



基于网络资源树及信息熵的 智慧家庭业务感知劣化定位方案

明萌, 时宇, 郑润含, 冯云喜
(中国电信集团有限公司, 北京 100020)

摘要: 随着科学技术和信息化发展的步伐日渐加快, 智慧家庭市场发展迅速, 给人民生活带来了极大的便捷, 也给运营商带来新的运维考验。提出了一种基于大数据和 AI 算法、结合网络资源树及信息熵对智慧家庭业务进行感知实时监测, 并对感知劣化类故障进行定界、定段、定位的方案, 旨在解决智慧家庭业务端到端环节较多、群障及单用户故障人工定位困难的问题。

关键词: 智慧家庭; 感知监测; 告警定位

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022082

Smart home service perception degradation positioning scheme based on the network resource tree and information entropy

MING Meng, SHI Yu, ZHENG Runhan, FENG Yunxi
China Telecom Group Co., Ltd., Beijing 100020, China

Abstract: With the accelerating pace of development of the science and technology and informatization, the smart home market is developing rapidly. It brings great convenience to people's lives, also brings new operation and maintenance tests to operators. A scheme based on big data and AI algorithm, combined with the network resource tree and information entropy was proposed for perceptual real-time monitoring of smart home services, delimiting, segmenting and locating perceptual degradation class faults, aiming to solve the problems of more end-to-end links of smart home services and difficulties in manual location of group faults and single user faults.

Key words: smart home, sensory monitoring, alarm location

0 引言

自 2013 年以来, 凭借“互联网+”的风潮, 国内外各大运营商纷纷借助宽带运营与家庭互联网产业链, 积极发展智慧家庭业务, 在推进人民

家庭生活的智慧化、信息化方面做着尝试与转变。传统的宽带业务是智慧家庭业务的基础, IPTV 以光宽带为接入方式, 向用户提供家庭娱乐/信息化一体化服务、视频通话/智能家居/智能音箱等新兴业务。天翼看家通过“光宽+4G/5G”网络接入,



依托云化开放的智慧家庭平台,为客户提供解决家庭视频安防需求的“云+网+平台+应用+服务”型产品。天翼看家与传统宽带+IPTV业务共同构建了更为丰富多彩的智慧家庭生态圈。

当前,对IPTV故障定位的研究方法较多,包括EPG(electronic program guide)响应故障定位^[1]、视频质量故障定位^[2]、网络系统故障定位^[3]等,对家庭宽带和其他智慧家庭产品的故障定位方法较少。文献[3]设计实现了一种大数据环境下的IPTV故障定位系统,采用相关性算法与聚类算法结合的方式筛选用户关键绩效指标(key performance index, KPI)中的代表性指标,建立质差用户筛选模型得到质差用户。文献[4]主要采用视频平均意见得分(video mean opinion score, VMOS)评价用户观看视频业务体验。文献[5]通过遍历获取的故障树,对各非根节点对应的故障原因进行检测,并从中定位导致所述待定位故障现象的故障原因。文献[6]介绍了基于主动探测技术家庭宽带运维质量监测手段。以上文献大多基于网元KPI进行故障定位,缺少对感知群障的定界、定段、定位。

本文提出了基于网络资源树及信息熵的智慧家庭业务感知监测及感知劣化定位方案,定义了统一的智慧家庭业务感知指标体系,提供故障精准定界、定段、定位方案,旨在解决智慧家庭业务端到端环节较多、群障及单用户故障人工定位困难的问题。

1 需求分析

随着智慧家庭业务的快速发展,接入网^[8]、承载网和骨干网等端到端网络的规模也在不断扩大,故障运维难度也逐渐增大。当用户感知发生劣化时,如何精准快速地定位故障原因已成为运维方面的重大挑战。目前,产业界对家庭宽带、IPTV、天翼看家等智慧家庭业务的感知质量评估侧重于网络质量,主要是在用户主动投诉或维护人员定期或不定期地对设备进行测试方面,需要耗费一定的人力和物力。而对用户端到端感知的

