

面向数字化运营的无线网问题根因定位算法研究

Research on Root Cause Localization Algorithm for Digital Operation Oriented Wireless Network Problems

史文祥¹,赵伟²,孟宁³,郭云霄¹,卢伟荣⁴(1. 中讯邮电咨询设计院有限公司,北京 100048;2. 中国联通浙江分公司,浙江杭州 310051;3. 中国联合网络通信集团有限公司,北京 100033;4. 中国联通山西省分公司,山西太原 030006)
Shi Wenxiang¹,Zhao Wei²,Meng Ning³,Guo Yunxiao¹,Lu Weirong⁴(1. China Information Technology Designing & Consulting Institute Co., Ltd., Beijing 100048, China; 2. China Unicom Zhejiang Branch, Hangzhou 310051, China; 3. China United Network Communications Group Co., Ltd., Beijing 100033, China; 4. China Unicom Shanxi Branch, Taiyuan 030006, China)

摘要:

在中国联通数字化转型过程中,为了快速定位无线网络问题,在识别出质差小区后,基于CM、PM、MR等数据,制定无线网络问题根因定位算法,建立根因定位算法库。根据根因定位算法实现对质差小区的问题自动定位和分析,从而快速、准确地定位网络问题,确保用户感知,降低人工成本,提升一线工作效率。

关键词:

无线网;网络优化;根因定位

doi:10.12045/j.issn.1007-3043.2022.01.013

文章编号:1007-3043(2022)01-0067-06

中图分类号:TN929.5

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

In the process of digital transformation of China Unicom, in order to quickly locate the wireless network problems, after identifying the poor quality cell, it formulates wireless network root cause location algorithm and establishes root cause location algorithm library based on CM, PM, MR and other data. According to the root cause location algorithm, the problem of poor quality cells can be automatically located and analyzed, so as to quickly and accurately locate network problems, which could ensure user perception, reduce labor cost and improve front-line work efficiency.

Keywords:

Wireless network; Network optimization; Root cause localization

引用格式:史文祥,赵伟,孟宁,等. 面向数字化运营的无线网问题根因定位算法研究[J]. 邮电设计技术,2022(1):67-72.

1 概述

随着近年来移动网络的大规模建设,4G移动网络覆盖已日趋完善,但网络的日常维护工作愈加繁重,网络优化问题日益凸显。传统方法主要通过一线工程师采集网络信息,依靠工程师分析各项数据,然后人工排查覆盖、干扰、容量等各类网络问题,通过他们的经验调整基站参数,从而对网络问题进行优化。这种优化方法效率低、准确度差,对优化人员的水平要求较高,但用户感知却可能并没有得到很大提高,难

以满足现在大规模网络优化的要求。在运营商数字化转型的背景下,中国联通以数字化运营为契机,基于大数据和人工智能技术的发展,采取多种措施提升用户感知。通过统一采集和解析全国基站和小区信息,建立数据分析库,以网络大数据为基础,实现对问题小区的识别和派单,把问题直达一线,主动进行维护和服务。但若依靠传统优化方法让一线工程师针对每个问题工单逐一进行分析,仍然无法避免效率低、准确率差的问题。在此背景下,本文旨在通过建立无线问题根因定位算法库,对问题小区进行根因定位,从而帮助一线人员快速定位网络问题,高效支撑生产运营。

收稿日期:2021-12-13

2 无线网问题根因定位算法分析思路

网络质量和用户感知之间相辅相成。评价用户感知,可以从“站的上、驻留稳、速率高”等方面来综合评价,当小区存在低接入、高掉线、低速率等问题时,则认为该小区为影响用户感知的质差小区。虽然能够通过掉线率、无线接入成功率、速率等指标识别出这些质差小区,但由于影响用户感知的原因多种多样,甚至可能是多种因素导致,一线人员定位质差小区产生的原因依然十分困难。

为了快速定位并解决无线网络问题,提升一线工作效率,助力中国联通数字化运营,结合省分公司实际,进行了根因算法的研究。在专家经验的基础上,将网络问题分为覆盖、干扰、容量、故障等类型,结合CM、PM、MR等数据,通过制定相应的根因算法,实现对质差小区的根因标注。

