



信息通信领域产学研合作特色期刊 | 十佳皖刊  
第三届全国期刊奖百种重点期刊 | 中国科技核心期刊

ISSN 1009-6868  
CN 34-1228/TN  
CODEN ZTJHAY

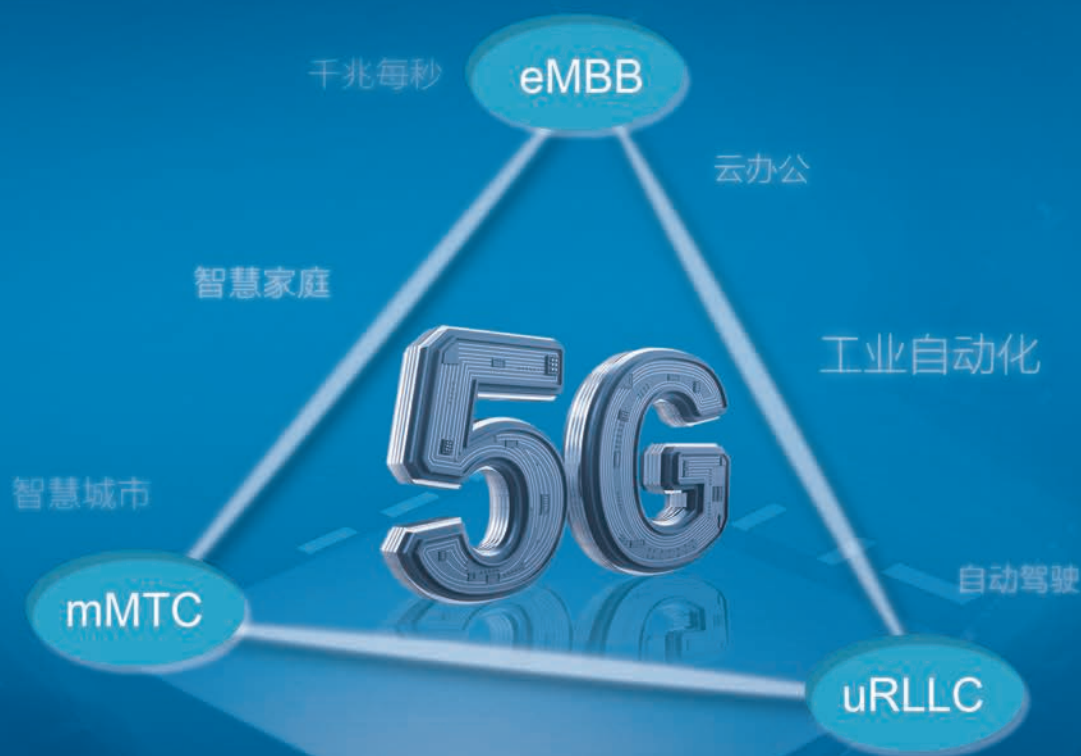
# 中兴通讯技术

## ZTE TECHNOLOGY JOURNAL

<http://tech.zte.com.cn>

2019年2月 • 第1期

专题：5G 商用支撑理论及关键技术



# 《中兴通讯技术》第8届编辑委员会成员名单

- 顾问

侯为贵（中兴通讯股份有限公司创始人） | 钟义信（北京邮电大学教授） | 陈锡生（南京邮电大学教授）
- 主任

陆建华（中国科学院院士）
- 副主任

徐子阳（中兴通讯股份有限公司总裁） | 糜正琨（南京邮电大学教授）

编委（按姓名拼音排序）

- 陈建平

上海交通大学教授
- 陈前斌

重庆邮电大学副校长
- 葛建华

西安电子科技大学教授
- 管海兵

上海交通大学教授
- 郭庆

哈尔滨工业大学教授
- 洪波

中兴发展股份有限公司总裁
- 洪伟

东南大学教授
- 纪越峰

北京邮电大学教授
- 蒋林涛

中国信息通信研究院科技委主任
- 李尔平

浙江大学教授
- 李红滨

北京大学教授
- 李建东

西安电子科技大学副校长
- 李军

清华大学教授
- 李乐民

中国工程院院士
- 李融林

华南理工大学教授
- 李少谦

电子科技大学教授
- 林晓东

中兴通讯股份有限公司副总裁
- 刘健

中兴通讯股份有限公司高级副总裁
- 刘建伟

北京航空航天大学教授
- 陆建华

中国科学院院士
- 马建国

广东工业大学教授
- 孟洛明

北京邮电大学教授
- 糜正琨

南京邮电大学教授
- 任品毅

西安交通大学教授
- 孙知信

南京邮电大学教授
- 谈振辉

北京交通大学教授
- 唐雄燕

中国联通网络技术研究院首席科学家
- 王文博

北京邮电大学副校长
- 王文东

北京邮电大学教授
- 王喜瑜

中兴通讯股份有限公司执行副总裁
- 王翔

中兴通讯股份有限公司高级副总裁
- 卫国

中国科学技术大学教授
- 吴春明

浙江大学教授
- 邬贺铨

中国工程院院士
- 徐安士

北京大学教授
- 徐子阳

中兴通讯股份有限公司总裁
- 续合元

中国信息通信研究院副总工
- 薛一波

清华大学教授
- 杨义先

北京邮电大学教授
- 杨震

南京邮电大学校长
- 易芝玲

中国移动研究院首席科学家
- 张宏科

北京交通大学教授
- 张平

北京邮电大学教授
- 张云勇

中国联通研究院院长
- 赵慧玲

工业和信息化部科技委信息网络专家组组长
- 郑纬民

清华大学教授
- 钟章队

北京交通大学教授
- 周亮

南京邮电大学教授
- 朱近康

中国科学技术大学教授
- 祝宁华

中国科学院半导体研究所副所长

# 目次

中兴通讯技术 (ZHONGXING TONGXUN JISHU)  
总第144期 第25卷 第1期 2019年2月

## 卷首特稿

对网络体系变革的思考 **02**  
邬贺铨

## 专题:5G商用支撑理论及关键技术

降低SCL译码错误的级联极化码 **05**  
王涛,屈代明,江涛

高频通信技术 **12**  
李萍,魏浩,黄静月

“太极混一”——极化码原理及5G应用 **19**  
牛凯

未来无线网络下的空中计算技术 **29**  
陈力,卫国

免调度非正交多址接入技术及其系统性能 **35**  
李卫敏,唐红,李剑,胡宇洲

**42** 智慧标识网络动态防御机制与应用  
于成晓,刘刚,张宏科

**47** 上行-下行联合优化的uRLLC传输  
成晶,沈超,夏树强

## 专家论坛

**55** B5G:泛在融合信息网络  
张平

**63** 基于非完备大数据的业务预测  
李建东,盛敏,文娟

## 企业视界

**67** 5G在垂直行业中的应用  
陆平,李建华,赵维铎

## 2019年第1—6期专题计划及策划人

### 1. 5G商用支撑理论及关键技术

中兴通讯股份有限公司执行副总裁 王喜瑜  
中兴通讯股份有限公司首席科学家 向际鹰

### 3. 边缘计算技术及其应用

清华大学教授 郑纬民  
佐治亚州立大学教授 潘毅  
韦恩州立大学教授 施魏松

### 5. 新型光互连与光接入技术

北京大学教授 李红滨

### 2. 云网一体化技术

中国联通网络技术研究院首席专家 唐雄燕

### 4. 5G通信安全技术

清华大学教授 李军

### 6. 5G通信系统示范应用

中国信息通信研究院科技委主任 蒋林涛

# CONTENTS

ZTE TECHNOLOGY JOURNAL Vol. 25 No. 1 Feb. 2019

## Guest Paper

- Reflections on the Reform of Network Architecture **02**  
WU Hequan

## Special Topic: Supporting Theory and Key Technologies of 5G Commercialization

- Concatenated Polar Codes for SCL Decoding Error **05**  
Reduction  
WANG Tao, QU Daiming, JIANG Tao

