

MDT 数据训练指纹定位模型的 误差分析及优化方法

Error Analysis and Optimization Method of Fingerprint Positioning Model Based on MDT Data Training

吕非彼¹, 许国平², 李纪华³, 朱佳佳¹, 张 琰¹, 只 璐¹, 王昭宁¹ (1. 中国联通研究院, 北京 100048; 2. 中国联合网络通信集团有限公司, 北京 100033; 3. 中国联通北京分公司, 北京 100032)

Lü Feibi¹, Xu Guoping², Li Jihua³, Zhu Jiajia¹, Zhang Yan¹, Zhi Lu¹, Wang Zhaoning¹ (1. China Unicom Research Institute, Beijing 100048, China; 2. China United Network Communications Group Co., Ltd., Beijing 100033, China; 3. China Unicom Beijing Branch, Beijing 100032, China)

摘 要:

无线指纹定位技术具有成本低、覆盖终端广、不触碰用户隐私等优点,是近年来运营商非常关注的定位技术之一。针对基于MDT数据训练的指纹定位模型存在较大误差的问题,基于现网数据定量分析了无线环境特征和GNSS位置信息偏差对指纹模型训练产生的影响。面向运营商维护优化应用的实际需求,提出并验证了根据不同场景差异化设置栅格大小以提升指纹模型定位准确性和实用性的方法。

关键词:

无线指纹; 定位; MDT; 特征工程

doi: 10.12045/j.issn.1007-3043.2025.11.003

文章编号: 1007-3043(2025)11-0012-06

中图分类号: TN915

文献标识码: A

开放科学(资源服务)标识码(OSID):



Abstract:

Wireless fingerprint positioning technology has the advantages of low cost, wide coverage of terminals, and without touching the user's privacy, which is one of the positioning technologies that operators have paid much attention to. To address the significant errors in fingerprint positioning models based on MDT data training, it quantitatively analyzes the impact of wireless environment characteristics and GNSS position information deviations on fingerprint model training using live network data. In response to the practical needs of operators in maintenance and optimization applications, a method is proposed and validated to enhance the accuracy and practicality of fingerprint models by setting differentiated grid sizes according to various scenarios.

Keywords:

Wireless fingerprint; Positioning; MDT; Feature engineering

引用格式: 吕非彼, 许国平, 李纪华, 等. MDT数据训练指纹定位模型的误差分析及优化方法[J]. 邮电设计技术, 2025(11): 12-17.

1 概述

在网络数字化、智能化转型的背景下,网络的智能运营技术已成为运营商研发的重点和热点方向之一。然而要实现全面的网络智能运营,精准的定位技术是不可或缺的。定位信息可以辅助网络评测、故障定位和规划优化等网络运营工作,同时面向客户服务

的定位需求也在日益增长。无论是导航、社交、广告还是服务推送等等,准确的位置信息都是提供个性化、定制化服务的基础。

目前移动网络中最常用的定位技术是卫星定位(Global Navigation Satellite System, GNSS),其具有良好的定位精度,但限于用户隐私、终端生态等因素,终端上报的位置信息量仍不足以支撑运营商的业务需求。对于运营商来说,寻找一种低成本且易于部署的定位技术至关重要。指纹定位作为一种有望满足这一需

收稿日期: 2025-09-26

求的定位技术,引起了行业的广泛关注。指纹技术利用无线信号在特定环境中的传播特征,如信号强度、延迟和多径效应等,构建了一个地理位置与无线信号特征之间的映射关系。相比传统定位技术,指纹定位有2个明显优势,一是它不依赖网络以外的硬件设备,降低了实施成本和运营复杂度;二是指纹定位可以实现室内位置定位,满足运营商对室内外综合定位的需求。

理论上通过机器学习的方法建立RF指纹能实现米级精度的定位^[1]。其中利用MDT数据作为指纹库模型的训练数据具有采集成本低、采集范围广、数据量充足等优势,是运营商最常用的训练数据来源^[2-6]。指纹定位技术也被写入了3GPP相关标准中^[7-8]。尽管在研究和实验环境中指纹定位模型已取得了良好成果,行业内也出现了众多指纹定位产品,但基于无线指纹定位的现网应用却非常有限,导致这一情况的原因主要有2个。一是现网无线环境的复杂性和时变性导致信号在时间和空间上不断变化,一个位置的“射频指纹”在何种时空尺度上稳定存在缺少科学的研究验证^[9-14]。二是用于模型训练的数据源质量缺乏检验,已有研究往往假设MDT或OTT等数据中提供的位置信息准确,并且模型验证也多采用抽样的形式进行。

