

# 运营商网络数据自主掌控及 场景化赋能探讨

## Discussion on Autonomous Parsing and Scenario Empowerment of Operators' Network Data

范琨,耿岩,高晓芳(中国联合网络通信集团有限公司,北京 100033)

Fan Kun, Geng Yan, Gao Xiaofang (China United Network Communications Group Co., Ltd., Beijing 100033, China)

### 摘要:

提出了运营商网络数据自主掌控及场景化赋能体系,以数据能力开放二级架构为管理保障,分层分级盘活网络数据。以场景化应用为牵引,通过数据自主掌控及跨域融合能力提供技术内生动力,实现数据对业务场景智能化赋能,激活数据要素活力,释放数据价值,强化自身数字化运营能力提升,赋能多行业的数智化发展。

### 关键词:

数据解析;场景化赋能;数据要素价值;运营商

doi: 10.12045/j.issn.1007-3043.2024.03.009

文章编号: 1007-3043(2024)03-0042-06

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:

A system for autonomous analysis and scenario based empowerment of operators' network data is proposed, with an open two-level architecture of data capabilities as the management guarantee to achieve hierarchical and hierarchical revitalization of network data. With scenario based applications as the traction, data autonomous analysis and cross domain fusion capabilities provide technological endogenous power, achieving intelligent empowerment of business scenarios by data, activating the vitality of data elements, releasing data value, and promoting the digital and intelligent development of the entire society.

### Keywords:

Data autonomous parsing; Scenario empowerment; Data element value; Operators

引用格式: 范琨,耿岩,高晓芳. 运营商网络数据自主掌控及场景化赋能探讨[J]. 邮电设计技术, 2024(3): 42-47.

## 0 引言

当前,数字技术与实体经济的深度融合不断加强,各行各业数字化转型持续向纵深推进,党和国家对发展数字经济提出更高的要求。数据要素作为数字经济的核心资源,对推动经济增长有放大、叠加、倍增作用。当前数据要素市场的运行规律和实践仍处于探索期。2023年3月,中共中央和国务院发文组建国家数据局,重新梳理数据生产要素,统筹推进数据资源整合共享,进一步加强数据流通,充分释放数据价值。电信运营商作为数字信息基础设施建设、运营和服务的国家队,日处理海量的通信数据,是数

据要素市场的重要参与者,在开展数据要素价值挖掘、创新数据场景化变现等方面有天然优势。

当前,运营商面临数据体量持续增长、价值深入挖掘难度大、解析技术自主性低等问题,数据的价值释放和创新应用不足。本文聚焦通信运营商在挖掘数据要素价值方面的实践,基于当前面临的各种问题和挑战,提出网络数据自主掌控及场景化赋能体系,探索和释放电信数据红利,推动数据要素市场健康发展。通信运营商在数字经济发展中将发挥“大国顶梁柱”作用,为自身数字化转型和千行百业数字化转型贡献更大力量。

## 1 当前运营商数据面临的挑战

5G商用以来,网络建设已基本实现室外连续覆

收稿日期: 2024-02-06

盖。根据工信部数据,截至2022年底,我国累计开通5G基站总数达228.7万个,占全球60%以上,已实现全国所有城区和超97%乡镇的5G网络覆盖。我国5G用户已达5.42亿,占全球用户近6成,5G手机出货量累计6.15亿部,占比超过80%。超高清音视频通话、XR、云游戏、超高清直播等个人应用愈发丰富,逐渐改变居民日常生活方式。《5G应用“扬帆”行动计划(2021-2023年)》发布以来,5G与各行业的融合应用蓬勃发展。目前,5G在文旅、物流、教育、智慧城市、安防、农业等行业均有融合应用场景的探索。5G时代,运营商网络价值凸显,电信业正在成为全社会数字化转型的底座和基石。与此同时,随着运营商数据体量的持续增长、价值深入挖掘难度提高,运营商在数据运营方面也面临着一系列的困难与挑战。

### 1.1 体量庞大且持续增长

运营商作为数字经济的底座和基石,有数亿的网络接入设备和丰富的服务场景,网络数据体量庞大,并且随着5G的规模化发展,该体量仍在持续增长。2018年至2021年,网络数据体量年平均增幅达到35.5%,未来还将持续快速增长。