监测和感知劣化定位手段相对缺乏,主要存在以下3点困难。

- 缺少统一的智慧家庭业务感知指标体系,数据存在厂商、型号、设备的壁垒。传统的用户感知评估模式通过调查问卷等方式判断用户感知,但是随着智慧家庭业务的快速发展,传统用户服务已无法满足用户感知提升的需求。为了科学、量化地评估用户感知并精准定位质差原因,需要建立统一的智慧家庭业务感知指标体系。终端存在多厂商、多型号的数据“壁垒”,需要定义统一的体验质量(quality of experience, QoE)^[9]、关键质量指标(key quality indicator, KQI)^[10]和KPI^[11],包括IPTV视频感知、家庭宽带感知和天翼看家使用感知的QoE模型,基于业务和应用层的KQI和网元KPI,建立统一的端到端精准定位方法。
- 缺乏主动预防群障手段,无法预先发现群障隐患。传统网络的维护方法只能在故障发生后才开始维护,不能满足智慧家庭群障预警、预防的需求。为了更好地提升用户使用智慧家庭业务的感知,先于用户发现故障并进行预见性维护可以减少用户投诉。
- 群障定段/定界、根因分析、单用户判障困难。传统的质量监测方法中网络性能指标与用户感知不对应,维护人员在处理用户感知劣化等故障时,较难判断真实的故障点,这导致群障和单用户等故障维护困难。

本文通过定义统一的智慧家庭业务感知指标体系,提供预见性维护和故障的精准定界、定段、定位方案,解决群障和单用户判障困难等问题。

2 智慧家庭业务感知质量监测及劣化定位架构

本文融合云、网、端到端资源数据、终端探针侧数据、媒资信息及各类告警性能等,进行实

时采集、汇聚及关联挖掘分析。在用户终端部署质量监测探针,覆盖多渠道采集来源,打破传统质量监测探针无法适配不同厂商终端的限制;定义统一的视频感知优良率、宽带网页访问优良率等感知算法,结合性能数据进行关联分析,解决了传统的质量监测方法中网络性能指标与用户感知不对应、用户体验难以准确定量和定位等问题。针对感知数据量大的特点,本文提出解耦化的分模块处理架构,主要依托 Spark 大数据计算平台基于多项大数据技术组合,采用“Flume^[12]+Kafka^[13]+SparkStreaming^[14-16]+Redis^[17]+MongoDB^[18]+ES^[19]+MySQL^[20]”的技术框架,通过设计逻辑统一、物理分散的“分省前置模块+中心处理模块”的数据采集处理架构,适用于智慧家庭业务感知群障告警实时分析、非实时隐患优化、单用户问题定位等多种生产维护需求。

2.1 感知数据采集与指标体系定义

本文采取在用户终端部署质量监测探针的方式,监测不同播放行为下的实际用户业务体验,经现网测试应用,该方式能灵敏、准确地感知用户的实际体验。针对宽带业务,通过家庭网络质量插件主动获取用户家庭网络设备信息、无线信息、异常信息等关键指标,通过宽带质量主动拨测探针插件模拟用户行为,获取浏览网页、观看视频、玩游戏时业务感知体验优良率。针对 IPTV 业务,通过部署机顶盒软探针,从芯片解码层面感知用户卡顿,获取用户的播放行为数据。针对视频监控类业务,在手机 App 采用埋点方式上报用户行为数据。

基于探针获取的感知数据,本文定义了一整套智慧家庭业务感知指标体系。针对各类业务特点,定义了光宽、IPTV、视频监控业务的主要感知指标,智慧家庭业务感知指标体系如图 1 所示。其中,网页访问、视频、游戏均属于传统光宽业务。针对每一类业务,选取关键的感知指标,根据重要程度赋予相应权重,可以达到客观、统一评价各类业务感知的目的。

