

面向5G的物联网大数据分析

Research on 5G Oriented IoT Big Data
Analysis Method System

方法体系研究

张涛,程新洲,徐乐西,高洁,曹丽娟(中国联通网络技术研究院,北京100048)

Zhang Tao, Cheng Xinzhou, Xu Lexi, Gao Jie, Cao Lijuan (China Unicom Network Technology Research Institute, Beijing 100048, China)

摘要:

5G将开启万物互联新时代。5G在为万物互联提供连接服务的基础上,更大的价值在于海量物联网行业终端接入带来的海量数据,在于隐藏在海量数据金矿中大量具备行业属性的终端行为。因此提出了一种面向5G物联网大数据分析的方法体系,探索如何更好地分析、挖掘5G物联网大数据来为各垂直行业赋能,全面提升5G物联网大数据应用价值。

关键词:

5G;物联网;大数据;分析

doi:10.12045/j.issn.1007-3043.2020.05.003

文章编号:1007-3043(2020)05-0013-05

中图分类号:TP391

文献标识码:A

开放科学(资源服务)标识码(OSID):



Abstract:

5G will open a new era of IoT (Internet of Things). On the basis of providing connection services for IoT, 5G's greater value lies in the mass data brought by the access of mass IoT terminals, and in the large number of terminal behaviors with industrial attributes hidden in the gold mine of mass data. Therefore, it proposes an analysis method system for 5G IoT big data, and explores how to better analyze and mine 5G IoT big data to better serve various vertical industries, so as to comprehensively enhance the application value of 5G IoT big data.

Keywords:

5G; IoT; Big Data; Analysis

引用格式:张涛,程新洲,徐乐西,等.面向5G的物联网大数据分析体系研究[J].邮电设计技术,2020(5):13-17.

0 前言

5G网络商用牌照的发放,标志着我国正式进入5G商用元年,5G产业链在牌照颁发后也得到了迅猛发展。信通院公开数据显示,截至2020年2月初,三大运营商共在全国开通5G基站约15.6万个,网络建设顺利推进。同时在5G终端的发展上,截至2020年2月初,我国已有43款5G手机终端入网,手机价格下探至3000元区间,大大降低了用户的5G准入门槛。另外

从5G应用的角度来说,5G的高带宽、高速率、低时延及海量高可靠连接的特性将驱动更多新应用场景及新业务的产生。总体上来说,5G网络的到来,将真正实现人与人、人与物、物与物的互联互通,创造万物互联的新时代。

运营商在万物互联新时代将迎来黄金发展机会,物联网终端的海量接入以及由此产生的海量数据,将为运营商物联网发展“连接+数据服务”创造巨大市场机会。据IoT Analytics咨询机构预测,运营商物联网连接数呈指数级增长,预计2021年全球的物联网连接数将超过人联网连接数。所以对于运营商来说,人口红利结束后,在5G时代将迎来海量物联网终端接入的

基金项目:国家重点研发计划资助项目(2018YFB1800800)

收稿日期:2020-03-23

红利时代,未来的业务增长点也聚焦于具有巨大市场空间的产业互联网2B市场。5G将成为各行业数字化转型的重要引擎,越来越多的传统企业将会把他们的产品、服务接入互联网,由此会产生海量的行业异构数据,这就为运营商探索5G物联网大数据运营提供丰厚的土壤环境。

对于中国联通来说,在5G为万物互联提供连接服务的基础上,更大的价值在于海量数据的运营,在于隐藏在海量数据金矿中大量具备行业属性的终端行为。因此在5G物联网大数据呈现井喷式爆发之前,亟需建立起5G物联网大数据分析的方法体系,探索如何更好地分析、挖掘5G物联网大数据来为各垂直行业更好地服务,从而全面提升5G物联网大数据的应用价值。

1 5G物联网大数据的机遇与挑战

与2G、3G、4G网络的服务主体为人联网不同,5G将开启万物互联、万物智联的新时代,服务对象也将变为以物联网为主体。图1描述了中国联通2G/3G/4G连接数前10的省份截至2019年9月份的物联网和人联网连接数的对比统计情况,可以看到在当前阶段各个省份人联网连接数占据绝对主体,连接数占比都在95%以上。而在一份IoT Analytics的咨询报告中显示,随着5G产业链的逐渐成熟,预计2021年物联网连接数将超过人联网连接数(见图2)。因此在未来5G时代,5G物联网大数据将呈现出很多新的特征和特点,也给运营商带来了很多新的机遇与挑战。