尽管DPI技术在网络数据采集中得到大量应用,目前全国实现集中采集的网络数据日均约为500TB,其中DPI数据占80%。但网络数据体量庞大,对存储和算力资源的要求过高,仍有大量数据未被采集和解析。未被采集的网络数据包括TRACE数据、eDPI数据,约有800TB和1400TB。如何有效平衡海量数据的处理效率和存储压力成为制约数据价值释放的关键因素。

### 1.2 数据价值解析不足

当前部分网络数据已投入生产应用,但仍有多类数据的价值需深入挖掘。目前已应用于生产场景的数据包括CM/PM、MRS、XDR、告警等,TRACE、MRO、而MDT及eDPI、DPI原始码流数据仍处于研究探索阶段,尚未形成研究成果赋能生产应用。

随着物联网、车联网、5G虚拟专网等新产业形态的出现和不断成熟,垂直场景的专业壁垒不断提升及产业需求的快速迭代,针对场景化的应用开展数据弹性解析,将O域、B域、固网、物联网、车联网、行业MEC等数据中与场景需求紧密相关的部分解析出来,实现跨域数据的融合打通,研究探索行业融合、人物互联、固移融合和BO融合等跨域融合成为数据要素场景化创新应用的必然趋势。

### 1.3 自主掌控及融合能力仍受制约

运营商数据包含海量信息,尤其是其中的网络数据,每日产生海量的网络运行和用户上网等各类日志数据。但是数据解析受到第三方厂家“卡脖子”等问题的制约,目前运营商在原始数据采集后的解析和处理环节自主性较弱。在面对业务一线的各类复杂实时场景时,按需采集、按需解析、按需融合的能力不足,影响数据在价值挖掘和创新应用方面的探索和灵活应用。

因此,在数字化不断发展的产业互联网时代,电信运营商的服务对象由个人向行业转变,网络形态和业务模式更加复杂多样。面对灵活多变的应用场景和持续增长的海量数据,运营商需要探索构建新的数据赋能体系,统筹数据统一集约与需求敏捷响应,提升场景驱动的注智赋能服务能力,充分盘活电信运营商数据,激发数据活力。

### 1.4 灵活性低限制一线数据应用

当前中国联通的全国网络数据一点集中,形成资源能力、数据能力和数据专题结果输出能力向省分赋能。其中资源能力以租户方式向省分开放,省分按需订阅汇聚层数据,专题结果由省分订阅开放。省分主要提供原始码流等数据分光、镜像等统一采集点,本地不做详单留存。由此带来数据取用时灵活性下降的问题,当一线有个性化需求时,灵活匹配能力较为欠缺。

在现有架构下,网络数据经过末端采集、省分转发、集团中台处理、数据订阅处理等多次迂回,增加不必要的时延及开销,难以满足一线业务分析对实时性的要求。例如,投诉处理场景中,业务感知、网络质量、位置轨迹数据的高时效性对问题的根因分析至关重要。现有架构下O域DPI数据、CM/PM、MR、TRACE经网络中台统一采集再迂回到省分网优一线的机制不合理,不利于客户投诉的及时处理。

## 2 网络数据自主掌控及场景化赋能体系建设

### 2.1 场景化赋能体系

场景是实现数据要素价值的核心要素,场景化数据赋能体系以数据开放二级架构为制度管理保障,以网络数据自主解析技术为技术引擎,通过建立底层数据采集、自主解析、跨域融合和场景化应用的联动体系,实现敏捷的数据提取、动态智能的场景数据匹配,快速响应场景化业务应用,在保障网络运行的基础

