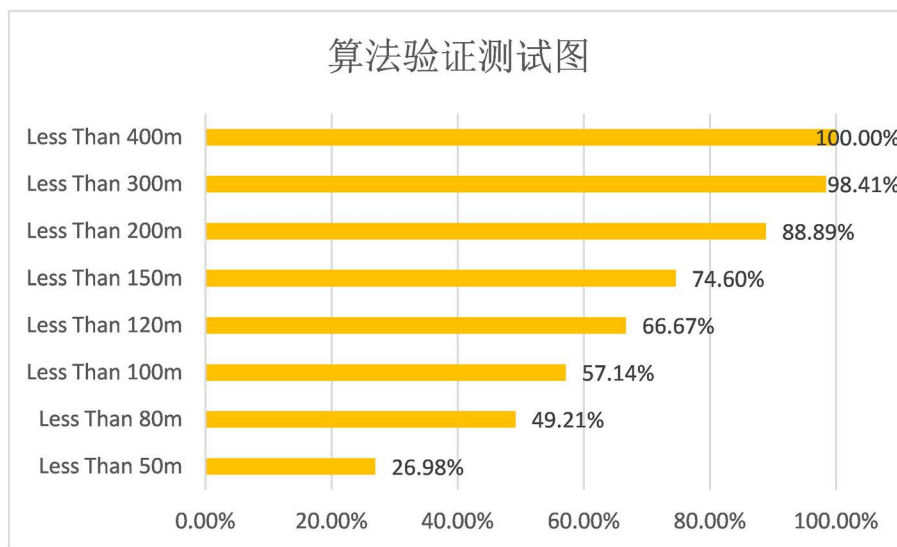


Figure 3. Algorithm validation test evaluation
图 3. 算法验证测试效果图

该算法与传统基于 AOA、TA 等算法相比具有精确度高, 适应场景高等特点; 对于网络覆盖的评估而言是一种较好的定位算法。

5. 结束语

本文描述了用户投诉的快速检索与反馈系统的基本结构与业务流程, 重点介绍了构成检索库的数据类型以及数据处理流程, 最终利用基于指纹数据的定位算法对 MR 数据进行定位处理, 以及其他告警数据与 MR 定位数据的关联, 最终形成了查询库; 整个系统设计架构能够使客服人员根据实际的网络质量情况对用户的投诉进行详细的回应, 并且可以在系统中对用户投诉区域进行标记, 后台技术人员也就能拿到准确的处理位置, 使处理更加具有针对性, 也更加有效。

基金项目

河南省科技攻关项目计划支持(项目编号: 192102310020, 182102210363); 河南省高等学校重点科研项目计划支持(项目编号: 17A520009); 河南省教师教育课程改革研究项目资助(项目编号: 2018-JSJYYB-020); 河南省社科联调研课题(项目编号: SKL-2018-771); 河南师范大学教育科学研究基金资助项目(项目编号: 2018JK10)。

参考文献

- [1] 冯雪元. 基于 MR 定位结合道路测试与电子地图的网络覆盖评估方法[J]. 中国新通信, 2019, 21(11): 54-55.
- [2] 徐珊珊, 余健, 王计斌. 基于 OTT 与 MR 数据的自适应 WKNN 室外定位方法[J]. 江苏通信, 2019, 35(1): 7-11
- [3] 董建, 范江, 王凌青. 利用 MR 定位精准优化 WCDMA 网络质量新方法[J]. 通信管理与技术, 2018(3): 21-23.
- [4] 方中圆, 宋巍, 顾翀. 基于 MR 定位的深度弱覆盖分布研究[J]. 数字通信世界, 2017(9): 137+197.
- [5] Chai, X.-Q., Dong, Y.-L. and Li, J.-F. (2016) Profit-Oriented Task Scheduling Algorithm in Hadoop Cluster. *EURASIP Journal on Embedded Systems*, 6. <https://doi.org/10.1186/s13639-016-0026-x>
- [6] Chai, X.-Q., Dong, Y.-L. and Li, J.-F. (2016) An Algorithm for Improved Similarity and Collaborative Filtering in Social Networks. *IJSSST (International Journal of Simulation: Systems, Science and Technology)*, 4.
- [7] 池刚毅, 商亮, 孙浩, 韩云波. 基于三维空间特征库的 MR 定位系统[J]. 电信工程技术与标准化, 2016, 29(3): 25-27.
- [8] 黄友亮. 基于 MR 的 LTE 网络结构优化分析应用[J]. 广东通信技术, 2016, 36(277): 50-54+83.
- [9] 柴旭清, 孙丽娜. 基于量子粒子群和 SARSA 算法的蜂窝网络信道分配[J]. 计算机测量与控制, 2015, 23(205): 312-314.
- [10] Chai, X.-Q., Liu, H.-T. and Dong, Y.-L. (2013) Switching Algorithms Based on Upstream Block Error Rate in TD-SCDMA. *JCIT (Journal of Convergence Information Technology)*, 8, 367-374.