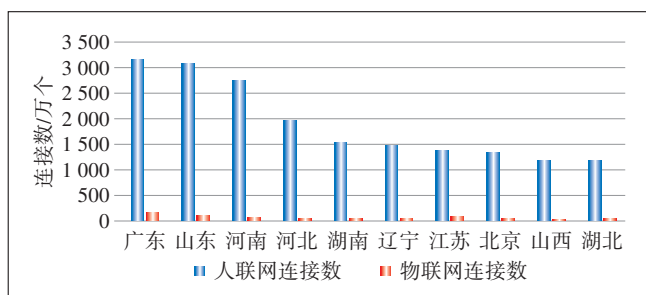


图1 中国联通人联网和物联网连接数对比

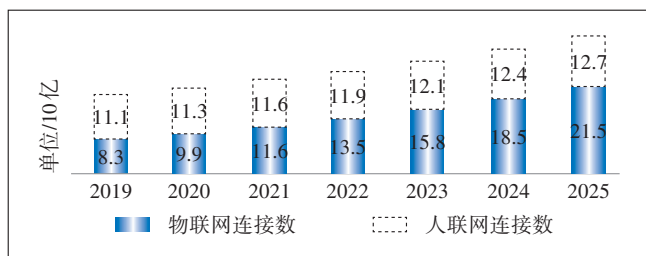


图2 全球人联网和物联网连接数预测

首先,5G网络中海量设备的接入以及更快的传输速率,将带来数据量的急剧增长。海量数据为运营商带来更大价值金矿的同时,也给运营商存储、传输、处理、挖掘海量数据带来了更大的代价和成本。所以运营商需要在5G物联网数据全面爆发前,对当前的大数据平台处理架构和流程进行重塑和升级。

其次,万物互联的开启,使得数据维度和种类会更加丰富,会产生很多诸如视频、图片等非结构化数据;众多行业数据的接入也会使得物联网大数据的应用场景更加丰富。图3展示了中国联通截至2019年8月份的物联网连接数在各个行业的分布,其中连接数占比最多的为车联网行业,达到64%;其次为运输与物流行业,占比为11%;后面依次为消费电子(6%)、能源与公共事业(6%)、零售与商业(6%)等。面对众多垂直行业,而且行业间业务多样化、差别大,场景碎片化更加严重,物联网行业解决方案难以形成规模效应。同时运营商也迎来了巨大的机遇,在掌握各行业大数据后,可以借此机会进入垂直行业领域,推动各行各业的数字化转型升级,开辟新的利润增长点;然而运营商的短板在于缺乏大量的垂直行业的专业知识和专家人才。运营商可以通过与行业客户进行深入合作,也可以通过自行储备培养垂直行业人才来补齐短板。图4从市场成熟度以及市场空间2个维度展示了各垂直行业的吸引力,运营商物联网大数据可在技术相对成熟、经济价值大的行业优先发力,如车联网行业、工业互联网行业等典型行业优先入局,后面随着时间的推移,为更多垂直行业赋能,将物联网、大数据、云计算、人工智能等技术设计与行业设计、生产、管理、服务等各个环节进行深度融合,在更多行业凸显5G物联网大数据价值。