- High Frequency Communication Technology **12**  
LI Ping, WEI Hao, HUANG Jingyue

- The Principle and Application of Polar Codes in **19**  
5G Systems  
NIU Kai

- Over-the-Air Computation for Future Networks **29**  
CHEN Li, WEI Guo

- The Grant-Free NOMA Technology and Its System **35**  
Level Performance  
LI Weimin, TANG Hong, LI Jian, HU Yuzhou

- 42** Dynamic Defense Mechanism and Application for  
Smart Identifier Networks  
YU Chengxiao, LIU Gang, ZHANG Hongke

- 47** Joint Uplink-Downlink Transmission Design for  
uRLLC  
CHENG Jing, SHEN Chao, XIA Shuqiang

## Expert Forum

- 55** B5G: Ubiquitous Fusion Information Networks  
ZHANG Ping

- 63** Traffic Prediction with Incomplete Big Data  
LI Jiandong, SHENG Min, WEN Juan

## Enterprise View

- 67** Applications of 5G in Vertical Industry  
LU Ping, LI Jianhua, ZHAO Weiduo

期刊基本参数: CN 34-1228/TN\*1995\*b\*16\*74\*zh\*P\* ¥ 20.00\*15000\*11\*2019-02

## 敬告读者

本刊享有所发表文章的版权,包括英文版、电子版、网络版和优先数字出版版权,所支付的稿酬已经包含上述各版本的费用。未经本刊许可,不得以任何形式全文转载本刊内容;如部分引用本刊内容,须注明该内容出自本刊。

# 专题:5G 商用支撑理论及关键技术

专题策划人 王喜瑜



中兴通讯股份有限公司执行副总裁,教授级高工;1998年入职中兴通讯,先后主管CDMA设计开发部、BBU研发中心、无线研究院、公司技术规划部,现负责中兴通讯系统产品研发业务领域;曾获国家科学技术进步奖二等奖、广东省科学技术进步一等奖、中国通讯协会科技进步一等奖等。

专题策划人 向际鹰



中兴通讯股份有限公司首席科学家;先后从事CDMA基站设计、WCDMA、WIMAX、LTE基站系统、全IP核心网和控制器等设计工作;获“2013年度中国通信产业技术贡献人物奖”、国家技术发明奖、国家科技进步奖、广东科技进步奖、深圳市科技进步奖、南粤科技创新奖,以及电子部科技创新奖等多项奖励;授权发明专利30余项,国际PCT专利约20项。

## 内容导读

随着5G第1阶段标准化工作的逐步完成,业界普遍认为5G于2019—2020年开始商用,因此在这个时间节点上梳理一些5G商用的支撑理论及关键技术显得很有必要。

目前,5G已经定义了3大类应用场景,即增强型移动带宽(eMBB)、高可靠低时延通信(uRLLC)及大规模机器通信(mMTC)。这3大类应用场景基本覆盖了目前生活、生产所能碰到的大部分场景。

5G的上述特性会影响到其系统设计的方方面面,如对eMBB的支持就需要考虑不同频段的不同特征:在低频段,需要充分考虑频谱的效率,除了对链路自身需有较好的自适应及链路之间的相互协调配合外,还需要有效利用空间资源;在高频段,波束是整个系统设计的关键,所有静态的信号和信道以及动态的过程都发生在波束中,因此波束管理在高频通信中至关重要。

5G虽然可以通过不同途径的灵活配置参数来支持以上3大类场景,但并不意味着只是技术的简单堆砌,而是在设计伊始时就考虑到前向兼容。前向兼容一方面体现在同一层面的设计保持在同一框架下,最典型的一个例子就是参数集和帧结构的设计,通过半静态或动态配置极其有限的几个参数,灵活实现对不同场景较优化的支持;另一方面体现在其“瘦”的设计,或称为“自包含”设计,即5G大量削减了周期性、小区级的广播信号,把所有相关信号信道控制在所配置的范围,这样既能通过

各种手段进行合理配置,又能做到不影响其他功能的实现。

5G基本实现了当初设定的目标,但也处于不断的发展和优化中,尤其在uRLLC和mMTC这2个场景的设计和优化方面,还有很多工作要做,如对uRLLC场景的支持,需要有更好的理论来说明各种结构化利用资源的方式,更加深入地理解延迟和可靠性之间的关系;对于中短包的编译码和链路自适应问题,需要继续深入研究极化码/低密度校验码(LDPC)在5G中已有的应用方式,并思考它们的扩展使用方式。在mMTC场景中,面对数量多、功耗低、价格低廉的节点,需要思考传统语音/数据的正交接入方式是否合适,是否存在更简单、高效的非正交接入方式;面对5G对各垂直行业的渗透,需要研究在接入网侧支持对各种业务的感知和汇合功能,以达到诸如“智慧城市”的远期目标等。

本期专题针对上述5G在商用过程碰到的理论问题及关键技术进行了初步探讨,并给出了一些初步的解决方案。论文主要来自高校和中兴通讯等相关团队的最新研究成果,相信这些成果对推动5G的商用以及5G未来的发展都具有积极的借鉴意义,也能给读者提供有益的启示和参考。在此,对各位作者的大力支持表示衷心感谢。

王喜瑜、向际鹰

2019年1月20日





# 对网络体系变革的思考

## Reflections on the Reform of Network Architecture

邬贺铨/WU Hequan

(中国工程院, 北京 100088)  
(Chinese Academy of Engineering, Beijing 100088, China)

**摘要:** 互联网已走过了 50 年, 其发展远远超越了初衷, 各类新业务与应用加大了网络体系变革的压力, 对网络体系颠覆性的探索和演进创新一直在进行。近年来关于 5G 网络体系的演进方向逐渐明朗, 从传送模式创新、虚拟化和云化 3 个方面, 分析了网络体系变革给运营商带来的机遇与挑战, 指出变革将开拓网络技术与业务应用创新的空间。

**关键词:** 互联网; 网络体系; 5G 传送模式; 虚拟化; 云化

**Abstract:** The Internet has been developed for 50 years, and its development is far beyond the original intention. All kinds of new business and applications have increased the pressure of network system reform, while the disruptive and evolution innovation of the network architecture has been constantly carried on. In recent years, the evolution direction of 5G network architecture is gradually clear. In this paper, the opportunities and challenges to operators are analyzed from innovation of the transfer mode, the virtualization and the cloudization. It is pointed that transformation will bring the innovation of network technology and business application.