上,充分挖掘数据中的业务价值,促进降本增效,激活数据要素价值,整体架构如图1所示。

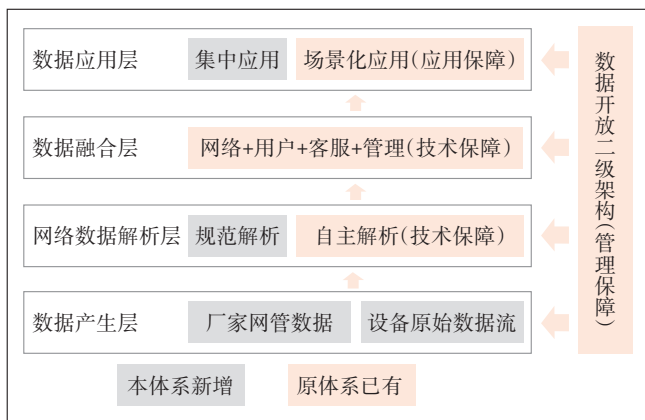


图1 场景化赋能体系架构

## 2.2 数据能力开放二级架构

网络数据体量庞大,使存储和解析压力倍增,在全国层面进行全部采集解析的投入成本高、收益低,运营商的数据自主解析能力不足进一步制约数据的定制化采集。但当前场景逐渐细分,需求逐渐深入,“一点集中按需订阅”的数据开放方式难以满足一线灵活性的场景化需求。部分场景特需的数据难以得到满足,大量数据反而无法使用,亟需在数据管理架构上进行调整,进一步提升运营商的数据自主解析能力,让运营商数据在企业内部得到充分灵活的使用。

因此,设置数据开放二级架构,实现数据统一集约和灵活运营的有效平衡。该架构下,集团公司重点建设全网的宏观分析、监控和推广,省分公司建设个性化的生产应用和本地化的实时场景应用。“集团+省分”优势互补,实现复杂问题闭环,充分盘活运营商大数据,降本增效,放大数据赋能价值,加快运营商数字化转型和产业赋能。

数据开放二级架构在现有数据管理架构的基础上进行逐渐演进,具体路线如图2所示。在现有数据架构基础上,新增省分部分。以省分前后台场景化需求为牵引,接入本地码流数据,按需解析压缩,构建本地化实时集群,实现本地数据的跨域融合,实现对场景化需求的跨域服务,实现一线层面的数据驱动智能运营,提升省分公司的数智化运营水平。随着场景化应用需求和集中应用的不断演变,部分全国性应用固化为集中统一采集,上报到集团中台,通过按需订阅方式调用。具有省分场景化需求的特色数据实现按需自主解析和场景化分析,释放数据价值。

## 2.3 场景化赋能体系的关键技术能力

### 2.3.1 场景化驱动能力

数据的应用包括集中应用和场景化应用。集中应用主要是面向全国性的赋能应用,基于IT集约化的数据底座来实现,场景化应用主要是面向垂直行业或一线生产的特定场景。特定场景对数据的需求有范围广而字段少的特点,应用的时效性更高,需要做出快速敏捷的响应,开展场景化的试点应用和反复迭代验证。因此,从数据应用层出发,驱动数据的定向跨域融合和自主解析,在满足集中应用需求的同时,充分激发网络数据的活力,具体的驱动方式包括如下3个过程。

- 场景化应用的驱动路径。主要通过设定匹配场景的数据需求来驱动网络数据按需弹性自主解析。
- 自主解析能力对场景化应用赋能的直接路径。根据场景化需求弹性定制解析,生成网络数据话单,在融合数据底座中进行跨域数据融合,满足场景化需要。该路径主要针对场景化应用的试点探索和验证阶段,能够快速迭代和分析,实施周期短且时效性高。
- 自主解析能力对场景化应用赋能的间接路径。