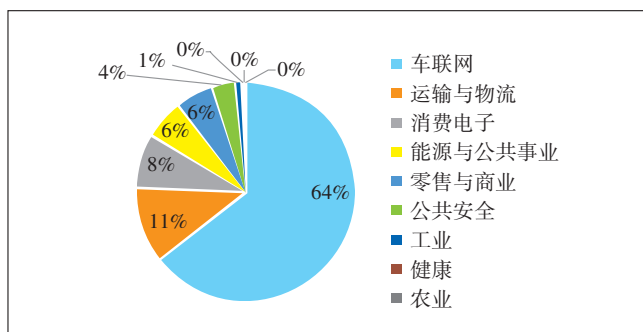


图3 中国联通物联网行业连接数占比

再次,物联网管道的业务识别会更加困难,物联网针对不同行业的应用层协议没有形成统一的标准,每个行业、甚至每个企业都使用不同的协议标准。物

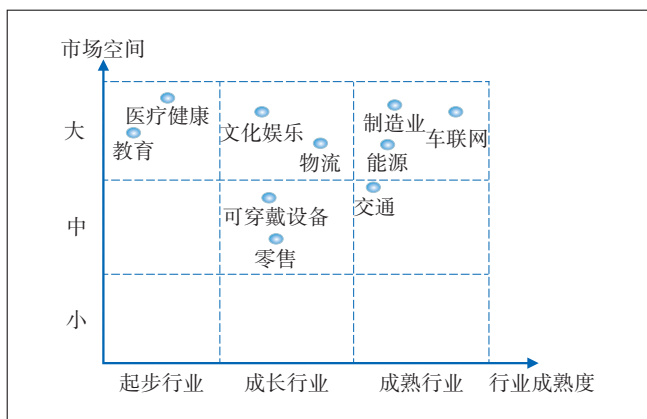


图4 运营商物联网重点突破典型行业

联网常用的应用层协议包括MQTT协议、CoAP协议、HTTP协议等,从当前物联网应用发展趋势来分析, MQTT协议具有一定的优势。物联网业务众多的应用层协议,对于运营商通过DPI深度包解析来对流经管道的物联网业务数据流进行流量应用识别提出了巨大挑战。

2 5G物联网大数据平台架构演进

5G时代,随着大数据急剧增长、数据维度和数据类型越来越多样化,对于大数据平台的要求会进一步提升,对于数据存储效率、数据ETL、计算算力、实时计算引擎、数据批处理引擎等提出更高的要求。传统的单一计算中心平台将难以有效应对如此复杂、多样、海量的数据采集、清洗、处理和挖掘带来的挑战。对于中国联通来说,4G时代全国数据单点集中,这种数据全集中的方式在5G海量数据时代将未必是最优的解决方案,海量数据的集中存储、计算和挖掘将带来巨大的硬件投资,分析处理的效率会大大降低,同时也会面临数据单点故障的巨大安全隐患。此外在今后服务众多垂直行业客户时,单点集中的方式不利于满足不同行业客户个性化的需求和业务应用场景,如车联网场景(对时延要求较高)和AR/VR场景(对业务速率和时延要求都较高)对平台的计算能力和响应速度的要求都是不同的。因此,在本文中,探讨了如何提前重塑当前的大数据平台架构,同时对于物联网大数据处理的流程进行了重新梳理。

5G时代的物联网大数据平台架构将向中心+边缘协同计算的方向演进。边缘计算将作为物联网大数据平台的“神经末梢”,旨在将一定规模的数据分析功能下沉到更靠近应用的地方,从而提升响应速度,避免所

有数据都上传到中心平台侧,给中心平台造成巨大的压力;而中心平台作为物联网大数据平台的“大脑”,会将边缘计算的结果回传到中心平台汇总统一整理,同时对于边缘计算无法处理的数据、算法、模型等进行处理、分析之后,再下发到边缘节点,此外中心平台也负责对边缘大数据平台节点的管理以及算法升级等工作。中心平台和边缘节点之间的协同工作模式,就成为5G物联网大数据的主流数据处理模式,二者只有紧密协同才能更好地满足各种需求场景的匹配。

图5展示了5G时代物联网大数据处理的基本流程。物联网行业终端产生海量数据并传输到离自己更近的边缘大数据节点,边缘大数据节点会根据业务场景的需求进行实时、短周期的局部数据统计挖掘分析,并将结果反馈到物联网行业终端,从而快速响应需求;而对于全局数据进行加工处理的需求,边缘大数据节点会与中心大数据节点进行交互,提供经过加工分析后的高价值数据,相当于为中心大数据节点做了数据的预处理服务,一方面极大地节省了网络带宽资源,另一方面降低了中心节点的计算压力。当中心大数据节点通过对全局数据的挖掘处理后,再将结果通过边缘计算节点反馈到物联网行业终端,从而完成物联网大数据处理的闭环。