**Key words:** Internet; network architecture; 5G transfer mode; virtualization; cloudization

DOI: 10.12142/ZTETJ.201901001

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190131.1438.008.html>

收稿日期: 2018-12-28 网络出版日期: 2019-01-31



邬贺铨

中国工程院院士, 曾任中国工程院副院长, 现任国家信息化专家咨询委员会副主任、国家标准化专家委员会主任、国家互联网+专家委员会主任、国家物联网专家组组长、国家新一代宽带无线移动通信网重大科技专项总师、中国互联网协会理事长, 以及 IEEE 高级会员等; 长期从事数字和光纤通信系统的研究开发工作, 近十多年来负责中国下一代互联网示范工程和 3G/4G/5G 等研发项目的技术管理及重要工程科技咨询项目研究; 曾获全国科学大会奖、国家科技进步二等奖、邮电科技进步一等奖等多个奖项; 出版专著 1 部。

从 1969 年美国 APARNET 算起到 2019 年, 互联网走过了 50 年, 全球互联网普及率超过 55%。中国全面接入互联网也有 25 年, 普及率高于全球平均水平<sup>[1]</sup>。互联网设计之初主要考虑的是抗毁性和异构设备互联及效率; 但现在在互联网的应用已经超越初衷, 从学术研究到商业应用, 从数据到语音、视频, 从固定业务到移动业务, 从面向人的通信到物联网, 现在又从消费互联网进入到产业互联网。以支持固定终端尽力而为数据业务为出发点设计的互联网体

系, 其承受能力超出了我们的想象; 但面对产业互联网对高可靠和低时延性及高安全性的要求, 现有互联网体系还能够继续适应应用的扩展吗? 关于下一代互联网的研究从 20 年前就开始了, 现在以国际电信联盟 (ITU) 为代表的机构又开启了面向大数据、人工智能、云计算和 5G 的未来网络研究<sup>[2]</sup>。

关于互联网体系的研究一直以来有 2 条路线, 即演进路线与革命路线。前者针对新出现的问题以修补或新增协议的方式来应对, 但会增加网络的复杂性而难于管理, 或带来过多的开销而失去效率。革命路

线以颠覆现有互联网体系为目标,例如:2005年美国的全球网络环境创新(GENI)计划本希望创造新的网络和分布系统体系;但其追求过于完美,试图解决互联网出现的所有问题,至今已过去了十多年,却与很多类似的研究项目一样尚未能取得实质性进展。GENI现在将关注点回到软件定义网络(SDN)和网络功能虚拟化(NFV)<sup>[3]</sup>。因为对互联网流量流向的研究未能得出接近实际的数学模型,仿真还不能反映实际应用环境,产品性试验床有实际用户但变化不够,研究性试验床有较多的变化但缺少实际用户;因此,与现有互联网规模和应用多样性及随机性相比,试验网难当此任。

面对互联网体系研究的两难境地,我们有必要回顾互联网的发展历史。在互联网概念提出之前有X.25面向连接的分组网,但只考虑支持数据传输。在互联网出现之后的80年代初,异步传输模式(ATM)被提出,它使用的是面向连接的分组通信,固定包长,可支持语音、视频与数据。ATM很理想,但在当时看相当复杂,输给了不理想但简单的IP。IP设计之初仅考虑支持无连接的数据通信,这是原有通信网络的弱项,而且在当时的网络中数据流量占比很少。另外,IP开放的架构可在数字化的任何网络上(包括电路交换的程控电话网)运行,对底层通信设施没有特别的要求,生存空间广

阔,而且互联网开始规模不大,也没有商业化的压力,从单一应用切入站住了脚。互联网于90年代才开始探索支持基于IP的语音传输(VoIP),逐渐渗透到电信网的所有业务,成为传统电信网技术的成功颠覆者。现在互联网规模大、应用广,要想通过革命性方式解决现有互联网的问题难度太大,难以验证所提出的全盘改造互联网的方案;因此探讨从某一点改进互联网也许是合适的切入点。

5G应用场景所要求的超宽带、大连接、低时延和高可靠对现有网络体系带来了新的挑战,加速了对网络体系研究的进程。按照思科的报告<sup>[4]</sup>,2018年移动IP流量占全网IP流量之比为11.03%。从流量看,移动网占全网比例不算高,移动网体系的改造对全网影响不大。从这一角度看,5G适合于作为切入点探讨互联网体系的变革或演进。

5G在网络体系方面<sup>[5-6]</sup>的一个特点是转发模式的创新。在网络层引入分段选路(又称为源选路)的概念,源节点按照每次通信首个IP包的特征来配置流表,用于指示报文的转发路径;该次通信的后续IP包被抽象为同一流,中间的节点无需再感知业务状态,只需维护拓扑信息并按流表执行转发功能,相当于面向连接分组通信。这与传统的IP网络按无连接方式工作不同。在无连接模式下,对具有相同源地址和相同目的地地址的

同属一次通信的各IP包独立处理而不考虑它们前后的关联,而且各IP包在沿途各节点均独立选路。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。5G的新转发模式还表现在以64 B/66 B码块为单位的以太网中继,可代替小颗粒性的IP包作为单元来选路,即以1.5层的交叉连接代替第3层的选路,转发效率更高,时延更短,特别适合支持宽带视频业务。以太网技术还被应用到5G的前传网络。前传如继续使用时分复用同步数字体系(SDH)技术,其传送带宽是固定的,不会随实际业务负荷而变,而且大规模多输入多输出(MIMO)下的带宽太高而难以承受,基于以太网的统计复用可动态适应业务流量并降低带宽压力。为了克服传统以太网的不确定性,利用以太网报头虚拟局域网(VLAN)中的服务类型指示来区分高优先级流量和后台流量,为高优先级的帧提供抢占功能,这被称为时延敏感网络(TSN)。

5G网络体系的另一个特点是虚拟化。基于SDN实现数据面与控制面分离,基于NFV从黑盒设备中解耦出网络功能软件,支撑了数据面与用户面的分离,并实现了硬件的白盒化。5G还将采用基于服务架构(SBA),借鉴IT领域的微服务理念,将软件分解为细粒度的模块化组件。每个微服务可独立地进行

部署、升级和扩展,这相当于传统电话网中智能网的思路;但不同的是5G通过开放应用程序编程接口(API)来实现微服务的集成,以提升应用开发的整体敏捷性和弹性。在上述基础上5G进一步发展到网络业务切片。5G既需要支持峰值高达20 Gbit/s的视频也要应对大量千比特每秒级的物联网连接,5G既要支持高铁类高速移动业务也有大量固定或低速移动业务,各类业务对可靠性和时延的要求有很大差异。在不改变物理基础设施的情况下,针对用户的不同业务需求通过网络控制面的协同编排,动态地将相应的网络资源组成切片,即为每一种业务建立逻辑上的VPN,各切片间业务隔离。5G的网络虚拟化还表现为异构接入组网,在密集微蜂窝的情况下,为了避免用户信令的频繁切换,信令接入到宏蜂窝而用户数据接入微蜂窝,即控制面与用户面分离。在5G的情况下,则取决于信道性能,用户的上下行可能不在同一个微蜂窝,甚至下行在5G而上行在4G,将过去的以小区为中心发展到以用户为中心。

5G网络体系特点之三是云化。5G的核心网有云平台,实现数据中心的互联和大数据分析,支持无线资源管理、能力开放和网络资源编排等与控制面有关功能,以及支持各业务流共享高速存储与防火墙及视频转码等业务转发功能。5G特别

重视移动边缘计算(MEC),可以部署在基站的集中单元(CU),为多个与微蜂窝对应的分布单元(DU)所共享。MEC配备缓存与计算能力,负责收集与分析用户空时行为的上下文信息,实现对数据整合、压缩与加密。具有缓存能力的基站使大众内容或需要就近处理的工作尽可能靠近用户,改进用户体验并减轻回传网的负载,支撑车联网和虚拟现实(VR)/增强现实(AR)等高可靠低时延业务应用。

上述5G网络体系的理念可概括为通信与计算的融合、硬件与软件的解耦、云管端功能的协同,它们源于业务发展的需要也得益于IT特别是云技术的进步。5G网络虽然目前还是基于IP体系但又不局限于IP体系,可以应用到其他类型的第三层协议。这些理念也可以用在信息中心网(ICN)、数据感知网(DAN)或内容中心网(CCN)。这些网络体系的变革并非对互联网现有问题的全面改进;但部分演进也不会是一帆风顺的,任何网络体系的变革都会涉及到产业链上下游各方利益的博弈,对运营商来说也是艰难的抉择。总体上,运营商对上述演进方向寄予很大的期望,希望通过白盒化摆脱对设备供应商的依赖,希望以服务化架构(SBA)提升业务生成能力,不被网络内容服务商(ICP)所隔离,希望通过边缘计算避免管道化。不过很多技术都是双刃剑,白盒化使运

