

# 大数据驱动5G网络与服务优化

邬贺铨

中国工程院, 北京 100088

## 摘要

5G时代的到来将进一步加速移动数据的发展。介绍了移动大数据的来源,分析了大数据分析在5G网络优化中的应用方向。针对5G网络的新特点,系统阐述了大数据在5G网络中的大规模天线与分布式天线、无线接入网资源管理、异构接入组网、云网、移动边缘计算、终端与云端的智能、SDN与NFV、网络切片、跨层联合优化、源选路优化等方面可能的应用。并结合5G大数据的特点,分析了其在智慧城市、智慧医疗、智慧交通、工业互联网方面的应用。

## 关键词

5G; 大数据; 移动数据; 网络优化; 服务优化

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.2096-0271.2018055

## *Big data driven 5G network and service optimization*

WU Hequan

Chinese Academy of Engineering, Beijing 100088, China

## *Abstract*

The advent of 5G era will further accelerate the development of mobile data. The source of mobile big data was introduced, and the application direction of big data analysis in 5G network optimization was analyzed. According to the new characteristics of 5G network, the potential applications of big data in 5G networks, such as large-scale antenna and distributed antenna, wireless access network resource management, heterogeneous access network, cloud network, mobile edge computing, terminal and cloud intelligence, SDN and NFV, network slicing, cross-layer joint optimization, source routing optimization, etc., were systematically expounded. Combined with the characteristics of 5G data, the applications of 5G big data in smart city, smart medicine, smart transportation, industrial Internet were analyzed.

## *Key words*

5G, big data, mobile data, network optimization, service optimization

## 1 引言

5G时代即将到来。和4G网络相比,5G网络在各个方面都有很大提升,不仅传输速率更高,而且在传输中还呈现出增强移动宽带、超可靠低时延、广覆盖大连接的特点。如果说1G~4G主要面向个人通信,那么5G则扩展到工业互联网和智慧城市应用。根据《Cisco VNI: global mobile data traffic forecast update, 2016-2021》的数据显示,2016—2021年全球移动数据流量增长7倍,平均年增47%,增长速度非常快。3G时代全球每个用户每个月的连接流量只有3 GB,4G时代这个数值已经达到了6 GB,5G时代将达到30 GB(是4G网络平均连接流量的4.7倍)。2017年我国每个用户每个月产生的移动数据流量为1.775 GB,是2016年的2.3倍,与全球平均水平相当。而且2018年上半年的数据已经远远超过了2017年的全年数据,也就是说,5G时代的到来将进一步加速移动数据的发展。按照《中华人民共和国国民经济和社会发展第十三个五年规划纲要》的要

求,5G网络将在2020年商用。

## 2 移动大数据的来源

移动大数据包括用户产生的数据和运营商产生的数据,其中用户产生的数据包括自媒体数据和富媒体数据,运营商产生的数据包括日志数据和基础网络数据。在运营商的网络上有很多环节可以进行数据采集,在终端可以采集路测(DT)/最小化路测(MDT)、测试报告(MR)、传输分组大小、使用习惯、终端类型等数据;在基站端可以获得用户的位置信息、用户通话记录(CDR)、链路状态信息(CSI)、接收信号强度(RSSI)等数据;通过后台的运维系统可以采集测量、信令、话务统计等数据;通过互联网可以采集新闻、资讯、地图、视频、聊天、应用等数据。也就是说,在运营商的网络中不但可以获得业务类型、上下行流量、访问网站等业务数据,还能掌握整个信道的状况。

