

基于人工智能的运营商故障分析 Research on Improving Fault Analysis Capability of Operators Based on Artificial Intelligence 能力提升研究

朱宏, 邓程, 王瑜, 宋文杰 (中国联通智网创新中心, 江苏南京 210019)
Zhu Hong, Deng Cheng, Wang Yu, Song Wenjie (Intelligent Network & Innovation Center of China Unicom, Nanjing 210019, China)

摘要:

传统的故障分析手段将运维经验固化为故障分析规则或脚本,这种方式针对特定故障模式较为有效,但是无法应对新网络业务、组网变化,一旦规则、脚本需要调整,需投入较大成本进行适配改造,且时效性较差。而人工智能技术在大数据统计、分析、推理、自适应学习上有着先天优势,能快速基于新网络、新业务的变化重训练AI模型参数,给出最佳推荐值。基于此,重点研究了如何利用人工智能提升运营商故障分析能力。

关键词:

故障分析;智能分析;人工智能;大数据;网络运维
doi:10.12045/j.issn.1007-3043.2024.06.015
文章编号:1007-3043(2024)06-0072-06
中图分类号:TN919
文献标识码:A
开放科学(资源服务)标识码(OSID):



Abstract:

Traditional fault analysis methods are fixed into fault analysis rules or scripts based on O&M experience. This method is effective for specific fault modes, but cannot cope with new network service and networking changes. Once these rules and scripts need to be adjusted, a large cost is required for adaptation and reconstruction, and the timeliness is poor. Artificial intelligence technology has inherent advantages in big data statistics, analysis, inference, and adaptive learning. It can quickly retrain AI model parameters based on changes in new networks and businesses, and provides the best recommendation value. Based on this, it focus on the research of using artificial intelligence to improve the fault analysis capability of operators.

Keywords:

Fault analysis; Intelligent analysis; Artificial intelligence; Big data; Network O&M

引用格式:朱宏,邓程,王瑜,等.基于人工智能的运营商故障分析能力提升研究[J].邮电设计技术,2024(6):72-77.

1 概述

在日常网络维护中,故障分析是最基础、最重要的工作,故障分析结果的时效性、准确性直接影响了故障处理的效率、网络质量及运维成本。如何从海量告警中准确识别出有效故障,做到告警不多、不漏、不错,一直是网络故障监控的痛点。常见的故障分析手段是基于规则的(例如单网元告警合并,或者基于某一区域统计告警数量),但这往往不能满足实际运维诉求,想要实现比较复杂的分析逻辑就要在规则基础上叠加脚本,这类脚本随着网络、业务以及运维要求

的变化会越来越复杂,持续开发和维护成本也会越来越高。那么是否能利用AI算法在大数据处理、自适应学习等方面的优势^[1],提升运营商故障分析能力是本文的研究重点。

2 传统故障分析手段的弊端

传统故障分析方法一般为基于白名单机制的主次关联、衍生关联。针对基站退服、射频单元掉电,一般会配置2条主次关联规则(见表1)。

若单纯采用主次关联规则,就会同时生成2个故障,而在实际场景中,故障原因可能是由射频单元掉电引起的基站退服。若不在派单环节对故障加以关联或合并,就会导致派单增多,带来额外的运维成本。

收稿日期:2024-04-22

表1 主次关联规则示例

规则ID	规则名称	告警厂家	告警领域	告警名称	告警主次类别
PC_001	基站退服	厂家A	无线	eNodeB退服告警	P
		厂家A	无线	小区不可用告警	C
PC_002	射频单元掉电	厂家A	无线	射频单元直流掉电告警	P
		厂家A	无线	射频单元交维维护链路异常告警	C
		厂家A	无线	BBU光模块收发异常告警	C

传统故障分析还存在以下弊端。

a) 当有新的故障模式出现时,已有规则不能及时调整,导致部分故障无法被及时识别出来,影响网络质量。

b) 规则的调整和适配依赖专家经验,而专家经验需要覆盖各厂家设备、各专业领域,才能完成规则的调整和适配,规则调测、上线周期长,成本高。

c) 多规则运行可能存在冲突,影响最终的故障分析结果。

3 整体设计思路

为了解决传统故障分析手段的弊端,本方案引入人工智能技术。人工智能的核心是机器可以自学习,那么在故障分析领域,机器应该学习什么呢^[2-3]? 当海量告警来临的时候,故障分析的本质是要完成告警的聚类、故障点及根因的有效识别。这就需要将告警的聚类、故障点及根因识别的内在逻辑抽象出来,形成一个特征,机器以这些特征为输入、以特定算法模