本研究基于现网测试和MDT数据分析空口指标的时空特征,采用数据统计和特征工程等方法研究运营商MDT数据中的GNSS准确度,及其定位误差对训练指纹定位模型所产生的影响,最后提出了对基于数据清洗的指纹定位模型精度优化提升方案。

2 现网无线覆盖特征对MDT位置信息的影响

2.1 测试环境和网络配置说明

本研究的实验环境为某商务写字楼周边2 km范围,区域内存在较多地理场景,包括视距可见(LOS)天线的开阔位置、高层写字楼间的非视距(NLOS)位置、建筑拐角、茂密树丛位置等。本研究的目的是观察无线环境特征,不关注网络性能和负荷,因此测试业务均采用PING测试,定点和拉远测试采用每次Log采集10 min,间隔5 min重复多次的方式进行。周边具备4G和5G网络覆盖,4G MDT功能已开启,终端能够上报具备经纬度的MR信息。

2.2 室外场景无线覆盖特征分析

2.2.1 空旷区域定点覆盖特征分析

首先建立基准位置点,选择环境空旷、视线可直

接看到基站的位置进行定点测试,位置环境如图1所示。



图1 LOS定点测试

测试数据统计结果如表1所示,在该位置采集到1个邻区信号的比例达93%,2个邻区信号比例为21%,3个及以上邻区信号的比例为1%。定点测试中服务小区非常稳定,一直驻留在PCI=173的小区,主服务小区RSRP标准差为2.67。主服最强邻区有2个(PCI=355和PCI=354),说明这2个小区到定点位置的覆盖强度较为接近。

表1 定点测试数据分析

字段	最小值	最大值	平均值	标准差	偏度
ServingPCI	173	355	175.879 31	22.712 88	7.763 5
ServingRSRP	-89.5	-74.5	-79.919 83	2.677 59	-1.152 58
ServingRSRQ	-20.06	-7.31	-13.321 04	1.459 7	-0.939 36
ServingSINR	-8.3	16	3.918 6	3.912 24	-0.087 7
RSRP1	-89.5	-74.5	-79.832 4	2.495 65	-0.978 81
RSRP2	-94.56	-79.5	-85.834 98	2.179 66	-0.428 93
RSRP3	-97.25	-81.69	-87.583 44	2.258 01	-0.767 81
DeltaRSRP1_2	0	40	6.010 88	2.329 97	0.334 69
DeltaRSRPs_n1	-5.06	40	5.836 55	2.738 34	-0.610 41