如图1所示,5G网络应是以用户为中心、上下文感知与先应式的网络,且5G无

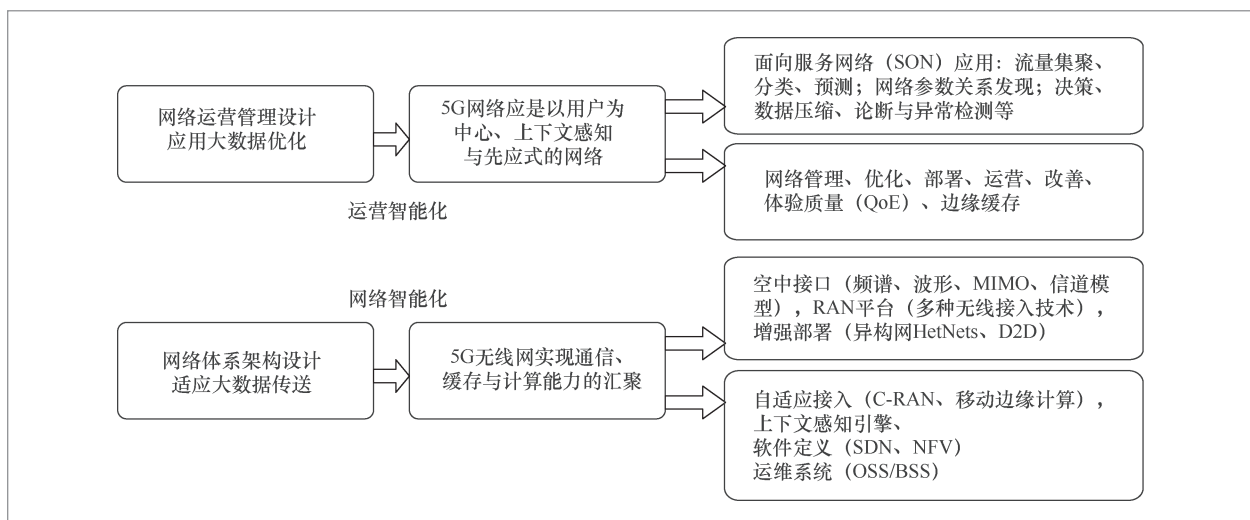


图1 大数据分析在5G网络优化中的应用

线网可实现通信、缓存与计算能力的汇聚,因此在网络运营管理设计时,需要利用大数据技术进行优化,在网络体系架构设计时要适应大数据的传送,以实现5G网络的运营智能化和网络智能化。

5G网络是数据终端到数据中心的主要通道,如图2所示。从互联网、物联网终端或移动用户处采集的数据,通过具有边缘缓存和计算能力的基站和无线接入网进行数据预处理与存储,最后通过核心网络将数据传输给数据中心和云计算中心进行数据分析<sup>[1-2]</sup>。5G网络除了传输数据终端的数据之外,智能终端的多功能业务还将触发终端与回传网络和核心网络内数百个服务器、路由器和交换机的各种交互。例如一个用户的HTTP请求可能只有1 KB,而内部数据流可能会增加930倍。因此,5G网络不仅要承载移动用户数据,还要承载来自不同后台、数据库、缓存服务器和网关以及回传链路的数据。

### 3 5G系统中大数据分析能力的设置

5G系统的数据分析应该在核心网大数据平台和基站端进行。在核心网大数据

平台应进行数据清洗、解析、格式化、统计分析、可视化等数据分析,按照内容预测算法执行计算并推断策略内容,然后主动地将决策指令存储在具有缓存能力的基站中,从而将决策行为从云传递到边缘(即基站)。而基站端负责收集上下文信息(如用户观看时长和位置信息),对用户空间-时间行为进行分析与预测,对数据进行汇总、压缩与加密,同时从核心网大数据平台获得决策指令。具有缓存能力的基站可以使大众内容靠近用户,改进用户体验并减轻回传网的负载。经过统计,当网络中有16个具有13 GB存储容量和30%的内容分级的基站时,采用主动缓存方式可以获得100%的用户满意度和98%的回传卸载<sup>[3]</sup>。缓存能力可以部署在无线接入网或核心网,或者两处均部署,缓存能力的分配也需要靠积累运行数据做出优化决策。