由于无线问题的复杂性,对无线网的问题根因定位注定是个长期、艰巨的研究课题,即便能够实现对问题的根因标注,但对于疑难问题,仍然需要网优专家实际环境勘察的辅助,才能制定最终的解决方案。根据研究的深入程度,本文将无线网根因定位的研究分为3个阶段。

a) 基于专家经验制定根因算法,实现对省分公司带根因派单,通过运行反馈不断完善根因算法,逐步提高根因标注率和准确率。

b) 引入AI算法对根因算法进行优化,通过对多种根因制定不同权重,实现规划、建设、优化问题的分拣,实现质差小区解决方案的输出。

c) 利用无线Trace数据补充完善根因算法,在之前根因算法的基础之上,提高根因定位的准确性,解决疑难问题。

受限于目前数据获取等实际情况,本文主要针对第1阶段基于专家经验的根因算法进行研究,并对根因定位研究的第2阶段和第3阶段进行了探讨。

3 基于专家知识图谱的根因定位算法

本章详细介绍了基于专家知识图谱的根因定位算法。基于优化专家在4G网络日常优化维护工作中的经验积累,结合网络中实际存在的问题,将问题分为覆盖、干扰、质差、容量等类型,从多个维度设置指标门限,定位质差小区的根因,在文献[1]中已经梳理了部分根因定位算法,如图1所示。

本文在前期研究的基础上,根据实际运行反馈,对根因定位算法进行了迭代优化。

3.1 覆盖类问题根因定位算法

无线网络覆盖类问题是网络覆盖范围不合理导致的,一般宏站场景下网络覆盖问题可以分为弱覆盖、过覆盖和重叠覆盖等。在本文所提出的根因算法中,通过RSRP、TA、重叠覆盖率等指标联合定位重叠覆盖、弱覆盖等覆盖类问题。

当前4G网络规模建设已趋于成熟,但城区、郊区/县城、乡镇/农村等不同场景站点规划的覆盖范围不同:城区站点密集,郊区/县城站点站间距较大,乡镇/农村站点规划多为广覆盖。因此,在制定根因算法时,需要根据不同覆盖场景设置不同的TA门限。

城区站点一般规划覆盖范围在1 km内,考虑城区内无线环境差异,以 $TA > 1$ km的比例 $> 25\%$ 作为判定城区过覆盖的先决条件;而郊区/县城站点由于站间距较大,考虑其覆盖区不宜过远,但同时也不能过近,以 $TA > 2.5$ km的比例 $> 25\%$ 作为判定郊区/县城过覆盖的先决条件;乡镇/农村站点虽然覆盖较广,但考虑到无线传播损耗问题,信号覆盖越远衰减越大、信号穿透能力越弱,覆盖距离不宜过广,以 $TA > 3.5$ km的比例 $> 25\%$ 作为判定乡镇/农村过覆盖问题的先决条件。

覆盖过近是指信号覆盖远低于规划设定的覆盖范围,例如质差小区的天线下倾角设置过大、覆盖方向有遮挡物、天馈系统存在隐性故障等导致覆盖过近,可以通过TA、RSRP等指标进行综合判断。

使用单一指标评判质差根因的准确率较低,可通过TA、MR等指标,结合不同场景的站间距、小区平均覆盖距离,制定弱覆盖、重叠覆盖、过覆盖、覆盖过近等根因定位算法,如表1所示。

针对宏站弱覆盖、过覆盖等覆盖类问题,在保证基站及天馈系统工作正常、参数设置合理的情况下,一般有5种解决方案:调整天线高度、调整天线下倾角和方位角、调整RS的功率、站点的重新选址搬移、增加RRU。其中,优先考虑调整天线下倾角和方位角的方案。