营商不依赖电信设备供应商但可能倒向IT企业,或更方便ICP自建网络设施;SBA有利于运营商开放业务能力,但业务外包有使运营商坐实管道化的风险;边缘计算有可能为非运营商的接入网企业所建设和管理,增加了争夺“最后一公里”的机会,把运营商推向产业链的更加边缘化位置。总之,网络业务发展至今,网络体系不变是没有出路的,变革又将面临技术与管理流程再造的挑战,但变革将开拓网络技术与业务应用创新的空间。

上述对网络体系变革的思考虽然是以5G为对象考虑的,但对固网来说也有重要的意义。在互联网50年的新起点和5G移动通信元年之初,面对网络体系变革的机遇与挑战,我们对网络体系的变革创新充满期待!

#### 参考文献

- [1] Internet World Stats. Internet Users in the World by Regions—June 30, 2018 [EB/OL]. (2018-06-30)[2018-12-18]. <https://www.internetworldstats.com/stats.htm>
- [2] ITU. SG13: Future Networks, with Focus on IMT-2020, Cloud Computing and Trusted Network Infrastructures [EB/OL]. [2018-12-18]. <https://www.itu.int/en/ITU-T/studygroups/2017-2020/13/Pages/default.aspx>
- [3] GENI (Global Environment for Network Innovations). About GENI [EB/OL]. [2018-12-18]. <https://www.geni.net/#>
- [4] Cisco Global Cloud Index: Forecast and Methodology, 2016–2021, White Paper [EB/OL]. (2018-11-19)[2018-12-18]. <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html>
- [5] 3GPP. 5G Phase 1 Underway in CT Groups, TR 24.890 [EB/OL]. (2017-06-13)[2018-12-18]. [http://www.3gpp.org/news-events/3gpp-news/1873-rel\\_14](http://www.3gpp.org/news-events/3gpp-news/1873-rel_14)
- [6] WEISSBERGER A. ITU-T SG15 Technical Report Transport network Support of IMT-2020/5G Network [R/OL]. (2018-11-05)[2018-12-18]. <http://techblog.comsoc.org/2018/11/05/itu-t-sg15-g-mtn-for-packet-slicing-network-transport-network-support-of-5g/>



# 降低 SCL 译码错误的级联极化码

## Concatenated Polar Codes for SCL Decoding Error Reduction

**摘要:** 串行抵消列表 (SCL) 算法是极化码的一种近似最大似然 (ML) 译码算法, 基于该算法的循环冗余校验 (CRC) 级联极化码、校验 (PCC) 级联极化码纠错性能优良, 已成为 5G 极化码标准编码方案。总结了 SCL 译码错误类型, 并从降低 SCL 译码错误的角度揭示了 CRC 级联极化码、PCC 级联极化码, 以及 CRC 辅助的 PCC 级联极化码, 三者提升 SCL 译码性能的原理。仿真结果表明: CRC 辅助的校验级联极化码可以显著降低 SCL 译码错误, 并在较高信噪比 (SNR) 范围内, 呈现出最佳的纠错性能。

**关键词:** 极化码; SCL 译码; 奇偶校验; CRC; 级联码

**Abstract:** Successive cancellation list (SCL) decoder with the proper list size works nearly as a maximum likelihood (ML) decoder for polar codes. With the modified versions of the SCL decoder, cyclic redundancy check (CRC)-concatenated polar codes (CRC-polar) and parity-check-concatenated (PCC) polar codes (PCC-polar) show the excellent error correction performance, and have been adopted as the standardized coding schemes in 5G technical specification. In this paper, the categories of SCL decoding errors are summarized, and the performance gain of the CRC-polar, PCC-polar and CRC-PCC polar are explained from the perspective of SCL decoding error reduction. The simulation results show that the CRC-PCC polar code could efficiently reduce the SCL decoding errors, and achieve the best error performance among the three concatenation schemes, in higher signal noise ratio (SNR) region.

**Key words:** polar codes; SCL decoding; parity-check; CRC; concatenated codes

随着移动通信技术的日趋进步和智能终端设备的高速发展, 为了满足日益增长的各类移动业务需求, 第 5 代移动通信技术 (5G) 逐渐成为学术界和信息产业界的研究热点<sup>[1-6]</sup>。在 2015 年, 国际电信联盟无线通信部 (ITU-R) 明确了 5G 3 大典型应用场景: 增强移动宽带

(eMBB)、大规模机器通信 (mMTC) 和高可靠低时延通信 (uRLLC), 并定义了各个业务场景的峰值速率和系统容量等关键技术指标<sup>[6]</sup>。为了满足业务场景的性能需求, 5G 信道编码标准技术经过多方研究和论证后率先于 2016 年被确定: 极化码和低密度奇偶校验码分别被选



王涛/WANG Tao  
屈代明/QU Daiming  
江涛/JIANG Tao

(华中科技大学, 湖北 武汉 430074)  
(Huazhong University of Science and Technology, Wuhan 430074, China)

DOI: 10.12142/ZTETJ.201901002  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190131.1032.006.html>

收稿日期: 2018-12-20  
网络出版日期: 2019-01-31

基金项目: 国家自然科学基金(61571200)、  
华中科技大学创新研究院技术创新基金  
(2018JYCXJJ032)

为 5G eMBB 场景控制信道和数据信道编码的标准技术方案<sup>[7]</sup>。

极化码于 2009 年由土耳其学者 E. Arıkan 教授提出,是第一类被证明容量可达的结构化信道编码方案<sup>[8]</sup>,并且因具有较低的编、译码复杂度等优势,受到广泛的关注和研究<sup>[9-18]</sup>。在 2011 年, I. Tal 等学者提出的串行抵消列表(SCL)译码算法是极化码的一种近似最大似然(ML)译码算法<sup>[9]</sup>,并且基于该算法的循环冗余校验(CRC)级联极化码展现出超越低密度校验码(LDPC)的纠错性能<sup>[9]</sup>。此后,文献[10]提出的校验(PCC)级联极化码引入校验比特,在译码过程中实时校验译码路径,性能可超越 CRC 级联极化码。这 2 类级联方案由于编、译码复杂度较低且纠错性能优良,成为 5G 极化码标准实现方案<sup>[7]</sup>。鉴于 CRC 级联极化码和 PCC 级联极化码的译码均基于 SCL 译码算法,本文中我们着重分析了 SCL 译码算法的错误类型,并从降低 SCL 译码错误的角度解释了 CRC 级联极化码、PCC 级联极化码,以及二者组合的 CRC 辅助的校验级联极化码提升纠错性能的原理。仿真结果表明: CRC 辅助的校验级联极化码可以显著降低 SCL 译码错误,尤其在较高信噪比(SNR)范围内,呈现出最佳的纠错性能。

## 1 极化码编码与 SCL 译码

极化码的容量可达性建立

在信道极化理论<sup>[8]</sup>上,具体为  $N$  个独立同分布的二元离散无记忆信道被极化为  $N$  个比特信道,且随着  $N$  趋于正无穷,  $N$  个比特信道的信道容量呈现两极分布:一部分信道容量趋于 1,另一部分信道容量趋于 0,且极化后的  $N$  个比特信道总容量与初始的  $N$  个二元离散无记忆信道总容量相等。极化码编码的基本思想是在容量趋于 1 的比特信道上传输信息比特,在容量趋于 0 的比特信道上传输收发双方已知的固定比特,从而实现可达容量的信道编码。