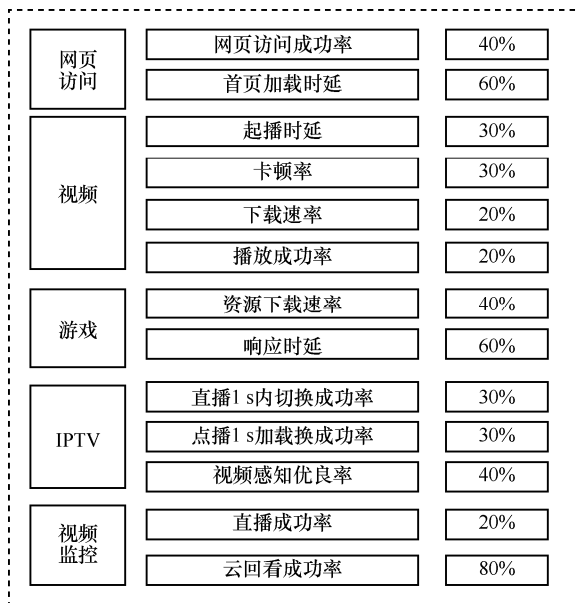


图1 智慧家庭业务感知指标体系

2.2 性能数据采集

性能数据主要通过设备网管进行各省分布式采集,本方案根据不同智慧家庭业务特点,主要选取设备端口利用率、CRC 误码率、CPU 利用率等 KPI,构建感知劣化问题定位的 KPI 模型。光宽网页访问业务 KPI 模型示例如图 2 所示,当发生光宽网页访问业务感知劣化问题时,可以通过 KPI 模型实现端到端定位能力。



图2 光宽网页访问业务 KPI 模型示例

2.3 系统架构

本方案采用 Spark 分布式计算框架作为核心引



擎,实现各种复杂的业务逻辑,对大规模数据进行快速的处理。在实践中通过对 Spark 任务进行多种的性能调优,包括代码优化、算子调优、参数调优、shuffle 调优以及数据倾斜调优等,实现所有 Spark 实时作业的执行时间都在 5 min 以内,且系统持续稳定运行,运行无时延。以单质量模块为例,实时计算每分钟每个用户的质量信息,每分钟用户话单合计约 1 500 万条(15 GB 左右)。算子优化及参数调整后,Spark 在 2~3 min 完成计算,比优化前效率提升了 5 倍,有效支撑了业务数据的实时处理。

智慧家庭业务感知质量监测及劣化定位架构如图 3 所示,共解耦为 5 个模块:数据采集、数据转发、数据清洗及分析、数据存储和页面显示/数据共享。其中,数据清洗及分析模块是系统的核心模块。首先,对采集的探针数据、性能数据、端到端资源数据进行数据清洗、关联、抽取,得到单个用户的基本信息、行为、感知指标数据,该数据一方面输出到 Elasticsearch 分布式集群支撑单用户查询需求;一方面输出至多维分析模块,按不同时间维度、不同区域、不同内容分发网络(content delivery network, CDN)节点等各维度进行收敛汇聚,生成多维结果数据,支撑实时监测及报表分析需求。告警模块对不同维度计算、分析、预测产生的数据进行监控,生成标准的告警数据,并对数据记录进行判断,将未达到优良率

阈值的记录判定为告警,在分析得出告警定位结果后,通过 Web 页面进行展示,并发送至综合告警系统,实现告警工单的及时派发。

3 感知劣化定位

本文针对网络告警关联复杂度高的问题,提出“信息熵”属性,将各层级设备下挂设备质差率分布均衡程度映射成信息熵,构建端到端特征库,对各层级信息增益及同级平行节点信息熵进行关联运算,实现感知劣化问题的精确定段。基于定段结果,结合网络性能数据和用户感知数据,通过关联分析实现接入网设备告警的根因定位。

3.1 端到端收敛

智慧家庭业务端到端环节较复杂,传统光宽业务,主要涉及家庭网络、接入网、城域网、骨干网、互联网数据中心(Internet data center, IDC)/服务提供商(service provider, SP)等;IPTV 业务涉及电子节目指南(electrical program guide, EPG)/CDN、节目源等;视频监控类业务涉及视频监控能力平台及服务节点。关联第 1.1 节的感知数据与用户资源拓扑数据、节目资源数据等,并定义感知优良率指标为:感知优良率=感知优良用户数/总播放(或使用业务)用户数,其中感知优良用户具体定义可参考第 1.1 节的感知指标。智慧家庭业务端到端拓扑结构如图 4 所示,包含从家庭