3.2 干扰类问题根因定位算法

无线网络干扰类问题一般包括系统外干扰和系统内干扰。在本文所提根因算法中,通过平均每PRB干扰噪声功率这一表征干扰强度的指标定位干扰类问题。

系统外干扰一般称之为外部干扰,是由外部干扰

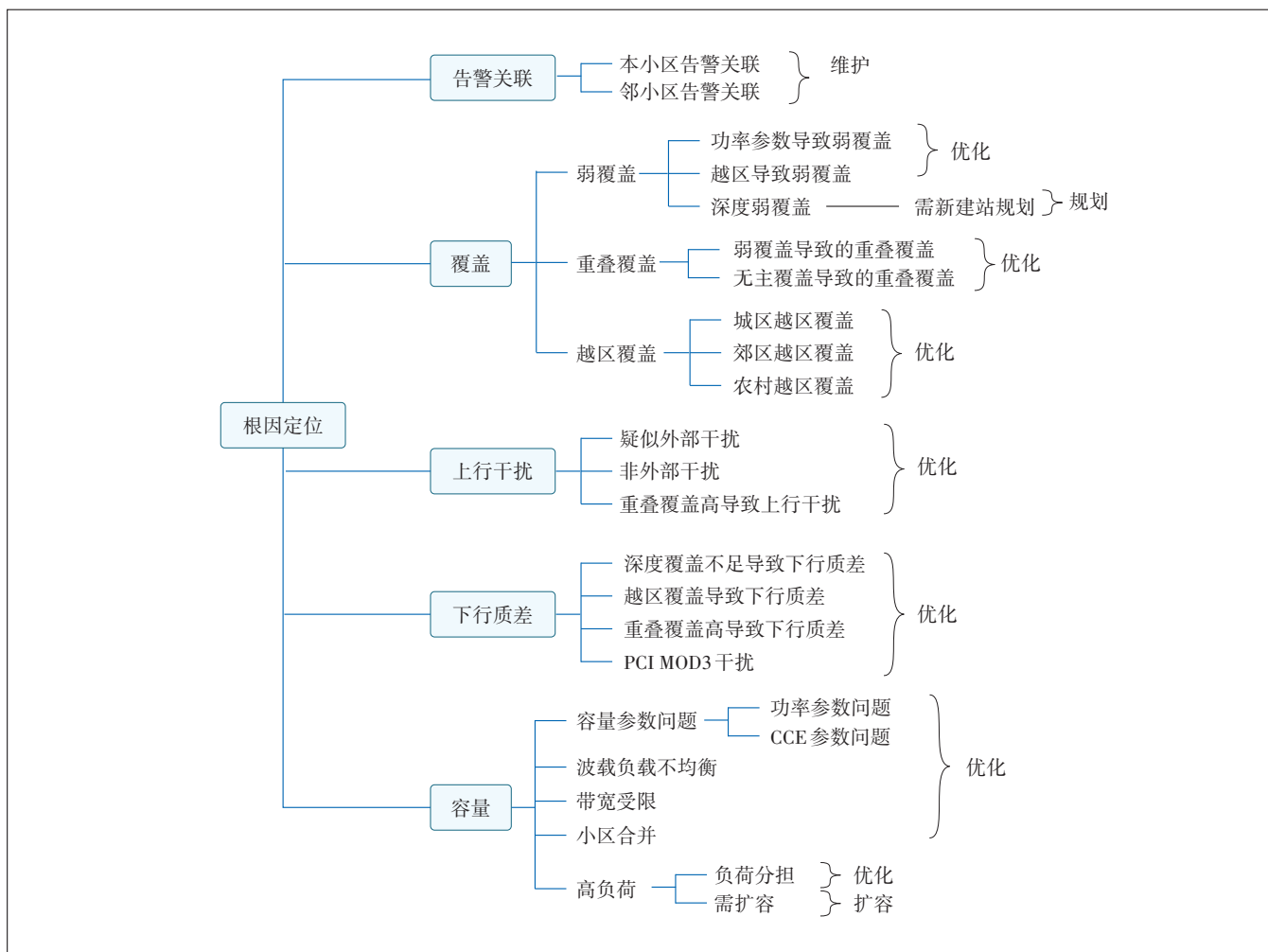


图1 根因定位算法梳理

表1 覆盖类问题根因定位算法

问题类型	质差根因	根因定位规则	适用场景
弱覆盖	过覆盖导致	(城区 TA>1 km 的比例>25% 或 郊区/县城 TA>2 km 的比例>25% 或 乡镇/农村 TA>3.5 km 的比例>25%) 且 RSRP<-112 dBm 采样点比例>20%	区分城区、郊区/县城、乡镇/农村
	深度覆盖不足	RSRP<-112 dBm 采样点比例>10% & 平均TA覆盖距离<覆盖方向最近宏站站间距	不区分场景
过覆盖	过覆盖	(城区 TA>1km 的比例>25% 或 郊区/县城 TA>2 km 的比例>25% 或 乡镇/农村 TA>3.5 km 的比例>25%) 且平均TA覆盖距离 > 覆盖方向最近宏站站间距	区分城区、郊区/县城、乡镇/农村
	越区覆盖	平均TA覆盖距离 > 覆盖方向最近宏站站间距	不区分场景
重叠覆盖	无强信号	RSRP<-112 dBm 采样点比例>20% 且重叠覆盖率>25%	不区分场景
	多个强信号	RSRP<-112 dBm 采样点比例<20% 且重叠覆盖率>25%	不区分场景
覆盖过近	存在弱覆盖	宏站平均TA覆盖距离<0.23 km 且 RSRP<-112 dBm 采样点比例>5%	不区分场景
	需一线核查	平均TA覆盖距离<0.078 km	不区分场景