极化码是一类线性分组码,其生成矩阵为  $G_N = B_N F^{\otimes n}$ ,其中,  $N$  为极化码码长,  $B_N$  为比特翻转排列矩阵,  $n = \log_2 N$ , 矩阵  $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ ,  $F^{\otimes n}$  表示矩阵  $F$  的  $n$  阶 Kronecker 积。给定极化码编码器输入序列  $u_1^N = (u_1, u_2, \dots, u_N)$ , 极化码码字即为  $c_1^N = u_1^N G_N$ 。序列  $u_1^N$  包含 2 个子序列  $u_A = (u_i, i \in A)$  和  $u_{A^c} = (u_i, i \in A^c)$ 。集合  $A \subset \{1, 2, \dots, N\}$  为信息比特索引集合,  $u_A$  为信息序列。给定信息比特数量  $M$ , 根据极化码编码思想,集合  $A$  对应的  $M$  个比特信道应具有更高的信道容量<sup>[8]</sup>,或者更低的误判概率<sup>[12]</sup>,文献[12-14]分别给出了集合  $A$  的不同构造方法。集合  $A^c$  为集合  $A$  的补集,  $u_{A^c}$  一般取值为全 0。

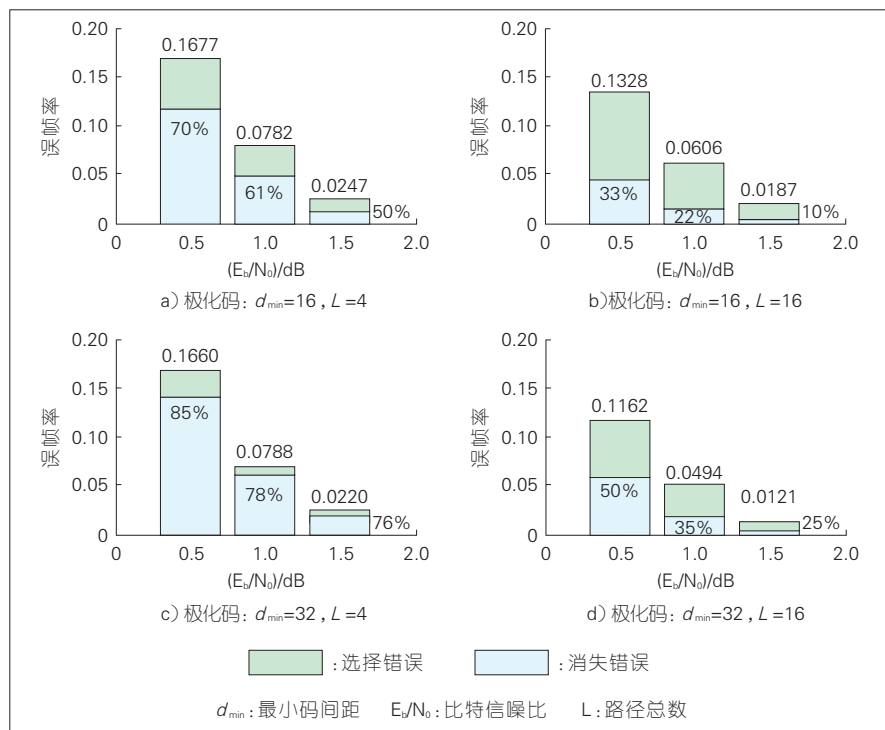
SCL 译码算法是极化码的一种近似 ML 译码算法<sup>[9]</sup>。该算法

的基本思想是在译码过程中保留  $L$  条似然概率最大的译码路径,当路径上的比特序列  $u_1^N$  判决结束之后,似然概率最大的路径上的信息序列作为译码结果被输出。SCL 译码过程中,记  $u_{i-1} (i=2, 3, \dots, N)$  处的  $L$  条译码路径为  $\hat{u}_{1,i-1}^l (l=1, 2, \dots, L)$ , 路径  $\hat{u}_{1,i-1}^{l-1}$  判决比特  $u_i$  扩展至  $\hat{u}_{1,i}^l$  的过程为:若  $u_i$  为信息比特,则每条路径  $\hat{u}_{1,i-1}^{l-1}$  在  $u_i$  处分裂为 2 条子路径  $\hat{u}_{1,i}^l = (\hat{u}_{1,i-1}^{l-1}, 0)$  和  $\hat{u}_{1,i}^{l+1} = (\hat{u}_{1,i-1}^{l-1}, 1)$ , 所得  $2L$  条子路径中似然概率最大的  $L$  条路径被保留;若  $u_i$  为固定比特,则每条路径  $\hat{u}_{1,i-1}^{l-1}$  在  $u_i$  处直接扩展为  $\hat{u}_{1,i}^l = (\hat{u}_{1,i-1}^{l-1}, 0)$ 。这种列表译码的思想促使 SCL 译码算法以较低的译码复杂度  $O(L \cdot N \log_2 N)$  即可达到极化码 ML 译码性能<sup>[9]</sup>。

## 2 SCL 译码错误分析

SCL 译码错误可根据正确路径在译码器中的存在状态分为 2 类:消失错误和选择错误<sup>[15]</sup>。消失错误是指比特判决之后,  $L$  条路径不包含正确路径,即正确路径在比特判决过程中从译码器中消失(被淘汰)。选择错误具体是指比特判决之后,正确路径存在于译码器中,但是正确路径的似然概率因小于某条错误路径而未被选择作为最终的译码结果。

SCL 译码的消失错误和选择错误主要受 3 个因素的影响:译码器路径总数  $L$ 、极化码的最小码间距  $d_{\min}$ 、SNR。图 1 的仿真



▲ 图1 串行抵消列表译码误帧率、消失错误比例对比图

示例直观地展示3个因素对译码错误的影响,该示例为2种不同最小码间距的极化码在不同路径总数  $L$  的译码器以及不同SNR下的误帧率、消失错误比例对比图。该仿真中,极化码码长  $N=256$ ,信息比特数量  $M=64$ ,信息比特集合  $A$  根据文献[13]和[14]分别构造,且最小码间距分别为16和32,调制方式和仿真信道分别为二进制相移键控调制(BPSK)和加性高斯白噪声(AWGN)信道。

由图1可知:1)当  $d_{\min}$ 、SNR一定,随着  $L$  增加,消失错误比例下降,也即  $L$  越大,则译码器越能包容正确路径,从而降低正确路径被淘汰的概率;2)当译码路径总数  $L$ 、SNR一定,  $d_{\min}$  越大,则消失错误比例越高,选

择错误比例越低,因为当比特判决之后,正确路径存在于译码器中,此时最小码间距越大,译码器误输出错误路径的概率越小,选择错误越少;3)当  $d_{\min}$ 、 $L$ 一定,随着SNR的增加,消失错误比例逐渐降低,即在低SNR下,正确路径在比特判决过程中往往被淘汰,出现消失错误。

根据上述影响SCL译码错误的3个因素可知:在给定信道和译码器参数(SNR、路径总数  $L$ 一定)的条件下,提升极化码在SCL译码下纠错性能的一个直接思路是增大  $d_{\min}$ 。对极化码级联外码是增大  $d_{\min}$  的一类有效方法<sup>[11]、[17]</sup>。在当前的级联方案中,CRC级联极化码<sup>[9]</sup>和PCC级联极化码<sup>[10]</sup>是2类有效的极化码级联方案,其有效性表现