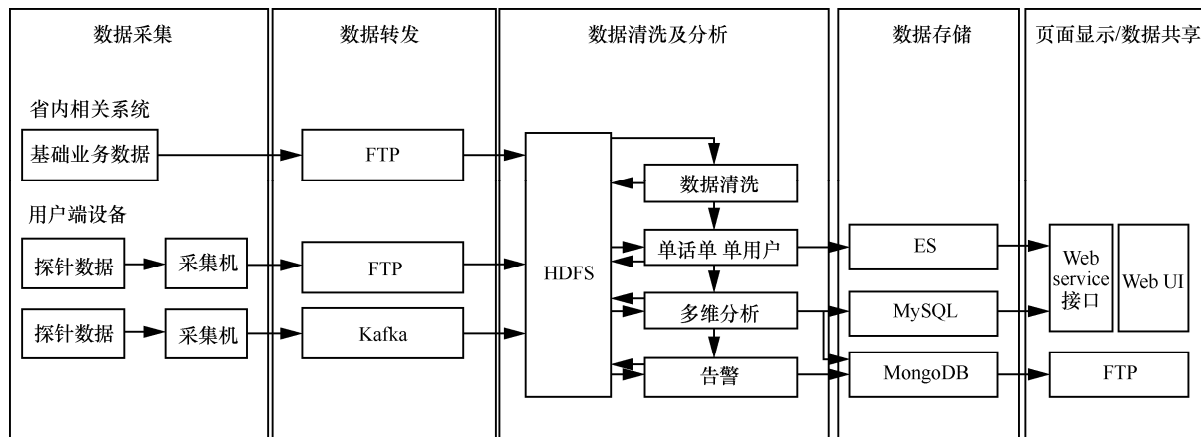


图3 智慧家庭业务感知质量监测及劣化定位架构

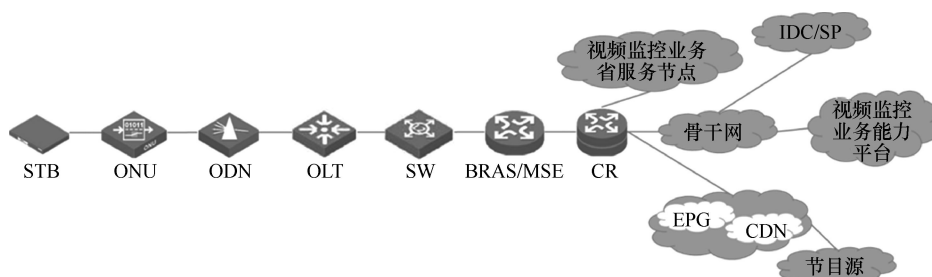


图4 智慧家庭业务端到端拓扑结构

网络至骨干网的各个段落设备 IP 地址、设备性能等数据，读取网络设备拓扑结构信息和感知优良率指标构造出网络资源树。

对各环节进行逐段收敛汇聚，即可得出各环节感知优良率。各环节感知优良率互相关联，感知劣化通常同时发生在多个环节。例如，当某节目源质量劣化时，观看该节目的用户感知劣化，用户所在的光线路终端（optical line terminal, OLT）、宽带接入服务器（broadband remote access server, BRAS）等设备感知指标均发生劣化。因此，当单用户或者某环节感知劣化超限时，无法直接定位故障原因。针对此问题，本文进一步提出了第 3.2 节中的定段算法，当单用户出现故障时，支持一键诊断，根据感知劣化情况收敛至某一环节，以实现感知劣化问题的精准定段。当群障发生时，根据第 3.2 节中的定段算法，定位故障原因。

3.2 基于信息熵的感知劣化定段

感知劣化定段即对每个设备节点，按照感知优良率指标计算信息熵，并在上下层的节点进行推理定界并定段。读取网络设备拓扑结构信息和感知优良率指标，构造网络设备树，即节点为各网络设备，父节点为同一链路上游的高层级网络设备，子节点为同一链路下游的低层级网络设备。同时每个节点包含设备 ID、感知优良率、层次编号、下挂子节点等属性。使用网络设备树结构储存数据的优点是在海量设备中查找指定节点速度更快，并为推理模块中对存在告警的节点链路进