型为基础,输出一组组故障模式,并在专家校准后,反向优化算法模型。

机器完成学习(训练)之后,需要在实战中进行检验(推理),其整体设计思路如图1所示。

a) 训练态。按照时间和空间2个维度将历史告警(3~6个月)及网络拓扑数据进行切片(分组),构建频繁性挖掘的AI算法模型(基于贝叶斯理论和离散度分析)。该模型支持专家经验注入(将CT网络的故障传导性作为专家经验注入),融合生成故障传播图。故障传播图表征了项目故障发生的传播路径,当推理态有新的告警发生时,可依据传播图快速确定故障根因。同时,故障传播图具备故障传播关系重复、冲突、成环检测能力。对于一般网络,故障传播图可一周训练一次以进行迭代更新,而对于故障模式变化较快的网络,可一天训练一次从而快速迭代更新。

b) 推理态。在日常监控场景中,通过流式处理的方式实时地对活动告警进行时空聚类,并调用故障传播图对聚类告警进行切割和根因网元推断,生成最终的故障组,快速识别故障影响范围、故障节点及根因。

训练态和推理态都依赖告警和拓扑数据的质量,告警的主要设计约束如下。

a) 告警数据标准化。需保证告警标准字段信息的完整、准确,包括告警标题、告警流水号、告警对象设备类型、网元名称、领域名称、站点名称、设备厂家名称、告警对象名称(或告警源)、机房名称、网络类型、专业网管名称、设备类型、省、地(市)、县级市名称