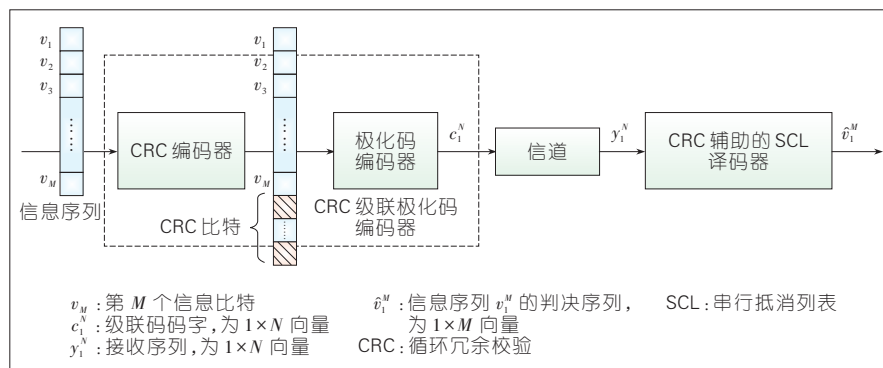
在:1)这2类级联码均可采用改进的SCL译码算法进行译码;2)这2类级联码均可通过码构造而增大  $d_{\min}$ <sup>[11]、[17]</sup>,从而降低SCL译码错误。这2类降低SCL译码错误的级联方案将在下一章节详细介绍。

### 3 降低SCL译码错误的级联极化码

#### 3.1 CRC级联极化码

(1)CRC级联极化码编、译码系统。

图2展示的是CRC级联极化码的编、译码系统示意图。在发送端,  $M$  长的信息序列  $v_1^M$  首先经过长度为  $L_{CRC}$  的CRC编码器编码得到长度为  $M+L_{CRC}$  的CRC码字,其中CRC码字的前  $M$  位为信息比特,后  $L_{CRC}$  位为CRC比特;其次,CRC码字经过内码极化码编码得到级联码码字  $c_1^N$ 。在接收端,接收信号  $y_1^N$  经过CRC辅助的SCL译码器输出信息序列的判决结果  $\hat{v}_1^M$ 。CRC辅助的SCL译码算法与传统的SCL译码算法主要区别在于2点:1)判决CRC比特时,译码路径按照信息比特的方式进行路径扩展;2)在比特判决结束之后,若译码器中存在满足CRC校验的路径,则能够输出满足CRC校验且似然概率最大的路径上对应的信息序列作为译码结果,否则则会直接输出似然概率最大的路径上对应的信息序列。



▲ 图2 CRC级联极化码编、译码系统示意图

(2) CRC 级联极化码性能提升原理。

CRC 辅助 SCL 译码器在比特判决之后, 通过 CRC 校验辅助选择正确路径, 从而显著降低选择错误。此外, 需要注意的是: 对于长度为  $L_{CRC}$  的 CRC 码, 内码极化码需要额外采用  $L_{CRC}$  个比特信道用于传输 CRC 比特, 以此保证整个级联码码率不变。由于额外引入  $L_{CRC}$  个信道质量更差的比特信道, 因此在相同的码率下, CRC 辅助 SCL 译码器往往比传统 SCL 译码器具有更多的消失错误。

图3展示的是码长  $N=256$ 、信息比特数量  $M=64$ 、 $d_{min}$  为 16 的极化码, 以及 CRC 长度为  $L_{CRC}=4$  的 CRC 级联极化码的误帧率、消失错误比例对比图。由图3可知: 1) 当路径数量较少时 ( $L=4$ ), 通过级联较短的 CRC 码几乎可完全消除选择错误, 但是由于  $L$  较小, 消失错误为主要成分; 因此额外引入 4 个信道质量较差的比特信道之后, CRC 级联极化码误帧率性能相比非级联极化码明显降低。2) 当路径

数量较多时 ( $L=16$ ), 通过级联较短的 CRC 码, 可显著降低选择错误, 但不能完全消除选择错误。此外, 由于  $L$  较大, 选择错误为主要成分, 级联 CRC 码之后, CRC 级联极化码误帧率性能相比非级联极化码明显提升。

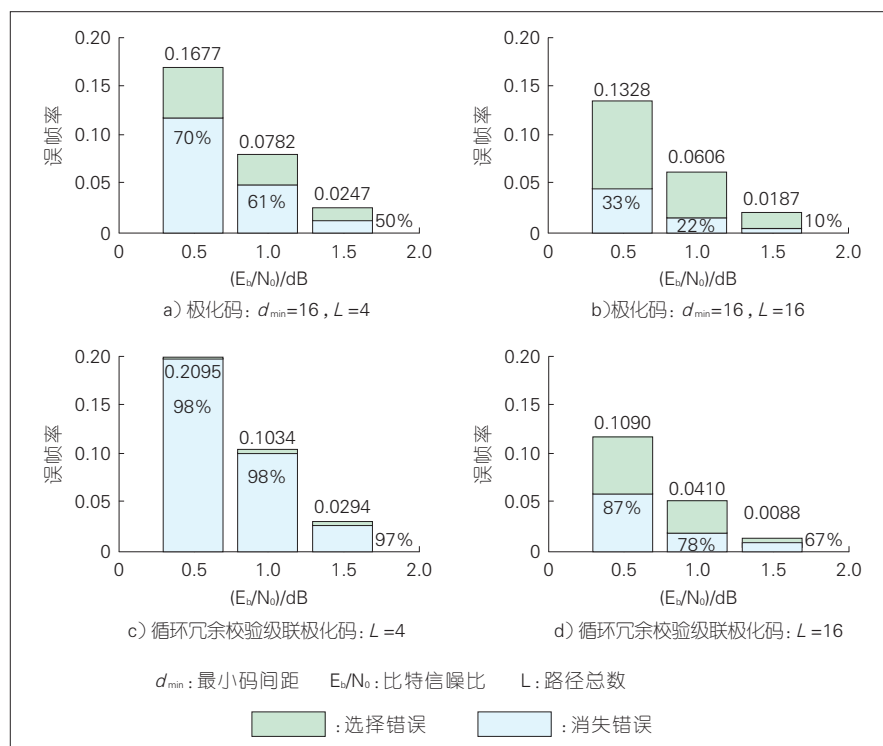
最后, 需要指出的是: 在 CRC 级联极化码中, CRC 码仍具有一定的链路层帧校验能力, 因

此往往为了保证 CRC 不可检测错误率 (UER), 在给定 UER 和译码器路径总数  $L$  的条件下, CRC 长度  $L_{CRC}$  满足  $L_{CRC} = \log_2 L - \log_2 UER$  [16]。

### 3.2 校验级联极化码

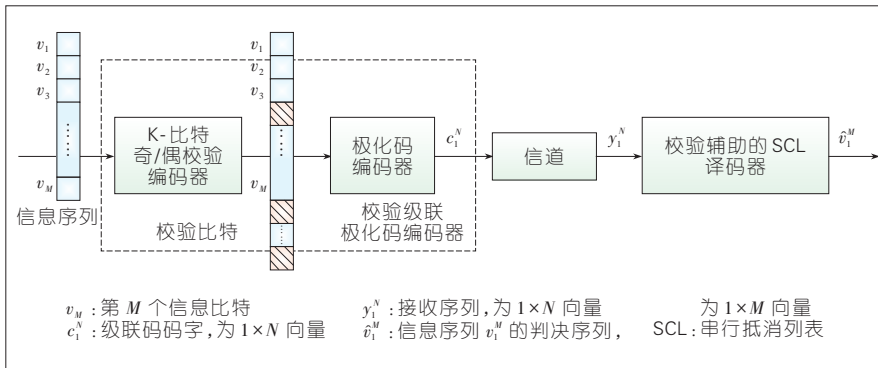
(1) 校验级联极化码编、译码系统。

图4展示的是校验级联极化码编、译码系统示意图。在发送端,  $M$  长的信息序列  $v_1^M$  首先经过  $K$  比特奇/偶校验码 (本文以偶校验码为例进行说明) 编码得到校验码码字, 如图5所示,  $K$  个偶校验比特可分散于信息序列中间, 而非集中于码字尾部。其次, 校验码码字经过极化码编码得到级联码码字  $c_1^N$ 。具体地, 校验级联极化码可采用四