对服务小区和主要邻区进行数据汇聚,分别对3个小区的RSRP、小区间RSRP差值以及各项指标进行相关性分析,统计如图2所示,可见邻区RSRP方差显

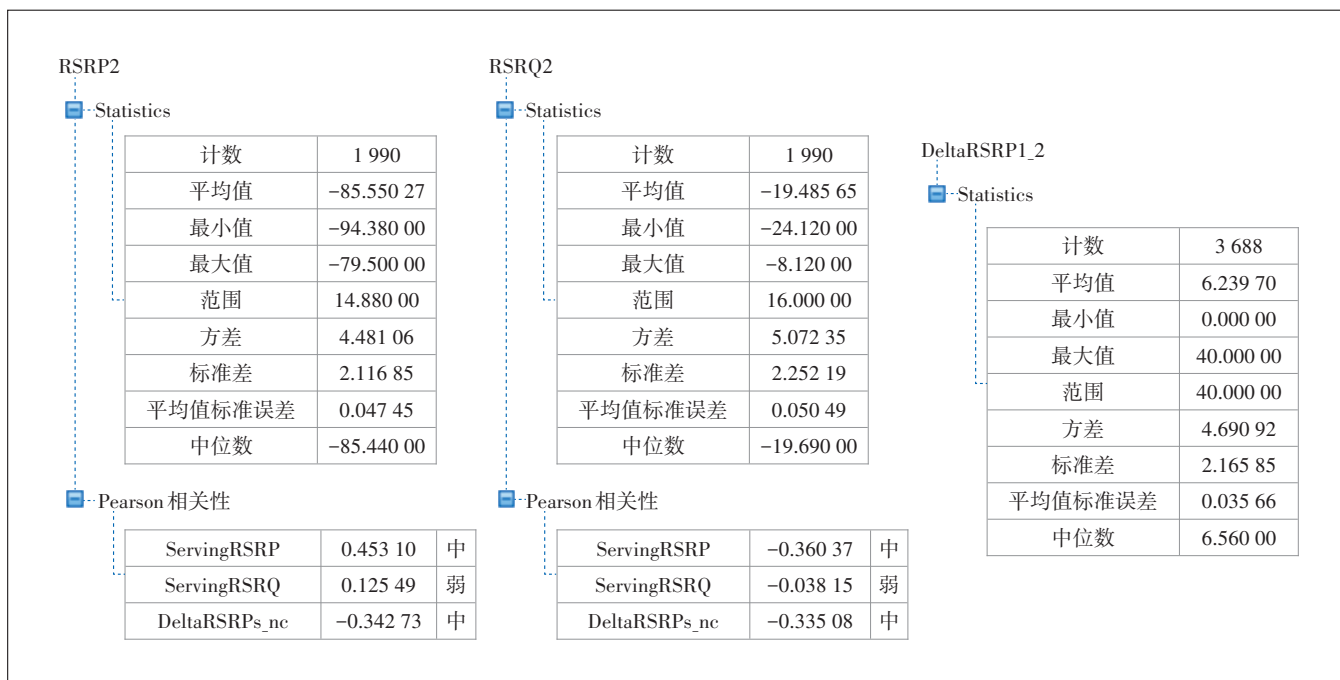


图2 定点测试邻区与服务小区间 RSRP 和 RSRQ 相关性

著大于服务小区。理论上同一位置下小区间 RSRP 差值相对稳定,主服小区和多个邻区间 RSRP 差值成为指纹定位模型的特征工程中最常用的构造特征。观察两两小区间 RSRP 差值统计特征,由图 2 可见主服小区和邻区 RSRP 存在中低度相关,但主服小区和邻区 RSRP 差值的方差与单小区的 RSRP 无显著差异。这表明 RSRP 差并非稳定、精确的 RF“指纹”特征。

2.2.2 空旷区域拉远测试数据分析

以第 2.2.1 节中的定点位置为基准 0 点,沿马路向南距基准点 5 m 位置设置一个测试点、10 m 位置再设置一个测试点,在这 3 个定点分别测试 10 min,对比距离与覆盖指标的关联程度。对 3 个点位所采集的数据进行统计分析,筛选出具备上述 3 个小区测量数据的采样点,设 $X=RSRP1$, $Y=RSRP2$, $Z=RSRP3$,进行 3D 可视化(见图 3)。

分析图 3 可知,0 m 位置的绿色点与 10 m 位置处的蓝色点有一定区分度,采用 KNN 分类准确率达 75%。但 5 m 位置的红色点与 0、10 m 位置的蓝绿点存在大量重叠,采用 KNN 分类准确率为 18%,这表明多维 RSRP 已较难实现 5 m 距离精度的准确定位。