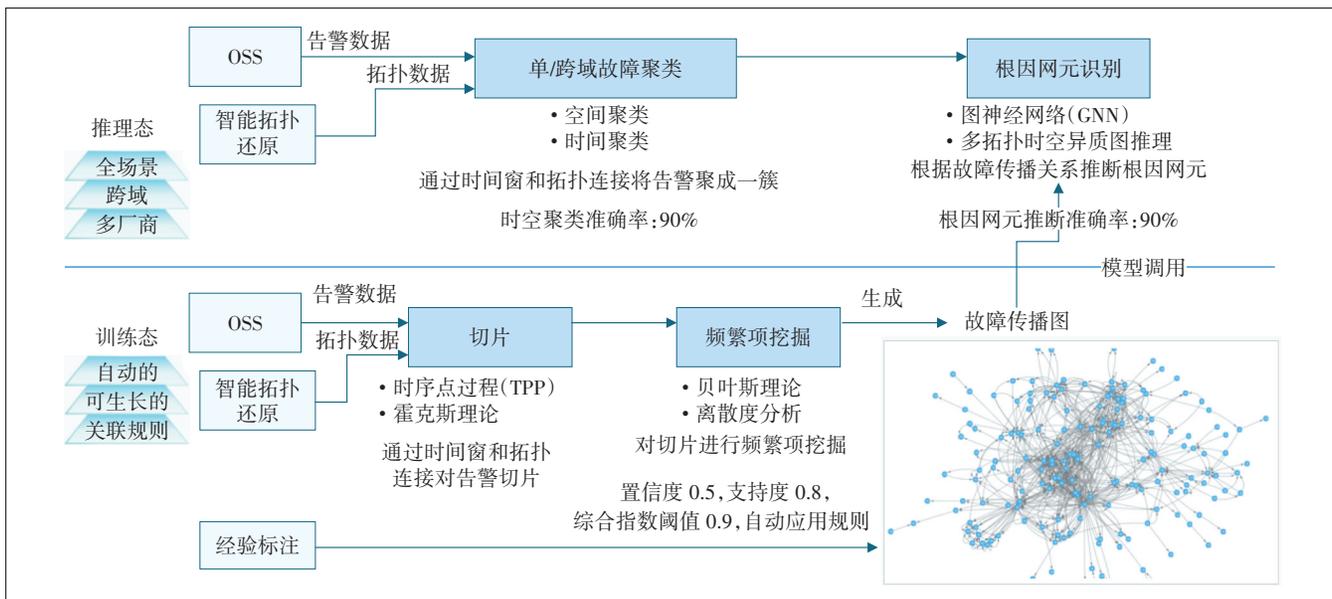


图1 智能故障分析方案架构

等。

b) 告警时间一致性。

(a) 不同网管和本系统之间的时间设置与实际时间同步。

(b) 网元与网管之间的时间同步。

(c) 本系统时间设置与实际时间的同步。

(d) 告警从网管北向上报到本系统活动告警的入库时延小于 30 s。

c) 拓扑还原完整度大于 95%(拓扑还原出来的链路/网络中实际链路数)。

d) 网元机房关系完整度大于 95%(机房信息准确的网元数/所有网元数)。

4 核心算法逻辑

如上所述,人工智能技术的关键是找到“好的特征进行训练”。对于CT领域来说,不同网元上的告警是有内在联系的,只有找到这种内在的联系,才能找到好的“特征”。

a) 时间连续性。同一个故障衍生的多个告警,其告警发生的时间接近,可能存在关联性。

b) 空间相关性。网元/设备、机房、相邻/相近的拓扑对象发生的告警,可能存在关联性。

c) 网络传导性。基于 3GPP 协议、TCP/IP 协议,协议栈间的告警存在关联性。结合 3GPP 协议及现网运维经验,提炼网元内及网元间的故障传播关系。3GPP 协议的影响是自下而上的,例如,如果物理层出现故障,一般链路层、网络层、业务层均会受到不同程度的影响,若物理层没有问题,链路层出现问题,则会导致网络层和业务层受影响,反之则不成立。不同的网络专业之间也会互相依赖、互相影响,例如无线基站设备掉电可能导致传输设备端口故障(收不到光),而传输设备掉电可能导致多个无线基站退服。

上述 3 个特征是本方案的核心理论依据,尤其是第 3 个特征,是将 CT 经验与 AI 算法模型结合的关键,也是人工智能在 CT 领域智能故障分析中取得较好应用的重要基础(业界通用的智能告警关联技术主要基于前 2 个特征)。但是,人工智能技术可能会自适应地学到大量无效的、甚至是错误的规则,这给后期专家标注带来了巨大的工作量。在 AI 算法模型中提前注入 CT 设备的“通信原理”,可以大大提升人工智能输出结果的有效性。

基于 3 种关联模式,在网络拓扑上挂载各个网元

的故障告警,使用图搜索算法完成点(告警)、边(告警关系)、边方向(告警影响方向)3 个维度的信息渲染,再叠加故障传播关系,形成故障传播图。

在拓扑上,若某个节点有告警发生,则称该节点被激活(被点亮)。在一定时间范围内,由被点亮节点所形成的最大连通区域中的所有告警形成的集合被称为子拓扑(subtopo)。子拓扑表示的是一定时空范围内的告警所形成的集合,且在相应的时间范围内,这些告警所形成的空间范围(连通区域)应最大化。

在挖掘过程中,子拓扑的生长方式为自底向上的。当多个子拓扑因为某个节点的点亮被连通起来时,这些子拓扑就需要进行合并,以始终保持连通区域的最大化。所以,故障就是子拓扑在一定时间范围内不断地在空间上进行扩展所形成的。一个故障在空间上的边界,就是相应子拓扑在一定时间范围内所能延展的最大范围。

子拓扑会随着时间的增长而不断地扩展(或者保持不变),那它什么时候结束呢? 一个子拓扑在时间上的定界由 2 个变量控制:一个是单步等待时长,另一个是整体等待时长。

单步等待时长,即一个子拓扑上相邻 2 条告警到达时间的间隔。整体等待时长,即一个子拓扑从第 1 条告警到达开始算起的整体时长。这 2 个变量存在 2 个阈值:一个是单步超期,当子拓扑的单步等待时长大于单步超期时,认为子拓扑结束;另一个是整体超期,当子拓扑的整体等待时长大于整体超期时,认为子拓扑结束。

图 2 给出了子拓扑单步等待时长示意,其中 $A_1, A_2, A_3, \dots, A_7$ 是先后到达同一个子拓扑的告警。对于子拓扑 1 来说,因为 A_2 和 A_3 之间的单步等待时长大于单步超期,所以子拓扑 1 在 A_2 后结束;对于子拓扑 2 来说,虽然 A_4 和 A_5 之间并未单步超期,但 A_3 和 A_5 之间的整体等待时长大于整体超期,所以子拓扑 2 在 A_4 后结束。