## 4 大数据支撑5G网络优化的方向

### 4.1 大数据支撑大规模天线与分布式天线

5G将使用大规模天线(MIMO),天线

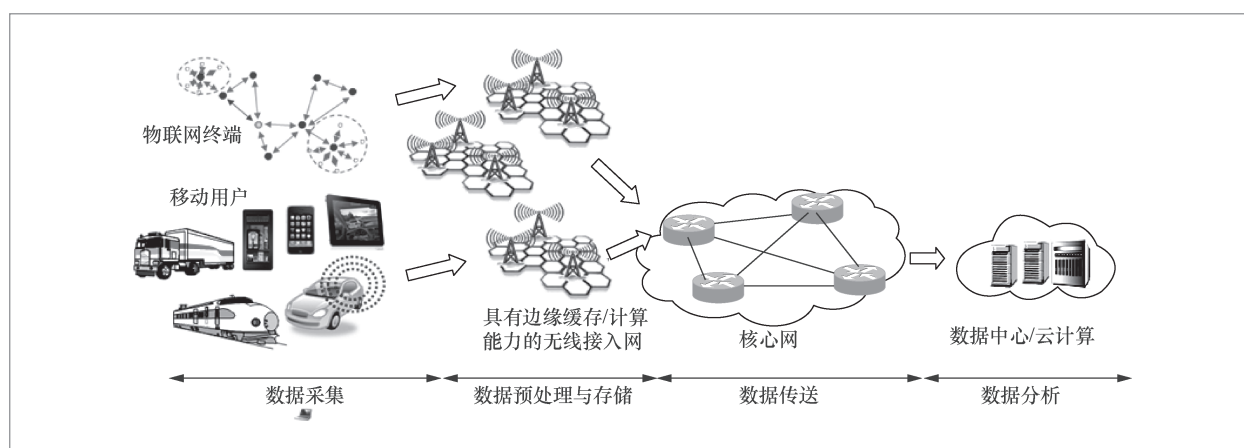


图2 5G网络是数据终端到数据中心的主要通道

数高达128个,甚至是256个。高阶MIMO为每条信道提供一条赋形的天线发射波束,实现空分复用,但各波束间存在干扰,降低了MIMO的效率,需要收集密集波束间的干扰数据,并基于系统的计算能力进行复杂的优化。此外,网络终端在基站中心接收的功率比较大,在基站边缘接收的信号比较差。此问题可以利用分布式天线解决,但是分布式天线互相之间也有干扰,如果能够收集到所有天线的信道数据和干扰数据,通过大数据分析技术对所有无线访问接入点(access point, AP)进行联合信号处理,就可以指导各天线和微基站实现对干扰的抵消,容量可较LTE系统提高约2个量级。此外,如果可以收集到MIMO数据和网络数据,并利用大数据技术进行分析决策,就可以提高定位精度。随着三维仿真、三维射线追踪技术的发展,通过室内天线和WLAN技术的结合,还能精确定位用户在室外或室内,甚至用户所在的具体楼层。

## 4.2 大数据支撑5G无线接入网资源管理

2G、3G时代的无线接入网是多层次的网络,在这种网络结构下,潮汐效应经常导致基站忙闲不均。因此4G系统将网络进行了扁平化设计,将基站分解为基带处理单元(BBU)和射频拉远模块(RRU),多个基站的BBU可集中为基带池,实现集约化资源利用。5G网络将BBU功能进一步分解为集中单元(CU)与分布单元(DU),CU可管理多个DU,实现干扰管理与业务聚合,DU实现多天线处理与前传压缩,灵活应对传输与业务需求变化,优化实时性能,降低硬件成本。这样的设计也可以更靠近用户,有利于集中化的管理。然而,一个CU管理多少个DU,需要基于大量用户空间-时间行为的大数据来优化设计,特别

是如何从能效的角度实现忙闲时不同的资源调配。

## 4.3 大数据分析支撑异构接入组网

由于5G网络的频段很高、带宽很大,若采用高功率的宏基站,则布设与运营成本高,但采用大量微基站,则干扰严重,且难以进行站点选址优化。以上问题可以通过以下几种方式解决。

- 宏微蜂窝混合组网。宏蜂窝负责广覆盖,支持高优先级业务;微蜂窝实现热点覆盖,面向低优先级高速业务。
- 控制面与数据面分离组网。大量微基站需要集中化管理,以防止干扰,将控制面信令数据与数据面用户数据分离,控制面信令数据接入宏蜂窝。这样,终端在微基站之间切换就不影响宏基站信令了,而且通过小区分簇化集中控制,可以解决小区间的干扰协调和负载均衡的问题。
- 上下行解耦异构组网。终端的MIMO数远少于基站,上行覆盖低于下行;在蜂窝边缘,可采用“5G下行+4G上行”的异构方式运行。