行递归推理做铺垫。

遍历每个网络设备节点，初始化“信息熵”属性，并先对其所有子节点的感知优良率进行分箱编码操作，再计算信息熵，更新“信息熵”属性，信息熵计算式为：

$$H(X) = \sum_i P(x_i) I(x_i) = - \sum_i P(x_i) \log_b P(x_i) \quad (1)$$

如果某网络设备节点的信息熵很小，说明子节点的感知优良率集中落入某一分箱编码，否则子节点的感知优良率可能分布较为分散。实际应用中，若某一设备下挂的众多下联口质差均衡，则该设备的信息熵偏小，就会将群障告警定段归为该设备；反之，若该设备下挂的下联口设备出现质差不均衡，则信息熵偏大，就会将质差设备定段到下联质差设备，其中信息熵阈值分别为 0.65、0.4。

例如，查看 OLT 上联 BRAS 的信息熵情况如下。

- 当 BRAS 的信息熵存在 < 0.4 时，认为该故障点发生在 BRAS 设备及其以上网络层。
- 当 BRAS 的信息熵都 > 0.65 时，认为该故障点发生在 BRAS 下联的 OLT 或 OLT 连接 BRAS 的链路。
- 其他条件无法判断，不做归并定段处理。

基于信息熵的感知劣化定段示意图如图 5 所示，设网络设备树中的最高设备层级编号为 0，最低为 3，向上推断阈值为 a ，向下推断阈值为 b ，对每个网络设备节点按照以下逻辑进行递归推理。

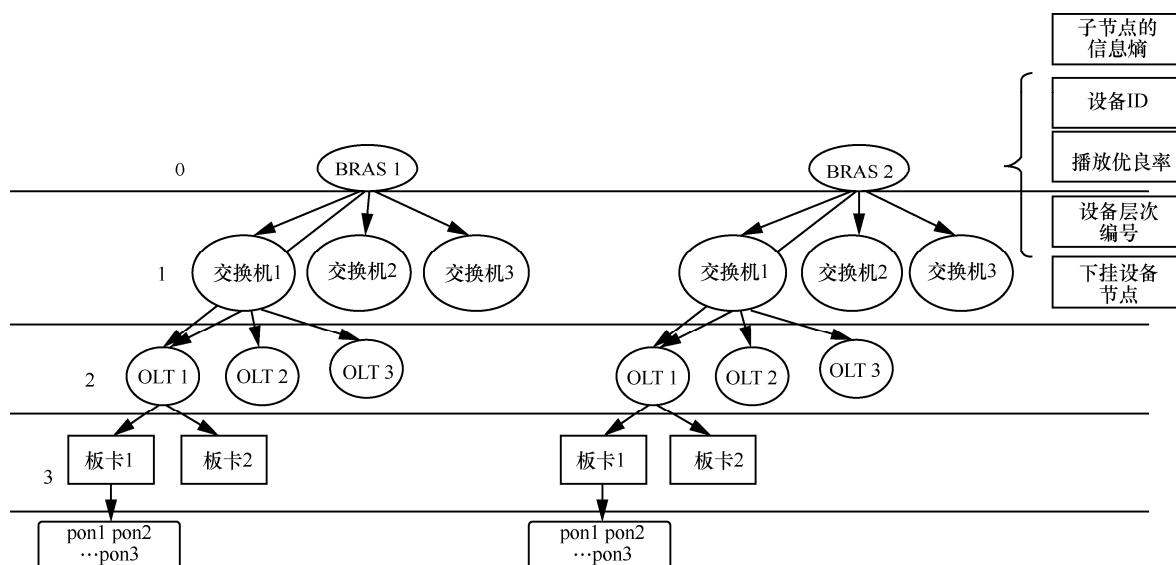


图5 基于信息熵的感知劣化定段示意图

如果该节点层级编号为0,且存在告警,若该节点的信息熵 $<a$,则推断故障发生在本节点;若该节点的信息熵 $>b$,则推断故障发生在下游节点。如果该节点存在告警,且不是最高层级,如节点的信息熵 $<a$,则查看父节点的推断结果,若父节点推断为下游,则判定为本节点故障,若父节点推断为上游或是父节点本身,判断故障位置参见父节点的定段结果;如该节点的信息熵 $>b$,推断故障发生在下游节点。