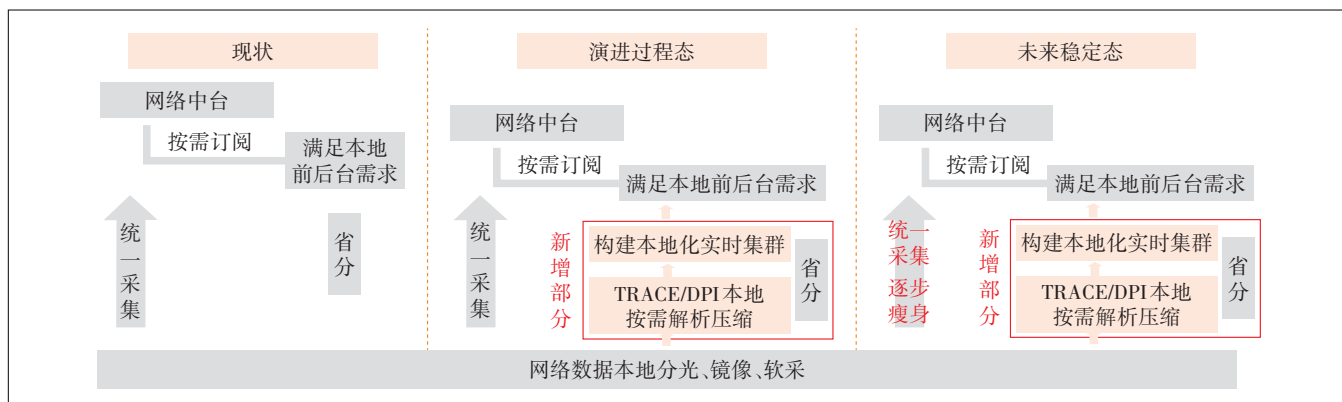


图2 数据开放架构演进路线示意

将自主解析能力内化到规范的解析流程中,对数据解析规范进行升级和补充,实现定制化解析能力的全国性统一。在此基础上完成全国层面的数据跨域融合,将试点验证过的场景化应用进行全国性推广,形成集中应用,并进行快速大规模复制,实现数据赋能的价值倍增。

### 2.3.2 数据自主解析能力

规范解析主要通过统一DPI采集规范实现,面向全网共性需求。但在各类创新场景中,灵活定制的需求日益增多,采用规范采集方式难以快速满足要求。像TRACE和全量XDR数据每日产生PB级的原始数据,在无明确的全网性需求时进行全网采集成本过高,收效不明显。同时,网络数据的采集能力依赖设备厂商,自主研发能力较弱,制约问题的闭环解决。因此,在省分一线场景端开展按需采集、弹性解析,根据具体的场景需求进行定制,能够实现对需求进行快速和实时的响应。

数据自主解析核心能力架构如图3所示,主要通过自主研发数据定制管理器、设备码流自主解码器、弹性自主定制能力平台来实现。数据定制管理器主要是接收来自场景化应用和集中应用的数据定制需求。通过以上各模块的协同配合解决处理效率和存储压力的问题,实现数据价值的高效提取。

根据定制类型的不同,定制需求可以进一步细化

为2个部分。一部分是直接从设备原始码流开始,定制解码元素,一部分是针对抽取内容进行按需自主定制。定制解码元素是将数据需求直接深入到原始二进制流的解析过程中,这一部分的定制主要是需求在协议层面的定制,由设备码流自主解码器实现。定制抽取内容是在原始码流解析生成话单的过程中,根据需求在加工层面的定制,主要由弹性自主定制能力平台实现。弹性能力自主平台根据数据定制管理器提供的需求,在码流解析和关联回填时进行字段抽取,在此基础上,对有效话单进行信息压缩,降低对传输带宽的要求,整个按需抽取和压缩输出的过程采用流式处理,确保加工的实时性,支撑实时场景应用。

自主解析的网络数据主要包括三大类:TRACE、MDT/MRO、XDR数据。不同数据的自主解析能力特点也不同。图4为TRACE数据的自主解析流程。TRACE原始码流数据导入后,通过配置数据管理实现按需弹性解码配置,可解析RRC、RRC\_NR、S1AP等8种协议;通过中心解码调度模块完成消息信元级的实时调度,实现解码输入的最大灵活性;最后通过ASN1文档编译和核心解码输出模块实现按需定制和弹性解码,直接跳过不需要的码流片段,从根本上提升解码效率。