2.2.3 LOS 和 NLOS 环境无线特征对比分析

为研究 NOLS 环境测试,选择基准测试位置西北侧的 2 栋高层写字楼中间进行测试数据采集,写字楼

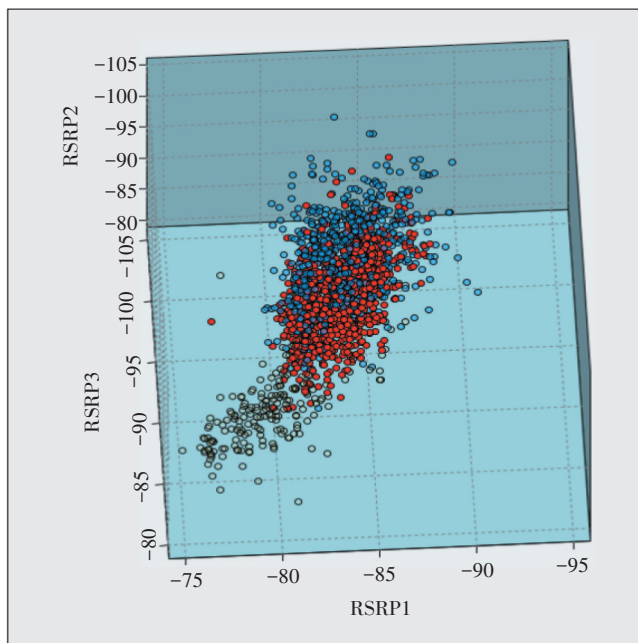


图3 拉远测试3个定点位置的 RSRP 3D 可视化示意

对周边天线的信号进行了直线遮挡(见图 4)。

数据分析结果如表 2 所示,服务小区 PCI=355 的 RSRP 标准差扩大为 4.2。由于测试位置接近建筑物,发现 GNSS 出现一定漂移,最大漂移约为 15 m。需要特别关注的是,GNSS 位置点分布并非是围绕真实测

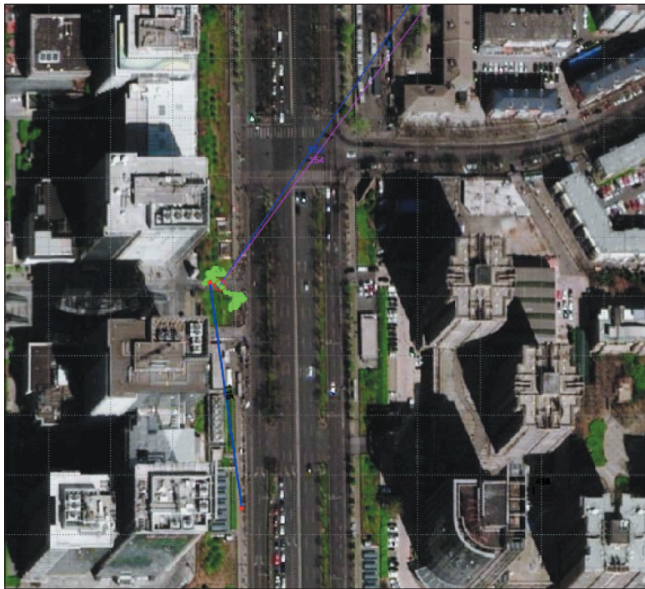


图4 NLOS环境定点测试位置

表2 NLOS环境定点测试数据统计

字段	最小值	最大值	平均值	标准差	偏度
DeltaRSRP1_2	0	20.69	5.452	3.293	0.789
DeltaRSRPs_n1	-3.88	20.69	5.396	3.385	0.642
DeltaRSRPs_nc	-6.28	40	3.387	3.72	1.054
ServingPCI	355	355	355	0	-
ServingRSRP	-87.25	-68.06	-78.6	4.233	0.633
ServingRSRQ	-16.19	-3.94	-11.6	2.518	0.208
ServingSINR	-0.9	19.2	8.78	3.501	0.326
ServingRSSI	-56.94	-34.25	-46.7	4.988	0.469

试位置的二维正态分布,而是呈现出一个“棒槌型”,也就是说指纹定位模型在训练过程中不能通过简单的平均或者离群值清洗来找到准确的真值。

2.2.4 树木环境覆盖特征分析

在研究树木环境下无线信号的特征时,选择的测试位置附近均为高3 m左右的榆树和梧桐,测试时正值夏季树叶茂盛。数据分析结果如表3所示,服务小区RSRP较为稳定但RSRQ波动较大,其余数据统计分布与定点测试差异不大。

2.2.5 室分定点覆盖特征分析

为研究室分信号覆盖特征,选择在写字楼内进行测试数据采集。为避免出现切换,测试位置选择服务小区RSRP超过-80 dBm的良好覆盖位置。由表4可知,同一位置信号最大波动幅度达17 dBm,但整体数据标准差仅为1.22,室分信号比室外RSRP稳定性好。