通过分析可以看出,传统移动网络的控制面、用户面、上下行数据链路都在同一个蜂窝小区内。而5G网络的控制面、用户面可以接入不同的基站,上下行可以接入不同的蜂窝,甚至分别在4G和5G系统,每个终端可能同时接入不止一个基站。因此,具体选择接入哪个基站和哪个系统,应该根据用户的分布数据和网络负载来决定,此时就要用到大数据的分析和决策方法。

## 4.4 大数据支撑5G云网

如图3所示,5G网络是一个云化的网络,包括接入云、转发云、控制云。接入云

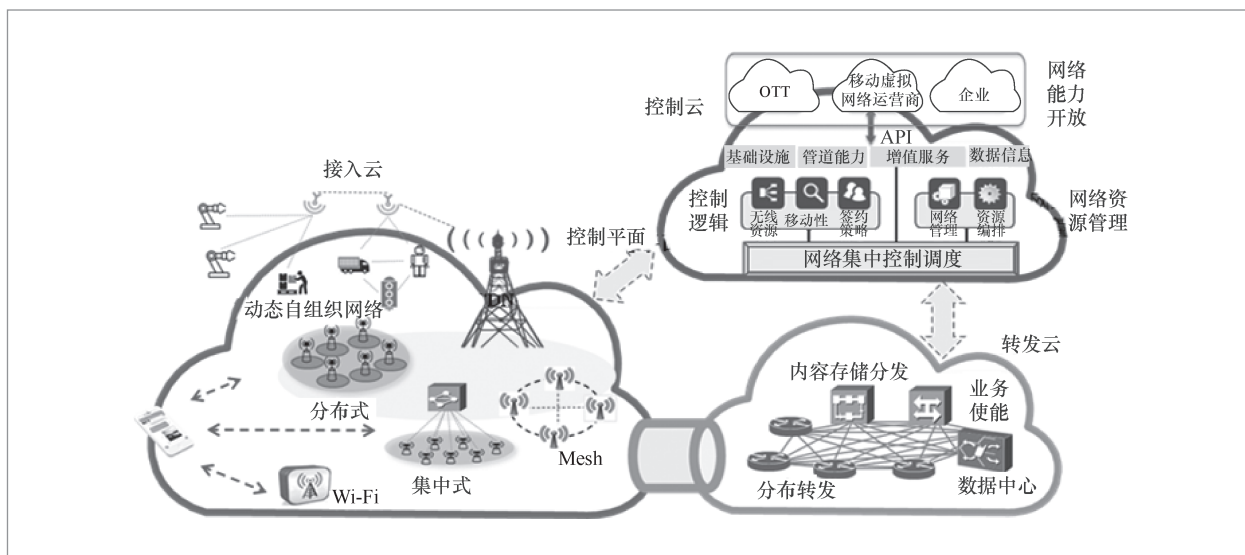


图3 云化5G网络

是指在微蜂窝超密集覆盖的场景下，一簇微基站组成虚拟小区，实现微基站之间的资源协同管理和干扰协调；转发云是指各业务流共享高速存储转发与防火墙及视频转码等各类业务使能单元；控制云包括网络资源管理、网络能力开放、控制逻辑等模块。此外，在5G的场景下还可以有移动云计算、移动边缘计算（mobile edge computing, MEC）、微云和飞云（femto-cloud）等多种云，它们可以被部署在无线网的不同位置，其配置需要借助网络 and 用户大数据分析来寻优。

#### 4.5 大数据优化5G移动边缘计算

为适应视频、虚拟现实/增强现实与车联网等业务的时延要求，减轻核心网带宽的压力，需将这些业务的存储和内容分发下沉到MEC处理。5G网络不仅可在边缘感知和分析数据，而且可在亚秒或毫秒内触发响应措施，所有的数据无缝地从云平台转到大量的端点或从大量的端点转到云平台。那么哪些业务需要放到云计算