MDT/MRO原始码流数据,主要通过解压、配置及解码过程实现MDT多类数据的自主解析、输出字段的

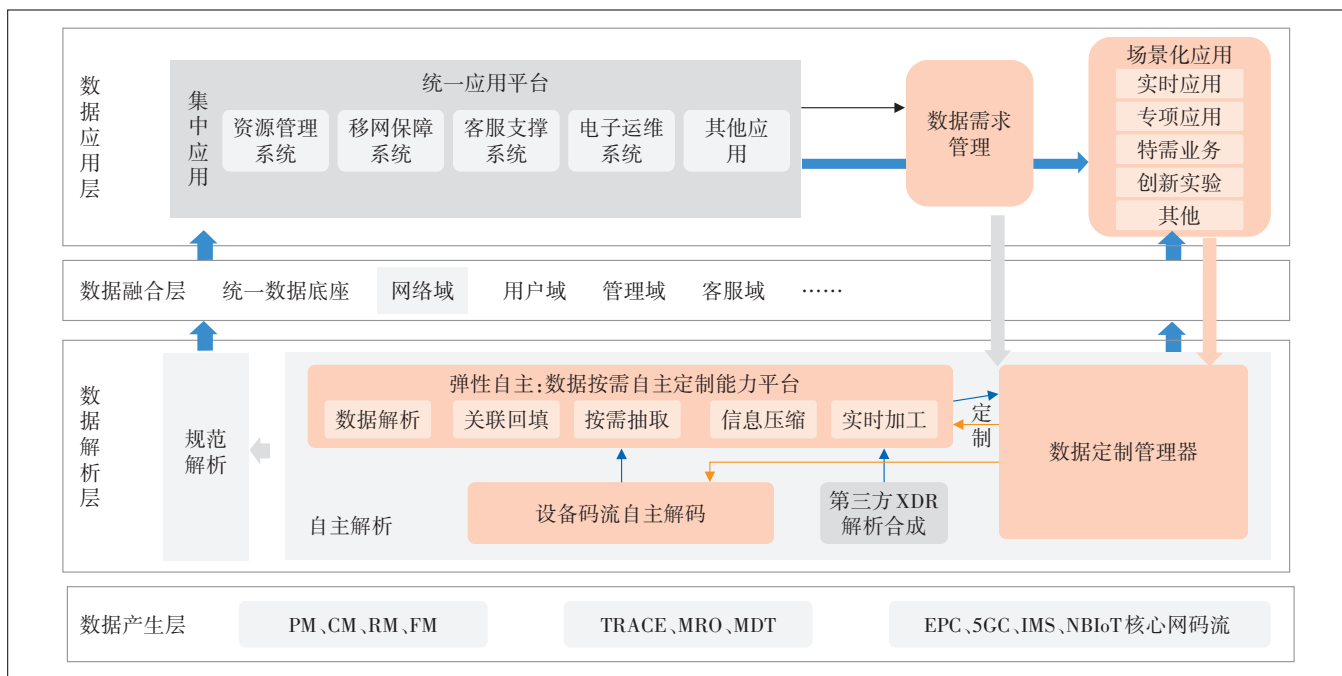


图3 数据自主解析核心能力架构

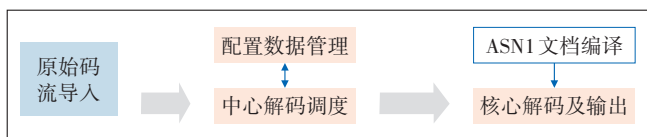


图4 TRACE数据自主解析流程

按需配置,进一步提升空口数据中网络异常事件失败原因值、经纬度等信息解析能力,助力一线网络问题根因定位闭环解决。

移网及固网XDR原始码流数据的解析,主要是通过特征解析配置实现价值信息深度挖掘,增强数据模型支撑一线的有效性。图5给出移网/固网原始码流数据的解析验证流程。其中特征解析配置环节是核心自主研发部分,能够有效提升用户行为数据识别的准确度。通过DPI业务特征解析实现对业务特征的获取,进一步形成统一业务特征库,并构建验证机制实现特征的迭代和优化。最终按需进行特征的灵活配置并生成定制单据,快速响应各类场景需求。