源引起的。系统内干扰是指系统内同频干扰、GPS故障、数据配置错误等引起的干扰。外部干扰是连续、持久的,判断外部干扰时需剔除白天因用户使用等因素导致干扰抬升的情况,一般考虑使用03:00点业务闲时的干扰指标进行评判;系统内干扰产生的时间段

不固定,受用户行为等因素的影响较大,一般考虑使用全天的干扰指标。系统外干扰、系统内干扰的根因定位算法如表2所示。

针对干扰类问题,在保证基站及天馈系统工作正常、参数设置合理的情况下,一般有5种解决方案:排

表2 干扰类问题根因定位算法

问题类型	质差根因	根因定位规则	适用场景
干扰类	外部干扰	凌晨3点平均每PRB干扰噪声值>-105 dBm(剔除0值)	不区分场景
	系统内干扰	日平均每PRB干扰噪声值>-105 dBm(剔除0值)且凌晨3点平均每PRB干扰噪声值<-105 dBm(剔除0值)	不区分场景

查外部干扰源、调整天线高度、调整天线下倾角和方位角、调整RS的功率。其中,优先考虑排查外部干扰源方案。

3.3 质差类问题根因定位算法

无线网质差相关指标能直接表征网络质量,例如,CQI表征信道质量,MCS表征资源调度情况。一般宏站场景下网络质差问题主要由弱覆盖、过覆盖和重叠覆盖等原因引起。在本文所提根因算法中,通过CQI、MCS、TA等指标联合定位质差类问题。

质差类问题中城区、郊区/县城、乡镇/农村站点的覆盖判决条件与3.1节中一致,在此不再详述。在质差类根因定位算法中,将CQI、MCS结合TA指标、小区平均覆盖距离、不同场景下的平均站间距等指标联合来定位质差小区根因(见表3)。

解决质差类问题与解决覆盖类问题相似,在保证基站及天馈系统工作正常、参数设置合理的情况下,通过天线高度的调整、天线下倾角和方位角的调整、RS功率的调整进行优化。其中,优先考虑天线下倾角、RS功率的调整方案。

3.4 容量类问题根因定位算法

高负荷直接影响网络的各项指标,在本文所提根因算法中,通过PRB利用率、RS功率等表征负荷的指

表3 质差类问题根因定位算法

问题类型	质差根因	根因定位规则	适用场景
质差类	下行质差:过覆盖导致	(CQI<7占比大于20%或平均MCS<10或MCS<10的占比大于20%)且[(城区TA>1 km的占比大于25%或郊区/县城TA>2 km的占比大于25%或乡镇/农村TA>3.5 km的占比大于25%)或平均TA覆盖距离>覆盖方向最近宏站站间距]	区分城区、郊区/县城、乡镇/农村
	下行质差:重叠覆盖	(CQI<7的比例>20%或平均MCS<10或MCS<10的比例>20%)且重叠覆盖率>25%	不区分场景
	下行质差:弱覆盖	(CQI<7的比例>20%或平均MCS<10或MCS<10的比例>20%)且RSRP<-112 dBm采样点比例>20%	不区分场景
	下行质差	(CQI<7的比例>20%或平均MCS<10或MCS<10的比例>20%)且不满足过覆盖、弱覆盖和重叠覆盖的条件	不区分场景

标联合定位载波间话务负荷不均衡、扇区间话务负荷不均衡等负荷类问题。

在同一RRU下,2个不同频点相同带宽的逻辑小区的最大功率相同,此时如果其PRB利用率差值>20%则为负载不均衡;实际情况中同一物理小区下,2个相同覆盖的逻辑小区为规避干扰等各种因素,其最大功率不一定都相同,但正常情况下差值不会超过3 dB。结合实际情况,以PRB利用率和RS功率联合定位载波间负载均衡问题。同时,此方法也同样适用于扇区间的负载均衡判决。