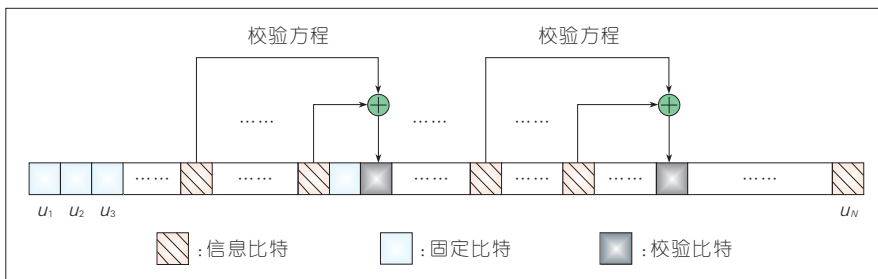


▲ 图3 极化码、循环冗余校验级联极化码误帧率和消失错误比例对比图





▲ 图4 校验级联极化码编、译码系统示意图



▲ 图5 校验级联极化码结构示意图

元组  $(N, I, P, \{T_k | k=1, 2, \dots, K\})$  表示, 其中,  $N$  表示极化码码长, 集合  $I$  表示信息比特索引集合, 集合  $P$  表示校验比特索引集合, 集合  $T_k$  表示参与第  $k$  个偶校验方程的信息比特索引集合。给定校验级联极化码  $(N, I, P, \{T_k | k=1, 2, \dots, K\})$ , 则第  $k$  个偶校验比特的编码公式如式 (1) 所示:

$$u_{p_k} = \sum_{i \in T_k} u_i \bmod 2, k=1, 2, \dots, K, \quad (1)$$

其中,  $p_k$  表示集合  $P$  的第  $k$  个元素, 对应第  $k$  个偶校验比特在序列  $u_1^N$  中的索引。在接收端, 接收信号  $y_1^N$  经过校验辅助的 SCL 译码器得到信息序列的判决结果  $\hat{v}_1^M$ 。校验辅助的 SCL 译码器与传统 SCL 译码器主要区别在于: 校验比特的判决值根据

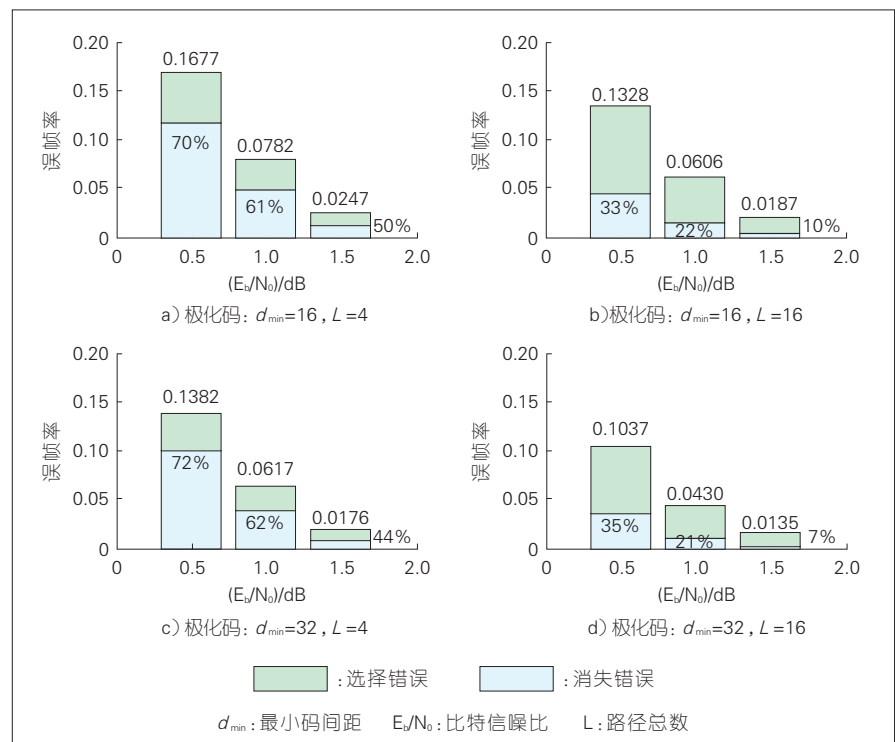
该校验比特所在的校验方程以及对应的信息比特判决值校验

得到。

(2) 校验级联极化码性能提升原理。

PCC 级联极化码相比 CRC 级联极化码具有更高的设计灵活性, 在不同的构造方法下, 其降低译码错误的原理不同。校验级联极化码的构造可分为 3 类: 1) 从降低选择错误的角度, 以最大化级联码最小码间距进行构造<sup>[17]</sup>; 2) 从降低消失错误的角度, 以最小化码字簇成对错误概率 (CPEP) 的构造<sup>[15]</sup>; 3) 从校验方程的硬件实现角度, 以循环移位寄存器为基础的伪随机构造<sup>[17], [18]</sup>。第 3 类构造方法一般可同时降低消失错误和选择错误。

图 6 展示的是码长  $N=256$ 、信息比特数量  $M=64$ 、最小码间距为 16 的极化码, 以及校验级



▲ 图6 极化码、校验级联极化码误帧率和消失错误比例对比图

联极化码的误帧率、消失错误比例对比图。其中,校验级联极化码的构造采用的是 5G 极化码标准方案中基于 5 位循环移位寄存器的伪随机构造<sup>[7],[18]</sup>。由图 6 可知:

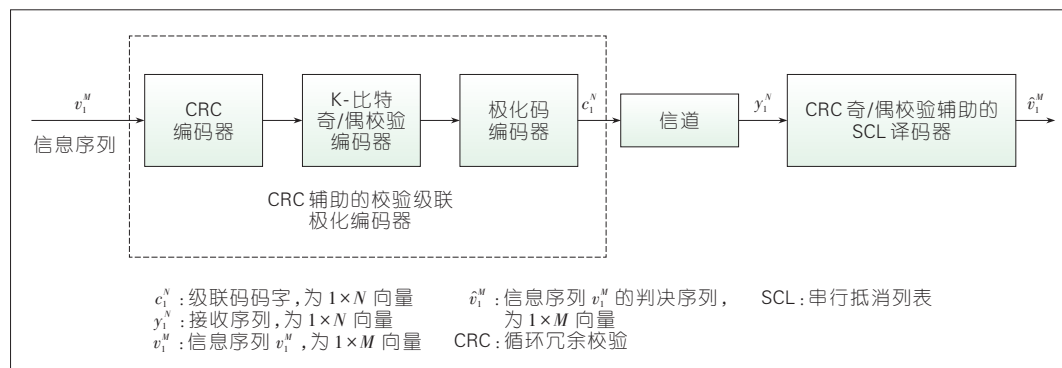
1) 在较低信噪比下 ( $E_b/N_0=0.5$  dB、1 dB),

对比极化码、校验级联极化码误帧率和消失错误比例可知,消失错误比例几乎不变,但是校验级联极化码误帧率明显较低,这表明伪随机构造的校验方程可同时降低消失错误和选择错误;2) 在较高信噪比下 ( $E_b/N_0=1.5$  dB),对比极化码、校验级联极化码消失错误比例可知,校验级联极化码消失错误比例明显降低,该结果体现了校验级联极化码在降低消失错误方面的优势。