整体超期往往由业务因素决定,如设置为 SLA (Service Level Agreement)。单步超期有 2 个确定方

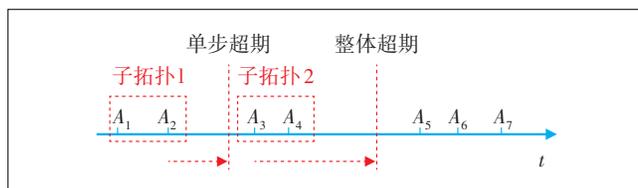


图2 子拓扑单步等待时长示意

法,一是根据经验直接指定,如所有故障统一指定为3 min;二是通过机器学习对单步等待时长进行建模,实时动态预测各个子拓扑的单步等待时长,以此来判定一个故障是否单步超期。

根据专家经验设置的故障单步等待时长为固定值,这种方法的缺点在于无法兼顾各种情景。单步等待时长会因为故障场景的不同而不同,有的故障在3 min内便会把告警上报完毕,有的故障则会持续十几分钟,甚至持续数个小时。人工配置的值只能尽可能保证对多数故障是合理的,但因为对各种故障都是“一刀切”,所以容易顾此失彼。

对历史数据进行分析,通过机器学习的方法来确订单步超期,可以解决根据经验设置单步超期的缺点。机器学习方法众多,本文采用按泊松过程对告警的到达事件进行建模。该方法的输入为一定量的历史告警,输出为各类告警的等待时长预测模型。

以某类告警 A_i 为例,根据历史数据,分析 A_i 发生后,后续告警的发生情况,建立 A_i 告警等待时长预测模型(见图3)。

对 A_i 中的某一条告警 A_{ij} ,当 A_{ij} 发生后,在一定时间间隔 $T_{tolerance}$ 内,在融合拓扑其所属网元的 n 阶邻域中,统计其后续告警发生的集合 S_{ij} 。

遍历所有的 j ,得到告警 A_i 的泊松过程参数——强度(intensity) λ , $\lambda_{ij} = \frac{|S_{ij}|}{T_{tolerance}}$, $\lambda_i = \sum |\lambda_{ij}|$ 。

设 A_i 对应的等待时长为 w_i ,在时间区间 $[N(t), N(t+w_i)]$ 中,告警发生次数服从泊松分布:

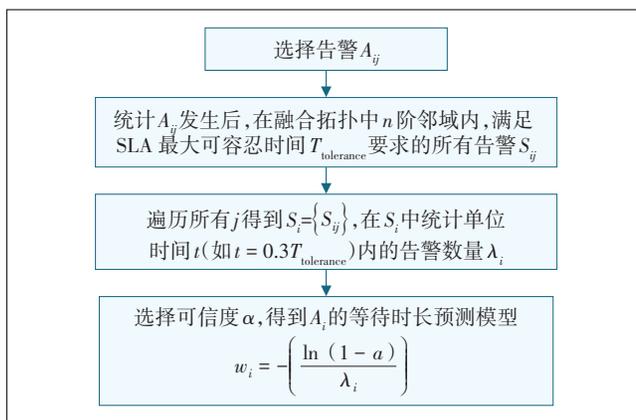


图3 告警等待时长预测过程

$$P\{[N(t+w_i) - N(t)] = k\} = \frac{e^{-\lambda_i w_i} (\lambda_i w_i)^k}{k!}, k = 0, 1, 2, \dots (1)$$

假设在 w_i 时间内有告警发生的概率为 α ,则无告警发生的概率为 $1 - \alpha$ 。所以, $1 - \alpha = e^{-\lambda_i w_i}$ 。

$$\text{进一步, } w_i = -\frac{\ln(1-\alpha)}{\lambda_i}$$

告警按照时空维度进行聚类,叠加故障传播图,使用图搜索技术识别故障的根因告警和根因网元。

时空流式聚类如图4所示。经上述算法,可快速将告警聚类形成一个告警组。针对时空聚类后的告警组,调用故障传播图,使用图搜索技术识别故障的根因告警和根因网元(见图5)。