通过PRB利用率、CCE利用率、用户数、流量、RS功率、小区合并等指标,制定了载波间话务负荷不均衡、扇区间话务负荷不均衡、用户数多、带宽不足、合并状态的逻辑小区负荷高、License受限扇区间话务负荷不均衡、CCE利用率高等根因算法,如表4所示。

表4 容量类问题根因定位算法

问题类型	质差根因	根因定位规则	适用场景
容量类	容量受限原因:载波间话务负荷不均衡	PRB利用率>70%且扇区载波间RS功率差值>3 dB且扇区载波间PRB利用率差值>20%	不区分场景
	容量受限原因:扇区间话务负荷不均衡	PRB利用率>70%且覆盖方向邻小区RS功率差值>3 dB且覆盖方向邻小区扇区间PRB利用率差值>20%	不区分场景
	容量受限原因:用户数多	RRC连接平均用户数多,5 MHz带宽:用户数>40;10 MHz带宽:用户数>75;15 MHz带宽:用户数>120;20 MHz带宽:用户数>150	不区分场景
	容量受限原因:带宽不足	PRB利用率>70%且小区带宽<20 MHz且小区频点非900 MHz	不区分场景
	容量受限原因:合并状态的逻辑小区负荷高	PRB利用率>70%且小区合并	不区分场景
	容量受限原因:License受限扇区间话务负荷不均衡	由于RRCLicense受限导致的RRC连接拥塞率>0,或由于RRCLicense受限导致RRC连接失败的次数>0,或License超限使用告警或资源分配失败而导致RRC连接建立失败的次数>0	不区分场景
	容量受限原因:需扩容小区	上行PRB利用率或下行PRB利用率大于80%,且忙时流量大于8 GB,且平均用户数大于50	不区分场景
	容量受限原因:CCE利用率高	上下行PRB利用率均小于60%,且CCE利用率大于50%	不区分场景
	容量受限原因:其他原因	PRB利用率>70%	不区分场景

针对宏站高负荷问题,在保证基站及天馈系统工作正常、参数设置合理的情况下,一般有5种解决方案:小区扩容、拆分合并小区、载波间负载均衡、站间负载均衡、扩容 License。其中,优先采用载波间负载均衡和站间负载均衡的方案。

3.5 维护类问题

通过操作维护中心(Operation and Maintenance Center, OMC)获取基站影响业务类的告警和故障信息,结合质差小区劣化指标、劣化时间,联合定位根因。

3.6 算法应用与效果验证

对某市开展了根因定位算法的应用实验,以验证本文所提算法的有效性,针对该市586个质差小区进行指标对比,有485个小区完全符合根因定位算法,标注率达82.8%。根据该市反馈,根因标注与小区实际问题定位一致的比例为69.3%(见表5)。

表5 根因定位算法验证结果

质差小区分类	根因标注率/%	根因标注准确率/%
覆盖类	57.1	68.4
干扰类	6.5	100.0
质差类	10.1	69.2
容量类	9.1	100.0
总计	82.8	69.3

4 根因研究探索

4.1 基于AI算法拓展根因算法

在基于专家知识图谱的根因定位算法中,算法门限是基于专家经验确定,可能会存在门限设置不太合理,根因定位不够精准和深入等情况,影响质差小区的解决效率。通过引入AI算法,实现对质差小区规划、建设、优化问题的分拣,优化根因的派单规则。同时,根据系统运行过程中的反馈,验证目前的根因算法是否合理,通过AI算法实现对根因定位算法的迭代优化,并逐步实现质差小区解决方案的自动输出。

以XGboost算法为例,首先输入获取的CM、PM、MR等数据,由于原始数据存在一些无效值和缺失值等问题,因此需要对数据进行清洗。在特征选择和环节,尽量挑选出对根因定位产生影响的所有特征值,并对特征值进行特征处理,构建出所需数据集。将数据集进行分割,训练集用于XGboost算法训练,并输出模型。之后将测试集导入训练模型中进行测试,