对告警进行定段处理,如果某设备故障告警定段于上游设备,且上游设备存在告警,同时上游设备定段在本节点或者上游,则删除本条告警。如果某网络设备告警故障定段于下游,则删除下游的所有告警。

本文实现了告警定段到具体的网络设备,通过在整条链路上进行递归推断,能够具体确定网络层级和编号。对于所有存在网络层级拓扑结构的设备故障定段场景,只需要确定信息熵计算依赖的用户体验指标以及上下推断的信息熵阈值即可复用。

3.3 结合性能数据的精确定位

为进一步实现感知劣化定位,需要基于网络

设备性能类、终端性能类数据等进行大数据建模,通过网络设备性能数据与设备感知优良率的关联分析,找到造成设备质差的性能阈值,从而实现感知告警的根因定位。

例如,若某感知告警定段至OLT,基于设备性能与用户感知关联对OLT故障进行根因定位。首先,对单个OLT历史时间内端口流量利用率与设备下的感知优良率(即卡顿且传输流(transport stream, TS)丢包用户占比)趋势做关联分析,发现OLT上联端口流量利用率(端口流量利用率为端口所用流量占端口总带宽的比例)与TS丢包率过高且卡顿用户占比过高有强相关性,OLT TS丢包卡顿率与端口流入流量利用率如图6所示。为找出端口流入流量利用率的阈值,进一步针对5 min 400万条OLT质差清单统计不同TS丢包卡顿率阈值下故障设备的占比情况。TS丢包卡顿率 $>5\%$ 占比随端口流量利用率分布变化如图7所示,当端口利用率超过70%,故障设备的占比开始显著上升。同理可分析得到OLT感知优良率与端口利用率、CRC误码率和CPU利用率等的关联关系。