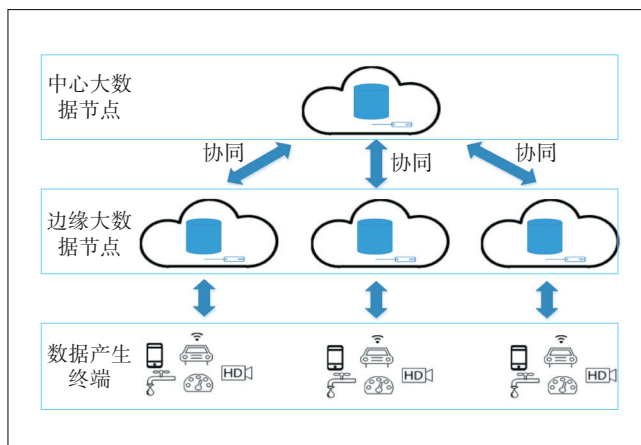


图5 5G物联网大数据处理流程示意图

从5G物联网对于大数据平台的演进要求来说,为了满足不同场景下的业务需求,各级数据中心要向智能、弹性、自适应、开放的分布式架构演进。要能够提供大规模的弹性存储容量与高效的计算能力支撑,平台可以根据行业自适应配置和使用不同的智能引擎,结合该行业领域的背景知识与数据建模,从海量数据中快速学习和提炼出高价值信息与策略。

3 5G物联网大数据分析体系

在前5G时代,运营商大数据以人联网为主体,积累形成的人联网大数据分析体系在5G物联网大数据时代不再适用。因此,本文旨在建立5G物联网大数据分析体系,探索如何更好地分析、挖掘5G物联网大数据来为各垂直行业赋能。

3.1 5G物联网大数据分析流程

5G物联网将服务于千行百业,而行业间的业务多样化、差别大,各行各业物联网的不同业务场景,将会呈现出不同的规律和特点,所以对于物联网大数据分析,需要针对于不同行业来分别进行处理。图6给出了对于5G物联网大数据的简要分析流程。

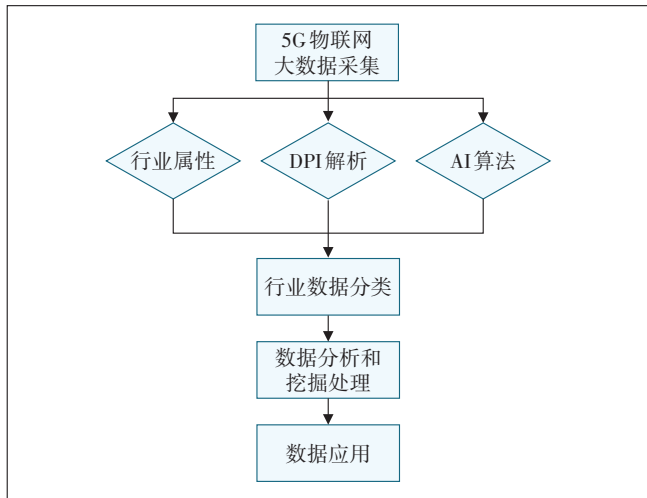


图6 5G物联网大数据分析流程

对于5G物联网大数据分析,最重要的是建立典型行业数据识别特征库,实现行业数据识别,这是开展后续处理、挖掘分析工作的基础。本文中提出可以通过属性标识、DPI协议特征识别或者AI智能算法3种方法来实现行业数据识别。

第一种方法是通过行业属性信息去识别,类似于人联网的实名认证。行业企业在入网时,运营商会根据其行业类别为该企业的所有物联网SIM卡分配一个行业标识。对于采集到的O侧物联网数据,可以根据IMSI或者MSISDN与物联网的用户管理平台中的数据进行关联,从而得到该物联网卡的行业类别。

第二种方法是通过DPI深度包解析技术,对物联网原始码流进行识别,通过积累形成的IP规则库、端口规则库、传输层&应用层解析规则库等,对物联网数据的业务内容进行识别,企业信息进行提取等从而识