表3 树木环境定点测试数据统计

字段	最小值	最大值	平均值	标准差	偏度
ServingRSRP	-95.12	-79.31	-88.069 93	1.627 9	-0.022 83
ServingRSRQ	-15.56	-5.81	-8.906 57	2.079 7	-0.645 81
ServingSINR	3.8	18.8	12.323 75	2.360 55	-0.184 32
ServingRSSI	-70.5	-44.88	-61.988 1	2.657 8	1.130 45
DeltaRSRP1_2	1.62	18.12	8.888 05	2.207 64	-0.135 38
DeltaRSRPs_n1	1.62	18.12	8.888 05	2.207 64	-0.135 38
DeltaRSRPs_nc	-1.47	18.12	6.569 83	2.438 51	0.020 29

表4 室分环境定点测试数据统计

Stream	IE	Count	Avg	Min	Max	Std
Indoor	ServingRSRP	48 431	-73.45	-82.13	-65.44	1.22

对信号进行时序分析,在过滤离群值后,RSRP波动范围仍有10 dB。

2.2.6 区域移动性覆盖特征研究

在测试区域范围内进行移动性测试,数据统计特征与定点测试差别不大,但层3 MR的RSRP稳定性优于层1,标准差分别为7.63(层3)和5.17(层1)。但在移动性测试中,在建筑物周边会出现拐角效应。当主服务小区在建筑物拐角5 m范围内移动时,RSRP快速衰减约10 dBm,而邻区RSRP波动正常。由此可知,若栅格下地理环境复杂或存在物理阻挡,该栅格将存在多种覆盖特征,并且这种覆盖特征受环境影响远比位置影响更大。一个典型的问题是栅格边界的特征相似度远超栅格内外,这无疑极大影响了栅格指纹定位模型的定位准确度。

2.3 层3测量与层1测量差异分析

MDT上报的带有经纬度的RSRP(SS-Reference Signal Receiving Power)是通过MR信令发送的,因此与MR执行相同的测量配置参数,具备周期性和事件性测量方法。现网周期性MR/MDT测量一般设置为5~10 s。并且MR中通过参数可设置层三滤波,即对一定测量周期内的RSRP进行平滑处理,降低无线环境波动的影响。

通过数据分析研究层3与层1的RSRP测量的指标差异,定量验证层3滤波能在多大程度上平滑无线环境的波动。基于2.2节中采集的定点位置数据,在终端Log中层1的RSRP指标由终端芯片采集,时间间隔为TTI级,层3RSRP来自MR/MDT信令。在数据分析时,只筛选周期性MR与层1RSRP进行对比,分析

结果如表5所示,其中服务小区层3 RSRP标准差相比层1略有减少,其他特征差异较小。

表5 终端层1 RSRP与层3 RSRP对比分析

ServeringRSRP	最小值	最大值	平均值	标准差	偏度
层1	-89.5	-74.5	-79.919 83	2.677 59	-1.152 58
层3	-87.0	-75.0	-79.414 20	2.153 36	-1.221 02

2.4 室内定位数据质量研究分析

由于MDT上报的位置信息数据受终端GNSS性能影响,本文在室内采集数据以研究浅层室内(即靠近门、窗位置)的GNSS定位准确度。测试终端在写字楼一层大堂内无法连接GNSS,由室内走至室外后,在写字楼门外约3 m处可以锁定GNSS信号,此时再从室外逐步走回室内并逐步进入室内较深位置。实验发现在距离门窗5 m范围内仍能连接GNSS,MR信令正常上报LocationInformation信元,直到距门窗8 m位置时GNSS断开(见图5)。

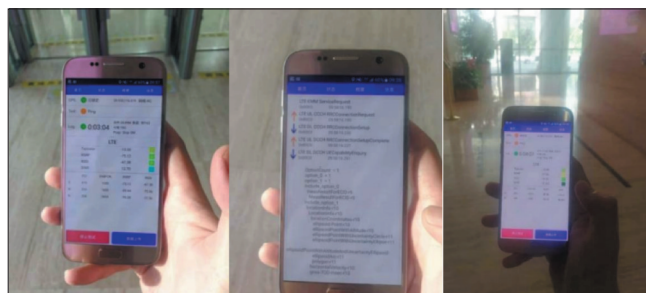


图5 室内一定深度下终端能够上报GNSS位置信息

对数据进行地理化分析,可见浅层室内位置MDT上报的GNSS定位漂移幅度较大,最小定位误差为55 m,最大误差已达110 m,平均定位误差为86 m。由此可见,在室内浅层区域的MDT数据作为指纹定位模型的训练数据输入将对模型的准确度造成较大影响。