当感知劣化定段至某具体段落,根据分析所

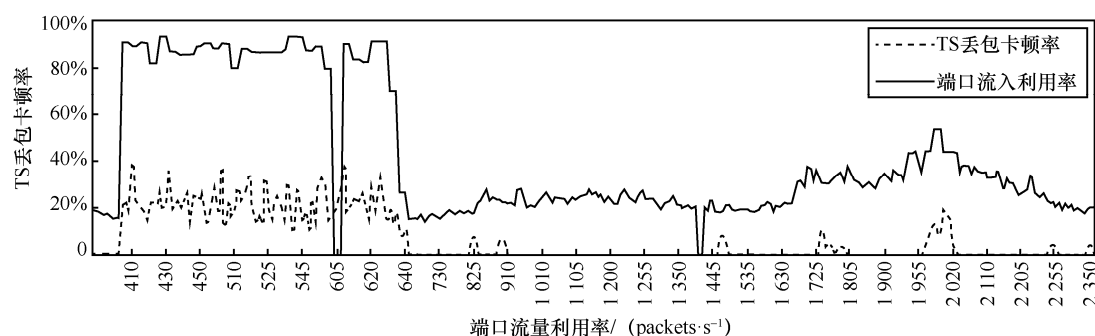


图6 OLT TS 丢包卡顿率与端口流入流量利用率

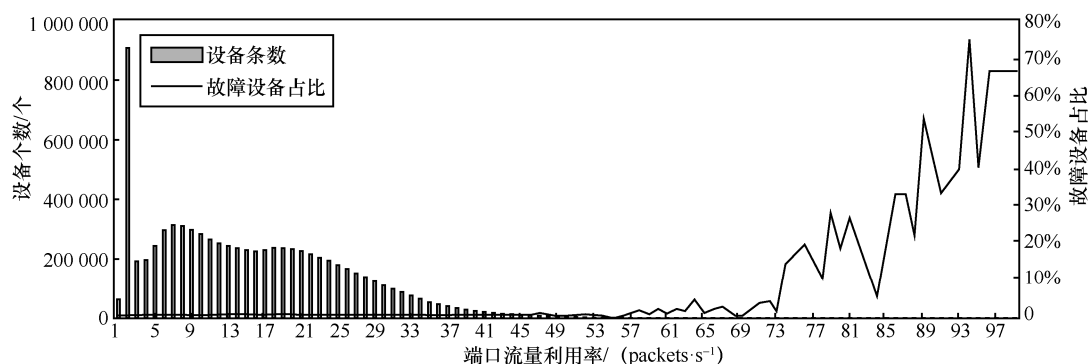


图7 TS 丢包卡顿>5%占比随端口流量利用率分布变化

得的关键性能指标与感知劣化的关联关系,可定位到某一个或几个具体性能指标。

4 结束语

本文提出了一种基于网络资源树及信息熵的智慧家庭业务感知监测及感知劣化定位方案,实现了智慧家庭业务感知劣化类故障的定界、定段、定位,解决了智慧家庭业务端到端环节较多、群障及单用户故障人工定位困难的问题。同时,该方案也适用于所有存在网络层级拓扑结构的感知劣化定位场景,在确定基于信息熵计算的用户体验指标以及信息熵阈值情况下,即可复用本文提出的定位方案,具有良好的可扩展性。

参考文献:

- [1] 谭晓敏, 方艾, 梁冰, 等. 基于决策树的 EPG 体验异常原因定位[J]. 电信科学, 2019, 35(8): 158-164.
TAN X M, FANG A, LIANG B, et al. Locating causes of abnormality of EPG experience based on decision tree[J]. Tele-

- communications Science, 2019, 35(8): 158-164.
- [2] 黄晓波, 黄志武. IPTV 视频质量故障定位及处理方法[J]. 信息通信, 2010, 23(4): 67-68, 60.
HUANG X B, HUANG Z W. IPTV video quality fault location and processing method[J]. Information and Communications, 2010, 23(4): 67-68, 60.
- [3] 刘奇峰. 基于大数据的 IPTV 故障定位系统的设计与实现[D]. 南京: 南京邮电大学, 2017.
LIU Q F. Design and implementation of IPTV fault location system based on big data[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2017.
- [4] 陈捷. 基于用户感知的 IPTV 端质量评估体系研究及实现[J]. 电信技术, 2019(9): 62-65.
CHEN J. Research and implementation of IPTV end-to-end quality evaluation system based on user perception[J]. Telecommunications Technology, 2019(9): 62-65.
- [5] 罗方方, 陈斌, 徐淑波. IPTV 系统故障定位方法及系统: CN107786897A[P]. 2018.
LUO F F, CHEN B, XU S B. Fault location method and system of IPTV system: CN107786897A[P]. 2018.
- [6] KARAGIANNIS T, PAPAGIANNAKI K, FALOUTSOS M. Blinc[J]. ACM SIGCOMM Computer Communication Review, 2005, 35(4): 229-240.
- [7] 张啸媛. 主动探测技术在家庭宽带用户业务感知质量监测中的应用[J]. 移动信息, 2017.