输出质差小区问题根因,算法流程如图2所示。

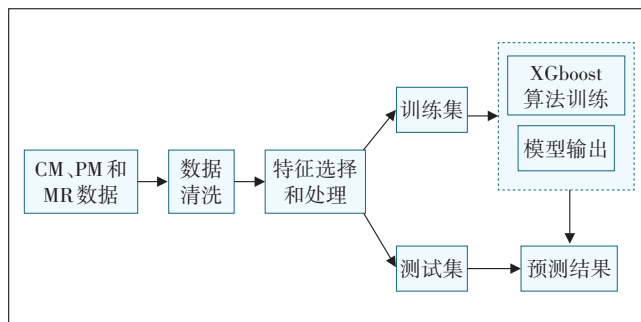


图2 基于XGboost算法拓展根因算法流程

4.2 基于无线Trace数据拓展根因算法

在无线环境复杂、难以使用常规方法定位问题原因的质差区域,可以采用无线Trace信令与MR、CM、PM等数据相关联,形成基于栅格的问题区域,利用仿真技术和场景参数训练,支撑问题区域的优化,并输出解决方案。

Trace信令即用户跟踪信令。依照《中国联通TDD-FDD-LTE数字蜂窝移动通信网无线信令及呼叫记录采集技术要求v2.0》规范,一个无线详细信令记录接口文件中采用二进制文件格式,依照指定规则涵盖特定时间段内某基站指定接口的信令数据。接口文件结构由文件头及Trace数据包2个部分构成。文件头包括Trace数据的采集起止时间以及相应规范版本信息等,Trace数据包则由数据包长度、通用包头、专用包头和原始信令数据4个部分构成。通过识别每个数据包的起始字符,可根据规范内的数据格式移位完成数据包的解析并生成对应的解析文本。目前4G默认订阅终端和网络的空中接口UU及UUextend、基站之间的X2接口、基站与核心网之间的S1共3个接口的数据,如图3所示。

通过信令数据、话单数据和MR数据中共有的MME_UE_S1AP ID字段,实现对用户终端信息及覆盖信息的回填,将基站级的信令拆解到基站下的单个用户,呈现单用户网络及业务接入、保持、切换、释放的全流程(见图4)。

4.3 需要注意的问题

数据准确性问题。中国联通采用省分公司上报、集团公司解析的方式进行数据处理,但集团公司采集到的数据与省分公司采集的部分指标可能存在差异,影响下一步的派单和问题分析。