3 主要结论与试点验证

3.1 MDT作为指纹模型训练数据的问题总结

上述研究分析的主要结论如下。

a) LOS定点信号稳定性:RSRP的 σ 在2 dB左右,主服邻区差值 σ 也在2 dB左右,最大最小值差在10 dB左右,GNSS定位稳定。

b) LOS拉远测试:位置相隔10 m以上,RF特征才有部分区分度。

c) 多径下信号稳定性:RSRP的 σ 在4 dB左右,

GNSS抖动在5 m以内。

d) 层3与层1 RSRP测量差异:定点位置RSRP分布基本一致,区域移动测试时层3滤波有一定平滑作用。

e) 穿透场景测试:树木只影响衰减,但RSRP稳定性与开阔点类似。

f) 建筑物拐角处快衰明显,15 m范围内信号衰减为10 dB。

g) 手机从室外进入室内5 m内仍能上报GNSS,但定位精度大幅下降,平均误差可达80 m以上。

基于上述分析结论,本文认为目前运营基于MDT数据训练无线指纹定位模型需要解决以下3个问题。

a) 由于无线环境的不稳定性,用户行为、位置分布、终端性能存在不确定性,依靠MDT数据训练RF指纹模型难以达到10 m以下的定位精度。在不引入新的辅助信号(如蓝牙或毫米波)的情况下,定位精度目标不宜过高。

b) 终端GNSS定位在浅层覆盖下存在定位漂移,MDT数据直接作为指纹定位模型的训练数据将导致模型准确度的不可控。不同场景下处于浅层覆盖或遮挡环境下的用户比例差异较大,因此应针对场景的地理特点对指纹模型进行差异化的调整优化。

c) 地理栅格的大小和边界将对指纹定位模型产生很大影响。一方面,无线环境的时变难以支持过小的栅格指纹;另一方面,在米级范围内覆盖特征受环境影响比覆盖距离更大,规则的地理栅格(正方形、六边形等)在边界附近必然存在高误差。而对大于100 m的栅格,指纹定位相比其他技术不存在优势。因此面向网络运维优化的实际需求,本文认为地理栅格划分在20~30 m较为合适,或采用不规则栅格绘制方法以提升定位准确度。

3.2 不同场景下指纹定位模型效果验证

为验证上述结论,本研究选择高校操场(开阔区域)、写字楼、公园和大型居民区4类典型场景,开展为期3个月的MDT数据采集,以前2个月的数据作为训练数据集,后1个月的数据作为验证数据集,训练指纹定位模型并验证不同场景、不同栅格大小对模型定位精度的影响。试点场景区域栅格数量最大不超过8 000,本研究选用了Xgboost算法将栅格定位问题转化为栅格多分类问题进行指纹定位模型训练。模型以栅格定位准确率作为评价指标,即每个MR数据模型给出的栅格是否是正确的栅格。

$$\text{栅格定位准确率 (Accuracy)} = \frac{\text{归类到正确栅格下的采样点数}}{\text{总采样点数}}$$

各场景模型定位精度验证数据如表6所示,模型在不同地理场景下展示出的性能差异较大,但采用20 m栅格参数的模型性能一直处于最优,10 m栅格在复杂环境中表现不佳。还需要说明的是,MDT数据本身是一种抽样数据,即抽取了能接收到GNSS信号的部分终端数据,但无线指纹定位模型最终应用对象是全量MR数据,相比试点中采用MDT数据作为模型验证集的模型效果,现网应用效果还会产生一定回退。基于试点结论,针对运营商运维优化场景,开阔场景采用20 m栅格可基于MDT数据直接训练应用指纹模型,而复杂场景应选择牺牲一定定位精度或者引入其他辅助定位技术以对GNSS数据不准的问题进行补偿或清洗。

表6 不同场景、不同栅格大小对指纹定位模型准确性的影响

场景	Accuracy/%			
	栅格大小 10 m	栅格大小 20 m	栅格大小 50 m	栅格大小 100 m
高校操场周边	69.3	85.6	84.2	88.7
商务写字楼区域	42.9	71.2	70.7	79.2
公园场景	68.5	82.2	75.0	90.2
大型居民区	48.9	65.8	55.9	71.4