用数学语言描述寻找根因网元的过程,具体如下。

- a) 建立超拓扑 HG ,将其转化为无向图 G 。
- b) 寻找 G 的所元件(component) $\{\hat{C}_i\}$,并按 HG 将

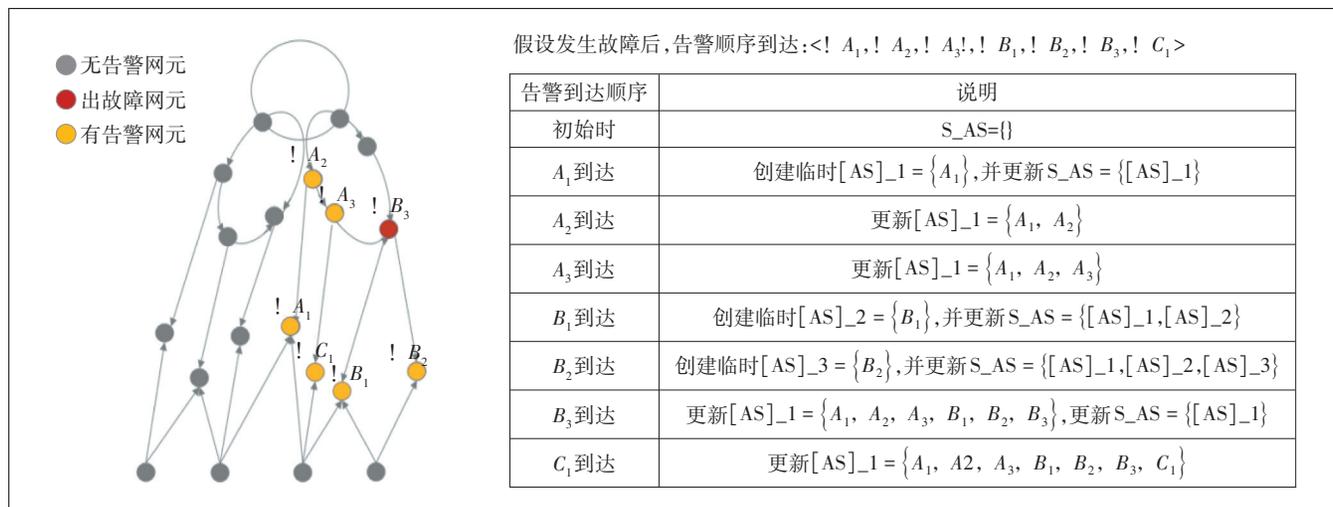


图4 时空流式聚类示意

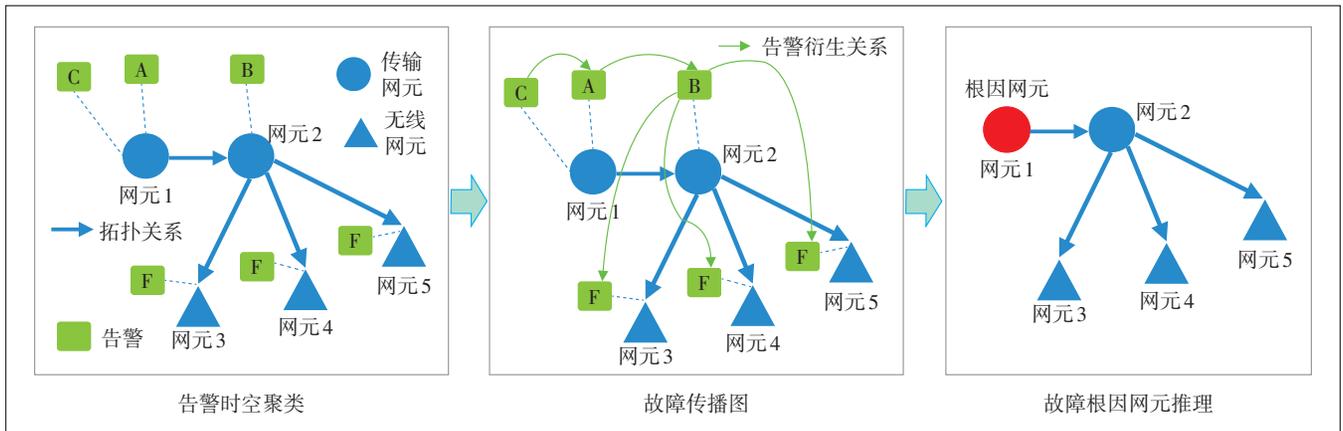


图5 故障传播示意

其恢复成有向图 $\{C_i\}$ 。

- c) 定义跨域边集合 $CE = \{ce_j\}$ 。
- d) 如果 $E(C_i) \cap CE \neq \emptyset$, 则 $C_i \in C^c$ 。
- e) 如果 $E(C_i) \cap CE = \emptyset$, 则 $C_i \in C^s$ 。
- f) 对于所有的 $C_i \in C^c$, 寻找 $V(C_i, d_{in} = 0)$ 。
- g) 如果 $V(C_i, d_{in} = 0) \neq \emptyset$, 则为 $V(C_i, d_{in} = 0)$ 的 C_i 的根因节点。
- h) 如果 $V(C_i, d_{in} = 0) = \emptyset$, 则说明一定存在双向边, 按照 i) 处理。