邻区数据问题。邻区数据一般由省内自行获取

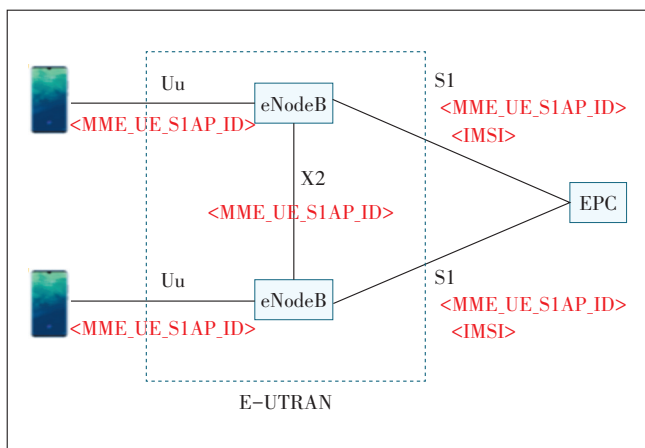


图3 4G信令接口数据

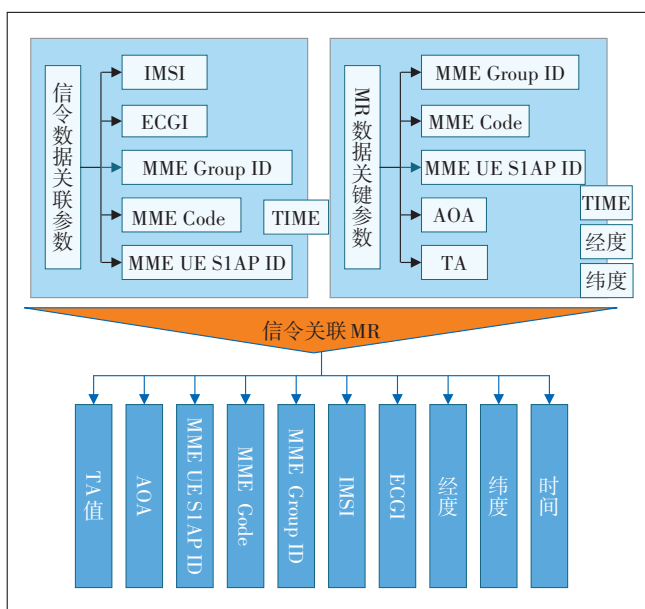


图4 信令数据关联MR

和分析,需关联邻区数据核查质差小区邻区的覆盖、告警、切换参数等情况,判断是否由周边邻区导致本小区质差。

维护数据问题。故障、告警等维护类问题对指标劣化存在直接影响,目前故障、告警信息与CM、PM、MR等小区无线指标分别属于不同系统,后续需将故障、告警信息与无线指标进行关联分析。

5 结束语

通过CM、PM、MR等数据能够有效分析4G网络中存在的问题,再结合AI算法和Trace数据实现对复杂无线问题的定位,进而有针对性地制定行之有效的解决方案。文中在前期研究的基础上,制定了基于专家

知识图谱的根因定位算法,实际算法应用结果表明,所提的根因定位算法对质差小区的标注率较高,可用于指导实际工作。后期随着能够采集到的数据逐渐丰富,可以针对更多数据源对根因定位算法进行拓展,提升无线网问题定位的准确性,提高网络维护和优化效率。

参考文献:

- [1] 蔡子龙,王晓刚,史文祥,等. 移动用户感知预判与数字化运营探讨[J]. 邮电设计技术,2021(6):35-40.
- [2] 杨磊. 基于FPgrowth机器学习的影响用户感知无线根因问题的快速定位方法研究[J]. 江苏通信,2019,35(2):56-62.
- [3] 杨飞虎,许国平,刘贤松,等. 基于Prophet时序算法的无线网络突变小区识别方法研究[J]. 邮电设计技术,2021(2):25-30.
- [4] 朱佳佳,马昱,杨洁艳,等. 基于聚类算法的5G网络覆盖问题智能定位[J]. 邮电设计技术,2020(8):6-10.
- [5] 徐青,宋琪,鲁志强. 机器学习算法在业务卡顿根因分析中的应用[J]. 数字通信世界,2019(1):214.
- [6] 王希. 确定基站小区质差根因的方法和系统[J]. 电子世界,2018(16):91-93.
- [7] 肖旭旭,门里,冯喆. 基于时间提前量的异常覆盖定位及优化方法浅析[J]. 邮电设计技术,2016(10):38-42.
- [8] 彭东升,黄华,王琪. 基于AI技术的5G基站告警收敛与根因诊断[J]. 网络安全和信息化,2021(9):106-109.
- [9] 付思冉,潘亮,崔岩. 基于机器学习的4G手机上网质差小区定位研究与实现[J]. 山东通信技术,2020,40(2):8-11.
- [10] 何蕊馨. 一种基于多数据源的楼宇弱覆盖小区评估分析方法[J]. 电信工程技术与标准化,2019,32(8):64-68.
- [11] 王希. 基于概率神经网络(PNN)的LTE质差小区分析方法[J]. 数字通信世界,2017(2):89-90,80.
- [12] 王勇,滕祖伟,周杰华,等. AI深度学习在移动网异常小区检测分类中的应用[J]. 邮电设计技术,2019(11):11-15.
- [13] 郭正,郭宁,黄蕴思. 基于多维度数据挖掘的自学习故障根因定位系统[J]. 电子技术与软件工程,2021(15):146-149.
- [14] 唐琳钧,冯喆,肖旭旭. 基于大数据的移动通信网稳定性差小区定位方法[J]. 邮电设计技术,2016(5):37-41.
- [15] 李翔明,杨晓康,孙磊. 基于大数据的无线网络智能分析优化方法[J]. 信息系统工程,2021(1):140-141,144.

作者简介:

史文祥,毕业于重庆邮电大学,高级工程师,硕士,主要从事移动网络优化及数字化运营工作;孟宁,毕业于北京邮电大学,工程师,硕士,主要从事重点场景专项优化及数字化运营工作;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的运用;郭云霄,毕业于重庆邮电大学,助理工程师,硕士,主要从事4G/5G无线网络优化及数字化运营工作;卢伟荣,毕业于山西大学,工程师,硕士,主要从事4G/5G无线网络优化及数字化运营工作。