别出行业类别。

第三种方法是通过人工智能AI算法来进行行业识别。物联网在不同行业的业务场景下的业务特征会存在显著差别,如车联网行业物联网卡的终端类别、传输时延、速率、流量、传输周期、是否高速移动等指标特征,与工业互联网中相同的指标会呈现出截然不同的规律和特点。可以根据对于不同行业数据的观察和积累,形成不同行业的指标特征库,对于采集到的物联网数据可以通过聚类或者分类的AI算法来进行行业的智能分类和识别。

完成行业数据分类后,可以根据行业类别形成基于不同行业种类的物联网行业大数据库,如可形成物联网能源大数据库、物联网车联网大数据库、物联网智慧医疗大数据库、物联网智慧交通大数据库等,从而可以针对不同垂直行业进行深度探索和研究,形成满足于该行业需求的物联网大数据上层应用及产品。

3.2 5G物联网大数据分析体系

类比4G人联网大数据分类,5G物联网大数据总体上也可以分为2类,网络侧OSS域数据以及行业属性BSS侧数据。网络侧OSS域数据重点关注通过各接口挂表采集合成得到的业务行为数据,行业属性BSS侧数据则重点记录了企业标识、开户地、营账数据及行业属性等信息。基于OSS侧和BSS侧的物联网大数据联合分析,可以从行业、业务、网络 and 模组等不同视角,实现物联网大数据的多维度关联分析,切实对运营商物联网内外部业务进行支撑,对内辅助决策,降本增效;对外探索如何为垂直行业更好地服务,创新合作提升物联网大数据应用价值,增加营收。

图7给出了5G物联网大数据分析体系的整体框架。整个分析方法体系建立在对物联网全域数据的收集,并将B侧数据与O侧数据打通,从而将行业静态数据与动态数据进行关联,完成底层数据关联层构建。然后基于B+O打通构建的关联层数据,根据典型行业数据识别特征库,构建行业、企业数据库,并根据不同行业进行标签化数据加工处理,为上层的关联分析打好基础。以上这2步为整个分析方法体系的基础工作,决定了上层专题分析和综合应用的准确度。

关联分析为整个分析方法体系的核心,主要思路是从网络、业务、行业、模组等多个维度对各垂直行业进行专题分析,从而可以对不同垂直行业进行全面分析,更有利于全面掌握各垂直行业网络运行状况、业务行为特征,发掘各垂直行业潜在需求。为满足上层百