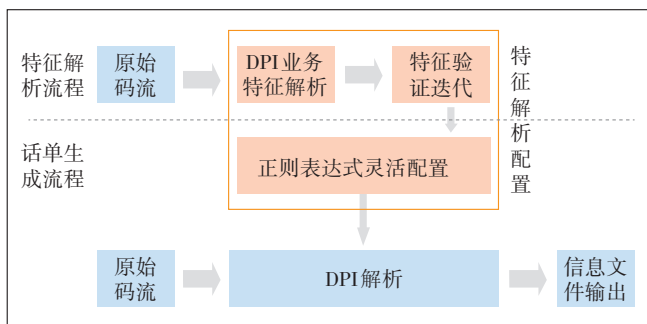


图5 移网/固网XDR原始码流解析验证流程

### 2.3.3 数据跨域融合能力

数据的融合包括2个层次,一个是在统一集约层面,将规范解析后的标准话单与其他域数据进行融合。以用户ID为主键,在全国性范围内实现数据的跨域拉通,主要针对全国性的集中应用,包括指标监控和统一结算等。另一个是在一线场景定制层面,搭建融合数据底座实现O域、B域数据的自动化ETL处理,实现数据拉通、离线/实时关联分析。按需弹性输出的定制话单,并与必要的其他域数据进行融合。以场景化应用为牵引,以用户ID、小区ID或产品ID等为主键,构建融合数据话单,为场景化应用提供全方位的数智化服务。针对一线的个性化或实时场景,构建用户融合标签体系,融合移网、固网、B域数据特征,能够实现用户全景和全天候洞察,有效支撑实时营销等前端场景。打通O域、B域及投诉数据,构建投诉问题实

时回溯及自动化分拣能力,有效支撑网络投诉问题的高效处理。

## 3 场景化赋能体系典型应用成效

过去3年面对新冠肺炎疫情,运营商的数据价值凸显。场景化数据赋能体系在部分省分开展试点应用,实现数据能力与一线各业务环节的深度融合,充分发挥数据要素的赋能价值。

### 3.1 面向市场赋能

省分在市场经营过程中,场景化、个性化需求较多,在使用集团提供的数据时遇到标签与用户的匹配程度不足、缺乏定制化数据等问题,难以精确满足个性化需求;可用的数据和标签维度/粒度比较单一,难以全面洞察用户行为特征;周期性差异化营销策略的制定需要数据和模型的快速响应,以当前能力很难快速构建和迭代。

基于本体系思路展开专项攻关,实现市场重点业务发展场景的精准赋能。例如,某省分在疫情期间针对网课用户开展宽带营销。基于本体系构建需求快速响应流程(见图6),在接到需求后第一时间开始自主解析,并进行跨域数据融合研发,在需求提出5日内识别并输出一批网课用户列表,同步一线开展培训和试点验证。20日内完成模型迭代优化,快速输出新一批网课用户列表。通过在本地搭建实时计算集群,快速完成O域和B域数据的融合及用户融合标签体系构建,实现网课群体精准挖掘、快速响应和验证优化闭环,有效提升宽带产品的办理率。

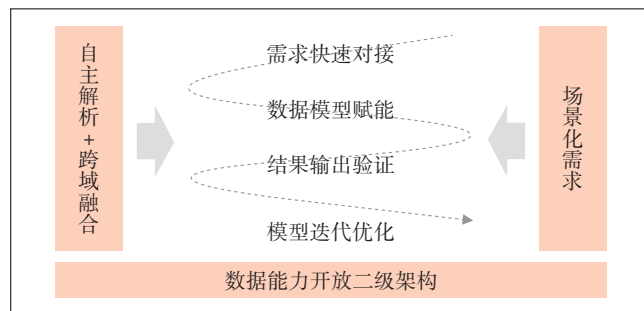


图6 需求快速响应流程

### 3.2 面向网络赋能

现有网络资源和属地化运营的特性要求需要场景化、差异化的云网运维优化方案来保障网络运营质量和用户满意度的提升。但是当前用户主观感知体验与客观网络指标之间存在脱节问题,网络优化大多

关注设备及指标,缺乏用户级端到端感知评价手段。网络优化工作大量依赖人工操作和人工经验,仍缺少有效的平台工具,智能化水平较低。

基于场景化赋能体系形成网络问题分解定位思路(见图7),通过网络数据自主弹性解析实现价值深度挖掘,通过B域、网络数据的跨域拉通支撑各类优化问题的分解研究。构建用户级感知评价、回溯分析及实时预警,实现端到端问题精准定位及根因回溯。通过问题分解驱动多部门联动,共同推进网络精准优化和网络服务提升。

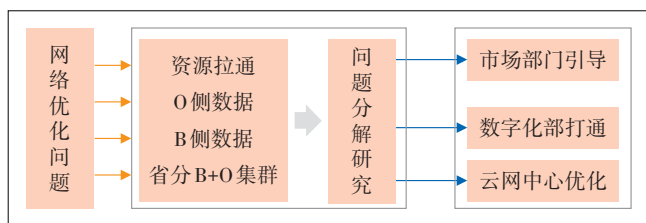


图7 网络问题分解定位思路

### 3.3 面向客服赋能

当前客服在处理用户投诉时,面临投诉内容的价值挖掘和发挥不够、投诉解决手段工具少、效率和智能化程度不足等问题,用户的问题定位和响应时间受到严重制约,客服处理投诉工单的效率低,影响网络问题根因诊断和问题回溯,因此亟需提升投诉分析和挖掘的前后端数字化水平。