4 结束语

针对无线指纹定位技术现网应用效果不佳的问题,本文通过现网实测,定量研究了无线信号时变特征和浅层室内场景下MDT中GNSS的漂移特性,总结了现网指纹定位技术需要改进的3个技术点,并通过试点验证了这一结论。受限于无线环境的复杂和当前MDT数据质量,低成本、高精度的指纹定位技术的实现仍需要进行更进一步的研究,下一步的工作重点包括引入高频Massive MIMO技术、创建非规则的地理化栅格^[3]、研发异常定位数据清洗等技术实现指纹定位模型的泛化性能和定位精度提升。

参考文献:

[1] MONDAL R, TURKKA J, RISTANIEMI T, et al. Performance evaluation of MDT assisted LTE RF fingerprint framework [C]//2014 Seventh International Conference on Mobile Computing and Ubiquitous

Networking (ICMU). Singapore: IEEE, 2014: 33-37.

[2] 晏志强, 李红艳, 王伟, 等. 立体化网络覆盖精准评估的应用与研究[J]. 邮电设计技术, 2018(8): 1-6.

[3] 赵俊德. 一种基于最小化路测(MDT)数据的天馈智能优化方法[J]. 电信工程技术与标准化, 2017, 30(12): 42-44.

[4] 汤凯, 李宗璋. 最小化路测应用研究[J]. 移动通信, 2017, 41(15): 1-6.

[5] 贺琳, 刘申建, 郭省力, 等. 最小化路测技术发展现状及应用分析[J]. 邮电设计技术, 2012(12): 13-17.

[6] 陈嘉明. 基于管理的Immediate MDT数据源的技术应用研究[J]. 电信工程技术与标准化, 2018, 31(11): 87-92.

[7] 3GPP. NR; radio resource control (RRC); protocol specification: 3GPP TS 38.331[S/OL]. [2025-02-30]. [ftp://ftp.3gpp.org/Specs/](http://ftp.3gpp.org/Specs/).

[8] Universal terrestrial radio access (UTRA) and evolved universal terrestrial radio access (E-UTRA); radio measurement collection for minimization of drive tests (MDT); Overall description; 3GPP TS 38.331[S/OL]. [2025-02-30]. [ftp://ftp.3gpp.org/Specs/](http://ftp.3gpp.org/Specs/).

[9] 吕非彼, 杨洁艳, 吴浩然, 等. 图像识别技术在5G网络覆盖率评估中的应用探讨[J]. 邮电设计技术, 2020(6): 40-44.

[10] LYU F, CHENG C, ZHU J J, et al. Coverage estimation of mobile network using supervised learning model on artificial estimation dataset [C]//IEEE. 2021 International Conference on Information and Communication Technologies for Disaster Management (ICT-DM). Hangzhou: IEEE, 2021: 214-220.

[11] SOHRABI F, KUEHN E. Construction of the RSRP map using sparse MDT measurements by regression clustering[C]//2017 IEEE International Conference on Communications (ICC). Paris: IEEE, 2017: 1-6.

[12] CHERNOGOROV F, NIHTILA T. QoS verification for minimization of drive tests in LTE networks[C]//2012 IEEE 75th Vehicular Technology Conference (VTC Spring). Yokohama: IEEE, 2012: 1-5.

[13] ZHANG P, CHEN H. A survey of positioning technology for 5G[J]. Journal of Beijing University of Posts and Telecommunications, 2018, 41(5): 1-12.

[14] VIEIRA J, LEITINGER E, SARAJLIC M, et al. Deep convolutional neural networks for massive MIMO fingerprint-based positioning [C]//2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). Montreal: IEEE, 2017: 1-6.

作者简介:

吕非彼,高级工程师,硕士,主要从事网络智能运营技术研究工作;许国平,高级工程师,博士,从事网络运营优化工作;李纪华,高级工程师,硕士,主要从事移动网络优化技术研究工作;朱佳佳,正高级工程师,硕士,主要从事网络智能运营技术研究工作;张琰,工程师,博士,主要从事网络AI技术研究工作;只璐,工程师,硕士,主要从事网络智能运营技术研究工作;王昭宁,工程师,博士,主要从事网络运营管理标准化研究工作。