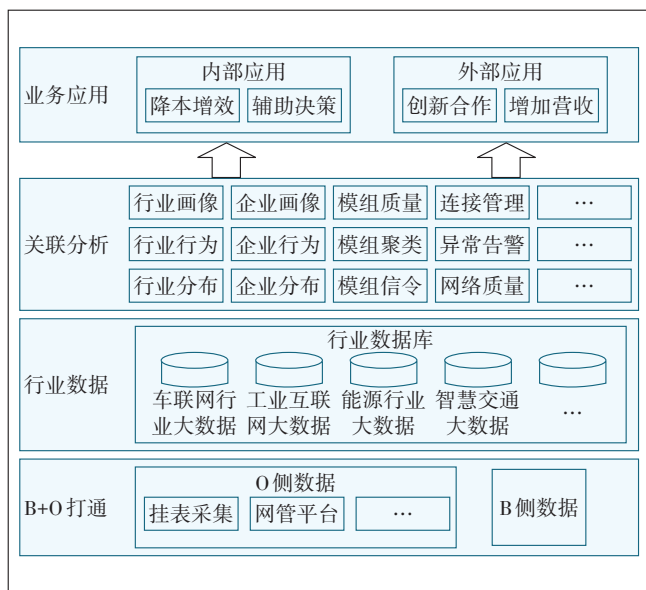


图7 5G物联网大数据分析体系总体框架

花齐放的业务应用需求,专题分析的数据分析与模型要分层构建,第一层要分行业构建多维宽表汇总数据及公共指标汇总数据,提升公共指标的复用性,减少重复的加工,通过对一些公共指标进行抽取,构建行业场景化应用标签库;第二层基于公共数据中心构建行业个性化的统计指标数据中心,直接为上层应用提供数据服务。这里指出关联分析内容的设定并不局限于图7中所包含的分析模块,可以根据行业类别和上层业务需求来进行定制化的开发。分析模块间的配合就像搭积木,不同模块间的组合可以满足特定的上层综合应用需求。如对消费电子类物联网企业,为了助力其进行市场营销,可以对企业画像、企业分布、人物互联等模块进行联合分析,得到该类消费电子产品的潜在用户以及潜在群体的消费能力,并对有消费能力的群体进行产品定向推广。再如对车联网企业,为了助力其为客户提供更优质的服务,可以对行为轨迹、业务行为等模块进行联合分析,为客户下班后的目的地进行预判并推荐避堵线路。

最上层为业务应用层,物联网时代业务应用层最大的变化为需求灵活多变,业务百花齐放,这就为整个数据底层的物联网大数据分析体系构建提出了更高的要求,也为大数据平台的可拓展性、敏捷性、轻量化,并注重与前台的交互能力提出了更高的要求。总之,整个物联网大数据分析体系的构建是需要经历业务闭环不断去完善,不断去验证和优化的过程,从而逐渐支撑越来越多的物联网行业业务应用。

4 总结

5G 的快速发展将支撑物联网在各行业迅速落地应用,物联网也将成为5G最重要的应用场景。海量的连接和多样化的应用将给运营商带来海量的物联网大数据,也会带来新的利润增长点。本文即在这种背景下,提出了一种面向5G物联网大数据分析体系,探索如何更好地分析、挖掘5G物联网大数据来为各垂直行业更好地服务,从而全面提升5G物联网大数据的应用价值。该分析体系将有助于运营商全面掌握各垂直行业网络运行状况、业务行为特征,发掘各垂直行业潜在需求,从而助力运营商不仅成为网络连接的提供者,也成为垂直行业市场的参与者和整合者。

参考文献:

- [1] IMT-2020(5G) Propulsion Group. 5G concept white paper[EB/OL]. [2020-01-12]. <https://wenku.baidu.com/view/2a32635a0066f5335b81215a.html>.
- [2] IoT analytics. Market-Growth, Trends and Forecast (2020-2025)[EB/OL]. [2020-01-12]. <https://www.giiresearch.com/report/moi389875-global-biscuits-market-growth-trends-forecasts.html>.
- [3] 中国信通院. 物联网白皮书[EB/OL]. [2020-01-12]. http://www.qianjia.com/html/2018-12/17_315737.html.
- [4] 中国联通网络技术研究院. 中国联通边缘计算技术白皮书[EB/OL]. [2020-01-12]. <https://wenku.baidu.com/view/ae34be2cbf1e650e52ea551810a6f524ccbfcfb2.html>.
- [5] Ericsson. Exploring IoT strategies[EB/OL]. [2020-01-12]. <https://www.ericsson.com/en/internet-of-things/trending/exploring-iot-strategies/download>.
- [6] 薛浩,英林海,王鹏,等. 云边协同的5G PaaS平台关键技术研究[J]. 电信科学,2019(22):89-97.
- [7] Vodafone Business. IoT Barometer Main Report[EB/OL]. [2020-01-12]. <https://business.vodafone.com/barometer2017>.
- [8] 吕华章,陈丹,王友祥. 边缘云平台架构与应用案例分析[J]. 邮电设计技术,2019(3):35-39.
- [9] 阳锐,刘娜,李俊珠,等. 泛在电力物联网大数据平台架构及应用探讨[J]. 邮电设计技术,2019(9):25-30.

作者简介:

张涛,毕业于北京邮电大学,工程师,硕士,主要从事运营商大数据分析及应用创新应用产品的研究工作;程新洲,教授级高级工程师,主要从事大数据研究与应用工作;徐乐西,毕业于伦敦大学,高级工程师,博士,主要从事大数据分析、LTE-A/5G移动通信研究及标准化工作;高洁,毕业于北京邮电大学,高级工程师,硕士,主要从事大数据分析技术及行业创新应用产品的研究工作;曹丽娟,毕业于北京邮电大学,工程师,硕士,主要从事大数据分析及构建大数据可视化方案等工作。