i) 在 $V(C_i, d_{in} = 0) = \emptyset$ 时, 取 C_i 的所有双向边连接的子图 C_i^{double} , 并将其转化为无向图 \widehat{C}_i^{double} , $V(\widehat{C}_i^{double}, d = 1)$ 即为根因节点。

g) 如果 $C_i \in C^s$, 则 $V(C_i) = \emptyset$, 即无根因节点。

某智能故障分析实例如图6所示, 其中左图为局部网络拓扑图, 右图为告警传播关系图。从接收到告警开始, 一般建议在 10 min 以内做出处理; 而采用传统方式对这样一组故障进行分析时, 一般需要 30 min 以上。

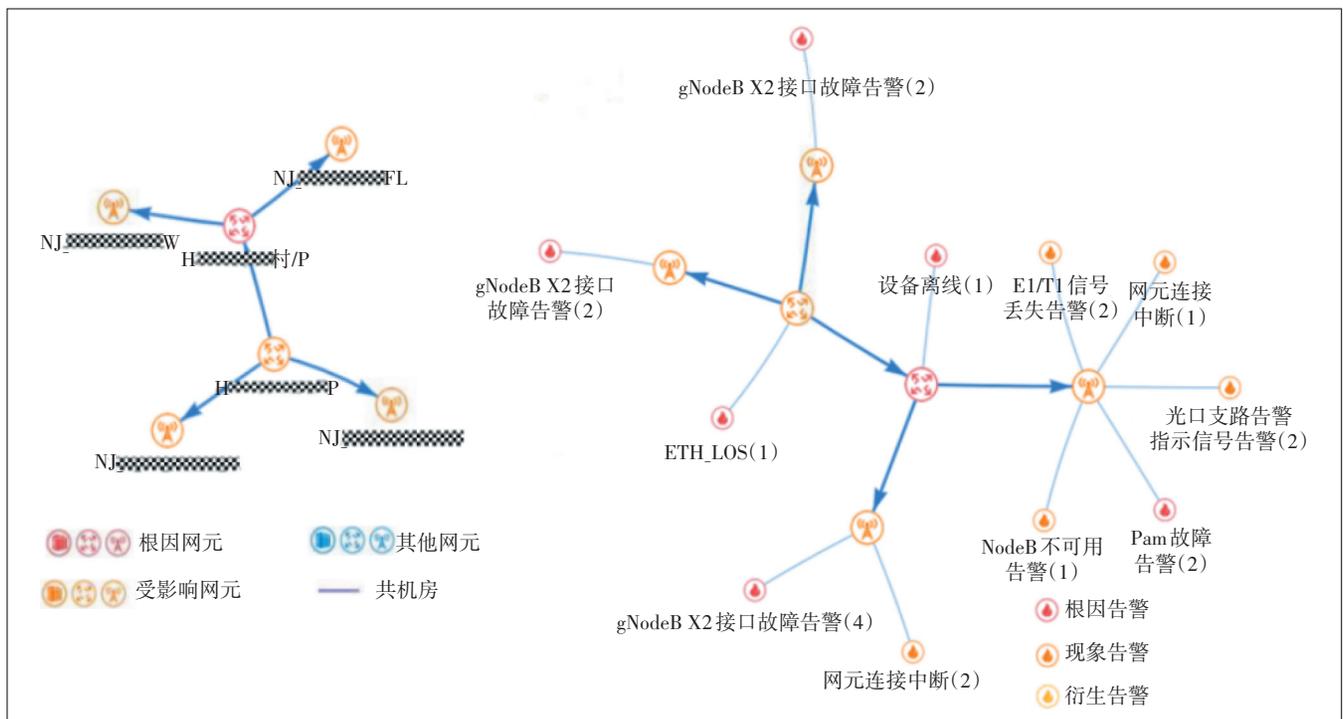


图6 智能故障分析实例

5 智能故障分析应用效果

如前所述,故障分析的本质是完成告警的聚类、故障点及根因的有效识别。为了评估智能故障分析的应用效果,制定2个指标:故障聚类覆盖率和根因识别准确率。

a) 故障聚类覆盖率。该参数主要是为了评估所聚类的结果是否能直接支撑派发工单,计算如式(2)所示。其中,故障组即智能分析聚类后,为这一组内的告警指定的唯一的ID,用于标识这一组告警是有传导关系的。