- ZHANG X Y. Application of active detection technology in service perceived quality monitoring of home broadband users[J]. Mobile Information, 2017.
- [8] 张中荃. 接入网技术[M]. 北京: 人民邮电出版社, 2017.
- ZHANG Z Q. Access network technology[M]. Beijing: Posts and Telecom Press, 2017.
- [9] 程卫华, 何肖嵘. 基于大数据的网页浏览质差分析方法研究[J]. 电信科学, 2020, 36(11): 174-181.
- CHENG W H, HE X R. Quality analysis method of Web browsing based on big data[J]. Telecommunications Science, 2020, 36(11): 174-181.
- [10] 陈森, 陈超, 张小勇, 等. 基于大数据分析的移动互联网用户感知评估系统[J]. 电信科学, 2015, 31(4): 154-161.
- CHEN S, CHEN C, ZHANG X Y, et al. Evaluation system of mobile Internet user experience based on big data analysis[J]. Telecommunications Science, 2015, 31(4): 154-161.
- [11] 王速, 卢华, 汪硕, 等. 智能运维中 KPI 异常检测的研究进展[J]. 电信科学, 2021, 37(5): 42-51.
- WANG S, LU H, WANG S, et al. Research progress of KPI anomaly detection in intelligent operation and maintenance[J]. Telecommunications Science, 2021, 37(5): 42-51.
- [12] 史蒂夫·霍夫曼, 斯里纳特·佩雷拉. Flume 日志收集与 MapReduce 模式[M]. 张龙, 译. 北京: 机械工业出版社, 2015.
- STEVE H, SRINATH P. Flume log collection and MapReduce mode[M]. Translated by ZHANG L. Beijing: Machinery Industry Press, 2015.
- [13] 妮哈·纳克海德, 格温·沙皮拉, 托德·帕利诺. Kafka 权威指南[M]. 薛命灯, 译. 北京: 人民邮电出版社, 2017.
- NEHA N, GWEN S, TODD P. Kafka authoritative guide[M]. Translated by XUE M D. Beijing: Posts and Telecom Press, 2017.
- [14] 比尔·钱伯斯, 马泰·扎哈里亚. Spark 权威指南[M]. 张岩峰, 王方京, 陈晶晶, 译. 北京: 中国电力出版社, 2020.
- BILL C, MATEI Z. Spark authoritative guide[M]. Translated by ZHANG Y F, WANG F J, CHEN J J. Beijing: Chinese Electric Power Press, 2020.
- [15] 汤姆·怀特. Hadoop 权威指南: 大数据的存储与分析[M]. 王海, 华东, 刘喻, 等译. 北京: 清华大学出版社, 2017.
- TOM W. Hadoop: the definitive guide storage and analysis at internet scale[M]. Translated by WANG H, HUA D, LIU Y, et al. Beijing: Tsinghua University Press, 2017.
- [16] NABI Z. Pro Spark Streaming[M]. Berkeley, CA: Apress, 2016.
- [17] 梁国斌. Redis 核心原理与实践[M]. 北京: 电子工业出版社, 2021.
- LIANG G B. Redis core principle and Practice[M]. Beijing: Electronic Industry Press, 2021.
- [18] 霍多罗夫. MongoDB 权威指南[M]. 邓强, 王明辉, 译. 北京: 人民邮电出版社, 2014.
- KRISTINA C. MongoDB: the definitive guide, second edition[M]. Translated by DENG Q, WANG M H. Beijing: Posts and Telecom Press, 2014.
- [19] 拉杜·乔戈, 马修·李·欣曼, 罗伊·罗素. Elasticsearch 实战[M]. 黄申, 译. 北京: 人民邮电出版社, 2018.
- RADU G, MATTHEW L H, ROY R. Elasticsearch in action[M]. Translated by HUANG S. Beijing: Posts and Telecom Press, 2018.
- [20] 施瓦茨男爵, 彼得·扎伊采夫, Vadim Tkachenko. 高性能 MySQL[M]. 宁海元, 周振兴, 彭立勋, 等译. 北京: 电子工业出版社, 2013.
- BARON S, PETER Z, VADIM T. High performance MySQL[M]. Translated by NING H Y, ZHOU Z X, PENG L X, et al. Beijing: Electronic Industry Press, 2013.

[作者简介]



明萌 (1994-), 女, 中国电信集团有限公司中级项目经理, 主要研究方向为智慧家庭业务感知优化、大数据和 AI 技术。



时宇 (1995-), 女, 中国电信集团有限公司初级项目经理, 主要研究方向为智慧家庭业务感知优化、大数据和 AI 技术。



郑润含 (1975-), 男, 中国电信集团有限公司高级项目经理, 主要研究方向为家庭带宽业务感知评价体系、家庭带宽网络质量分析和优化、无线网络优化、无线网络和固定网络融合等。

冯云喜 (1976-), 男, 现就职于中国电信集团有限公司, 主要研究方向为家庭带宽业务感知评价体系、家庭带宽网络质量分析和优化、无线网络优化、无线网络和固定网络融合等。