中心处理，哪些业务需要下沉到MEC处理，这需要基于网络收集到的业务流数据进行分析。

#### 4.6 大数据支撑5G终端与云端的智能

现在智能终端的能力非常强，但是在终端上的人工智能处理能力还是有限的。比如手机智能可完成2D人脸识别，但识别效果容易受光线、角度和表情及化妆等的影响，而且识别的是照片还是真人也不好区分，现在的3D人脸识别就需要利用网络云端的智能来处理，从而提供安全的识别能力。云计算可强化无线网，有效支持诸如增强现实等计算强度的应用，将用户端很重的计算任务卸载到云端。然而，无论是终端还是云端的智能，都需要基于大数据的分析。如AI需要训练与推断，训练包括前向计算和后向更新（通过大数据调整模型参数），推断主要是前向计算，将训练得到的模型用于应用。通常云端负责训练和推断，终端只负责推断。因此仅靠终端的计算与软件能力的发展还不够，很多智能

应用也需要云端的支持,如云端训练和云端推理、云端训练和终端推理。

#### 4.7 大数据支撑软件定义网络与网络功能虚拟化

软件定义网络(SDN)全局优化路由的能力来自对全网流量流向、跨层网络资源大数据及业务流QoS需求大数据的掌控与分析,需要很强的计算能力支撑,以实现网络路由的快速收敛和稳定。网络功能虚拟化(NFV)功能的选择也基于网络大数据的分析。

#### 4.8 大数据支撑5G网络切片

5G很重要的功能是网络切片。5G需要支持不同的业务需求,如超宽带业务、低时延业务、大连接业务。若带宽不一样,那么对网络的性能要求也不一样,切片是网络转发资源的分割,不同切片间的业务相互隔离,切片的实现涉及转发面与控制面功能,每个切片上可以运行不同的L2/L3网络协议。为不同业务需求的用户组织不同的切片,需要利用深度分组检测(DPI)数据建立预测模型,精准预测热点数据请求。网络资源在切片间分配的联合优化,也需要利用网络资源大数据进行学习和分析。

#### 4.9 基于大数据实现5G跨层联合优化

5G网络中,IP层的选路适用于细颗粒的业务流,但时延大;MAC层的交换适用于大颗粒的业务流,但时延也较大;灵活以太网中继的业务流颗粒较大,但优点是时延低。对于每一种业务流来说,选择在每一层做交换或路由是一个跨层联合优化问题,可以借助网络大数据进行优化。

#### 4.10 借助网络大数据优化5G源选路

切片分组网(SPN)基于切片以太网和分段选路(SR)技术,用于中传和回传。

传统的IP网络按无连接方式工作,对具有相同源地址和相同目的地址的同属一次通信的各IP分组进行独立处理,不考虑它们前后的关联,同属一次通信的各IP分组在沿途各节点均独立选路,甚至会走不同的路由,这是在互联网之初的网络可靠性不高的情况下,以时延和效率为代价换取灵活性和生存性。现在的网络性能已有很大的改进,如果按照每次通信中首个分组的特征来配置数据平面的设备(即配置流表),那么该次通信的后续数据分组被抽象为同一流,同一次通信的后续各IP分组无需再选路。由于在源节点已设置了有序的指令集,标识了沿途经过的节点或链路,这些节点无需感知业务状态,只需维护拓扑信息,简单地按配置流表执行转发功能,这就相当于面向连接分组的通信,显著提升了网络效率。因此,分段选路又称源选路,它无需LDP/RSVP-TE等信令协议,适合接受SDN的控制。源选路指令集的设计需要借助网络大数据来优化。

#### 4.11 大数据支持5G核心网基于服务的体系

在基于服务的网络体系(SBA)方面,网络功能在4G是网元的组合,在5G是通过API交互的业务功能的组合,业务被定义为自包含、可再用和独立管理。业务的解耦便于快速部署和维护网络;轻型的接口便于引入新特性;模块化为网络切片提供灵活性;使用HTTP的API代替Diameter作为公共控制协议,更容易调用网络服务。

然而,针对每一次会晤所调用的服务是否最优、网络资源分配是否公平合理、同时进行的多个会晤所用的服务有无冲突、以API方式新增加的服务类型与功能是否与网络能力兼容这些问题,不能仅依靠运维人员的经验,还需要利用网络大数据来支撑。

#### 4.12 大数据对视频业务传输体验的保障

视频是5G的主要应用场景,也是运营商的每用户平均收入(average revenue per user, ARPU)的重要支撑。但视频的传输质量受到以下因素的影响:

- 回传路径太长,时延超标;
- 因无线空口信道干扰或系统负载忙,使每一移动终端可获得的带宽受限,导致视频信号传输控制协议(TCP)端到端时延过长,吞吐率下降;
- 将OTT视频仅作为一般互联网业务对待,没有服务质量(QoS)保障;
- 对于视频会议业务,可能因上行分组数据汇聚协议(packet data convergence protocol, PDCP)丢失而停止视频接收。

因此,需要考虑基于业务流的大数据,

对无线接入网(RAN)和App进行相互感知,例如网络协助基于HTTP的动态自适应流(dynamic adaptive streaming over HTTP, DASH)、视频感知调度等,实现对基于Web的视频流的深度跨层优化和本地内容缓存。目前,3GPP已经开始研究改进上下文感知引擎,允许空口和核心网将流数据的指示传到5G的控制面,实现对单个用户或整个网络的流管理。

## 5 5G大数据在其他领域中的应用

前文介绍的是大数据在5G网络中的应用,然而结合了大数据的5G网络更大的应用领域应该在社会生活中,如智慧城市、智慧医疗、智慧交通、工业互联网等。5G大数据可以提供上下文感知、预测、规律发现、认知等功能,分别实现实时修正、下一步最佳措施、下一步自动应对、最好的决策的目的,见表1。

## 6 结束语

5G时代的到来正在加快无线大数据的增长。大数据在社会与产业各领域都将

表1 5G大数据在其他领域中的应用

功能	目的	网络	智慧城市	智慧医疗	智慧交通	工业互联网
上下文感知	实时修正	上下文感知网与应用服务	电子商务; 宽带化; 城市交通	可穿戴设备; 数字病历	驾驶员行为; 在线诊断; 服务提示; 胎压检测	资产管理
预测	下一步最佳措施	自学习网与应用服务	流量管理; 废品管理; 金融/保险	先进诊断; 卫生预警	保险溢价评估, 预测 全方位距离导航系统提示	预防性维护; 产出预测; 收入预测
规律	下一步自动应对	自愈网与应用服务	公共安全; 家庭安全; 可生存性	预防保健	突破Fix; 智能路径; 里程管理	能效管理; 供应链优化; 工作岗位调度
认知	最好的决策(似人的直觉)	自驱动网与应用服务	城市设施规划; 应急响应系统; 城市维修	远程伤害照护	自动/无人驾驶	控制与自动化

有广泛的应用,并产生重要影响。大数据对5G网络的发展(如网络体系架构的设计、运维的提效和服务体验的提升等)将起到强化和优化的作用。大数据在5G网络的应用有很广阔的空间,同时也面临不少挑战,需要解决数据挖掘计算复杂度、时效性、能效、安全性等问题,同时也给5G网络标准化和实现提出了很多创新课题。

### 参考文献:

[1] ZHANG N, YANG P, REN J, et al.

Synergy of big data and 5G wireless networks: opportunities, approaches, and challenges[J]. IEEE Wireless Communications, 2018, 25(1): 12-18.

[2] FARRINGTON N, ANDREYEV A. Facebook's data center network architecture[C]// IEEE Optical Interconnects Conference, May 5-8, 2013, Santa Fe, USA. Piscataway: IEEE Press, 2013.

[3] ZEYDAN E, BASTUG E, BENNIS M, et al. Big data caching for networking: moving from cloud to edge[J]. IEEE Communications Magazine, 2016, 54(9): 36-42.

#### 作者简介



**邬贺铨**(1943-),男,中国工程院院士,光纤传送网与宽带信息网专家。先后从事光纤传输系统和宽带网研发、中国下一代互联网(CNGI)和3G/4G/5G等技术管理及工程科技咨询项目研究。曾任电信科学技术研究院副院长兼总工程师、中国工程院副院长。现任国家信息化专家咨询委员会副主任、中国标准化专家委员会主任、国家“互联网+”行动专家咨询委员会主任、中国互联网协会理事长、国家“新一代宽带无线移动通信网”国家科技重大专项总师、中国下一代互联网示范工程(CNGI)专家委员会主任、IEEE高级会员。

收稿日期:2018-10-26

\* 本文为2018中国国际大数据大会演讲约稿