$$\text{聚类覆盖率} = \frac{\text{智能分析有效故障组个数}}{\text{日故障量(有效工单)}} \times 100\% \quad (2)$$

b) 根因准确率。根因准确率中错误根因的故障组需要运维工程师反馈,在实际操作中可通过人工抽检并结合专家经验进行验证。

$$\text{根因准确率} = \frac{\text{智能分析根因准确故障组个数}}{\text{智能分析有效故障组个数}} \times 100\% \quad (3)$$

经验证,N市无线+IPRAN接入环网络,涉及7800个以上的4G网元以及多厂家组网,该方案故障聚类覆盖率为90%,根因网元识别准确率为93%,压降了7%的工单数。某日具体告警数据统计结果以及6930个有效故障组的故障场景分布分别如表2和表3所示。

深入分析发现,大部分根因不准确都与基础数据

表2 告警统计实例

日告警总量	无线告警数量	传输告警数量	动环告警数量	日故障量(有效工单)	智能分析有效故障组个数	聚类覆盖率/%	智能分析根因准确故障组个数	根因准确率/%
778 000	466 800	194 500	116 700	7 681	6 930	90.2	6 504	93.9

表3 故障分类实例

故障组名称	发生次数	根因准确个数
DU小区退服	162	150
LTE小区退出服务	638	648
NR小区不可用	454	451
NodeB不可用	48	48
NodeB退服	186	186
UMTS小区不可用	38	36
gNodeB退服	6	6
厂家B 3G同一逻辑站多小区退服	862	812
厂家B 4G同一逻辑站多小区退服	1 076	998
厂家B 5G同一逻辑站多小区退服	1 382	1 228
厂家B 5G系统时钟不可用	2	2
厂家A 3G同一逻辑站多小区退服	50	50
厂家A 4G同一逻辑站多小区退服	268	204
厂家A 5G同一逻辑站多小区退服	469	451
厂家A电调天线ALD故障	2	2
多基站退出服务	12	12
小区不可用告警	349	324
小区退服	842	815
网元连接中断	74	71
网元链路断	10	1
总计	6 930	6 504

的缺失或不准确有关,例如机房信息不准确导致同机房的网元根因分析不准确(约占30%)、网元间拓扑连接关系的缺失(约占20%)、告警延迟上报(约占10%)等,这都会导致根因分析有误。也存在少数新组网场

景下,算法模型不适配的情况,这种情况需要结合专家经验持续进行模型调优。

6 结束语

通过对传统故障关联技术的分析,发现其很难应对新网络业务的变化。本文创新性地将人工智能技术引入了故障关联分析领域,同时叠加了CT领域的专业知识作为第3维关键特征,极大地提升了智能分析结果的准确性。此外,还存在因基础数据缺失,算法模型不能适应特定组网而导致的分析错误,后续一方面可完善基础数据治理的方法,另一方面可持续更新算法模型以适应更多的组网模式。

参考文献:

- [1] 伏玉笋,杨根科. 人工智能在移动通信中的应用:挑战与实践[J]. 通信学报,2020,41(9):190-201.
- [2] 余萍,曹洁. 深度学习在故障诊断与预测中的应用[J]. 计算机工程与应用,2020,56(3):1-18.
- [3] 中国人工智能产业发展联盟. 电信行业人工智能应用白皮书[R/OL]. [2024-01-16]. <http://221.179.172.81/images/20210322/68111616377652538.pdf>.

作者简介:

朱宏,毕业于宁波大学,工程师,硕士,主要从事智能运维产品的研发工作;邓程,毕业于南京师范大学,工程师,硕士,主要从事IDC产品的研发工作;王瑜,毕业于南京邮电大学,智网创新中心总监,高级工程师,主要从事智能运维产品的研发工作;宋文杰,毕业于南京师范大学,工程师,硕士,主要从事故障中心产品的研发管理工作。