通过数据的跨域融合,构建智能化投诉语义理解、问题分拣、投诉画像及时空轨迹感知洞察,实现回溯并复现用户描述问题、网络问题自动化诊断和定位定界,为一线投诉处理人员及省分网络问题定位处理提供有效支撑,投诉处理时长缩短70%,投诉定比压降及满意率定比提升均高于目标设定值,有效提升网络问题定位优化及管理效率(见图8)。

## 4 总结

数字中国、网络强国战略不断推进,社会的数字化、网络化和智能化水平持续提升,各行各业数字化

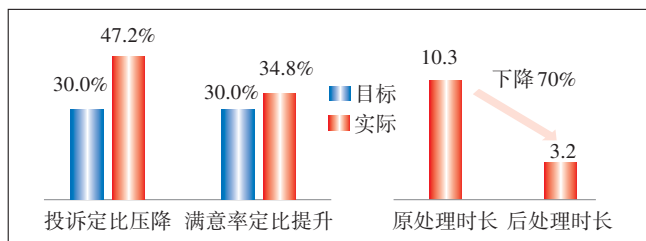


图8 投诉处理效果提升示意

转型进入关键期。国家数据局的组建将进一步加快数据要素市场的发展,数据要素价值将在数字化转型关键期发挥更大作用。

运营商在企业数字化转型中扮演“排头兵”和“筑基石”的双重角色,网络数据价值凸显。在加快推动企业内部数字化转型的同时,通过与垂直行业的深入融合向toB用户提供个性化和实时化服务等方式,为千行百业的数字化转型提供网络、数据和算力支撑。

数据自主解析及场景化数据赋能体系将以数据开放二级架构为保障,以内外部的场景化需求为牵引,通过数据自主解析、跨域融合技术驱动构建场景化应用智能引擎,能够为一线提供全面的智能化支撑和个性化快速服务,实现数据在业务场景、产品研发、渠道运营等多环节的贯通。场景化的数据赋能将进一步发挥数据要素活力,沉淀数据价值挖掘和赋能经验和能力,实现数据要素价值的放大、叠加和倍增,使数据赋能新生态,促进数字经济与实体经济的深度融合,从而推动社会智能化发展。

### 参考文献:

- [1] 刘烈宏. 创新引领赋能产业助力数字经济健康发展[J]. 中国产经, 2022(7): 48-50.
- [2] 中国信息通信研究院. 电信业数字化转型发展白皮书(2022年)[R]. 北京: 中国信息通信研究院, 2022.
- [3] 中国信息通信研究院. 中国5G发展和经济社会影响白皮书(2022年)[R]. 北京: 中国信息通信研究院.
- [4] 程新洲, 贾玉玮, 成晨. 数据驱动的网络智能运营体系思考[J]. 信息技术与政策, 2021, 47(6): 68-73.
- [5] 程新洲, 朱常波, 晁昆, 等. 掘金大数据[M]. 北京: 机械工业出版社, 2019.
- [6] 中国信息通信研究院. 取数之道[M]. 北京: 人民邮电出版社, 2017.
- [7] CHENG XINZHOU, XU LEXI, ZHANG TAO, et al. A novel Big Data based Telecom Operation architecture [C]//Proceedings of the 1st International Congress on Signal and Information Processing, Networking and Computers (ICSINC), 2016: 385-396.
- [8] 张晴晴, 张涛, 韩玉辉, 等. 基于大数据分析的潜在固移融合用户挖掘研究[J]. 邮电设计技术, 2021(11): 31-34.

### 作者简介:

范琨, 高级工程师, 硕士, 主要从事网络线科技创新管理、网络新技术研究、数字化转型、网络业务优化等方面的工作; 耿岩, 高级工程师, 学士, 主要从事网络线科技创新管理、网络新技术研究、数字化转型、网络业务优化等方面的工作; 高晓芳, 工程师, 硕士, 主要从事网络线科技创新管理、网络新技术研究、数字化转型、网络业务优化等方面的工作。