#### 4 CRC 辅助的校验级联极化码及其性能

校验级联极化码可针对不同的 SCL 译码错误进行构造,相比 CRC 级联极化码具有更高的设计灵活性;但是其缺点在于不具备链路层帧检错能力。一个直接的改进方法是将 CRC 码和校验级联极化码结合,从而保留 CRC 码的帧校验能力。

图 7 展示的是 CRC 辅助的 PCC 级联极化码编、译码系统示意图。在发送端,CRC 辅助的校验级联极化码可视为传统的

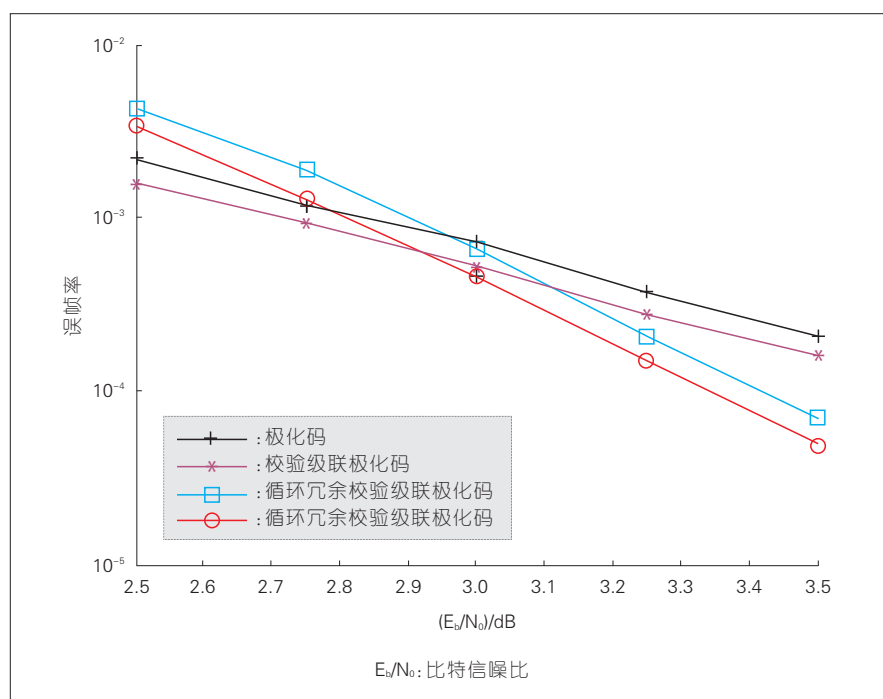


▲ 图 7 CRC 辅助的校验级联极化码编、译码系统示意图

CRC 级联极化码、校验级联极化码的组合。在接收端,CRC-校验辅助的 SCL 译码器与传统 SCL 译码器主要区别在于:1) 按照校验辅助的 SCL 译码器对每条路径上的比特进行判决;2) 按照 CRC 辅助的 SCL 译码器选择最终的输出路径。

图 8 展示了 BPSK 调制和 AWGN 信道下仿真了极化码及其 3 种级联方案的误帧率性能

能。仿真中,码长  $N=256$ ,信息比特数量  $M=64$ ,信息比特索引集合按照文献[13]构造,译码器路径总数  $L=8$ ,系统 UER 性能设定为  $1 \times 10^{-4}$ ;因此 CRC 长度设定为  $L_{CRC}=16$ ,CRC 生成多项式为  $g(x)=x^{16}+x^{12}+x^{11}+x^9+x^8+x^5+x^3+x+1$ 。在校验级联极化码、CRC 辅助的校验级联极化码中,校验方程采用 5G 极化码标准方案中基于 5 位循环移位寄存器



▲ 图 8 极化码及其级联方案的误帧率性能

的伪随机构造<sup>[7][18]</sup>。

从图 8 可知:1)对比极化码和校验级联极化码,对比 CRC 级联极化码和 CRC 辅助的校验级联极化码可知,引入校验比特可显著改善对应方案的纠错性能,在误帧率等于  $10^{-3}$  下,编码增益接近 0.1 dB。2)在较低信噪比下 ( $E_b/N_0 < 3$  dB),由于译码错误主要为消失错误,校验级联极化码呈现出最佳纠错性能。由于 CRC 级联极化码引入 16 个错误概率更高的比特信道传输 CRC 比特,使其消失错误更为严重,呈现出最差的纠错性能;3)在较高的 SNR 下 ( $E_b/N_0 > 3$  dB),由于译码错误主要为选择错误,其非级联极化码具有最差的纠错性能,此外, CRC 辅助的校验级联极化码可同时降低消失错误和选择错误,因此呈现出最佳的纠错性能。由于实际系统往往包含 CRC 码进行帧校验,在 3 种级联方案中, CRC 辅助的校验级联极化码展现出更高的实际应用价值。

## 5 结束语

CRC 级联极化码、校验级联极化码等级联方案纠错性能优良,具有较高的实际应用价值,因此分析这些级联码降低 SCL 译码错误的原理对于改善码的构造具有重要意义。本文中,我们总结了 SCL 译码错误类型,并从降低 SCL 译码错误的角度揭示了 CRC 级联极化码、校验级联极化码,以及 CRC 辅助的校验级

联极化码三者提升 SCL 译码性能的原理。最后,需要指出的是:当给定译码路径总数,使得 SCL 译码器不为 SC 译码器,也不能近似为 ML 译码器时,目前尚缺乏 SCL 译码器的误帧率性能分析理论表达式,该工作的推进将有助于最优级联极化码的构造。

### 参考文献

- [1] SHAFI M, MOLISCH A F, SMITH P J, et al. 5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice [J]. IEEE Journal on Selected Areas in Communications, 2017, 35(6): 1201–1221. DOI:10.1109/jsac.2017.2692307
- [2] BIOGLIO V, CONDO C, LAND I. Design of Polar Codes in 5G New Radio[EB/OL]. [2018–12–20]. <https://arxiv.org/abs/1804.04389>
- [3] RICHARDSON T, KUDEKAR S. Design of Low-Density Parity Check Codes for 5G New Radio [J]. IEEE Communications Magazine, 2018, 56(3): 28–34. DOI:10.1109/mcom.2018.1700839
- [4] 徐慧俊. 5G 商用, 蓄势待发[J]. 中兴通讯技术, 2018, 24(1): 2–5. DOI:10.3969/j.issn.1009–6868.2018.01.001
- [5] 史治平. 5G 先进信道编码技术[M]. 北京:人民邮电出版社, 2017
- [6] IMT. IMT Vision – Framework and Overall Objectives of the Future Development of IMT for 2020 and Beyond: ITU–R M.2083–0 [S]. 2015
- [7] 3GPP. Technical Specification Group Radio Access Network, NR; Multiplexing and Channel Coding (Release 15) [S]. 2017
- [8] ARIKAN E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels [J]. IEEE Transactions on Information Theory, 2009, 55(7): 3051–3073. DOI:10.1109/tit.2009.2021379
- [9] TAL I, VARDY A. List Decoding of Polar Codes [J]. IEEE Transactions on Information Theory, 2015, 61(5): 2213–2226. DOI: 10.1109/tit.2015.2410251
- [10] WANG T, QU D M, JIANG T. Parity-Check-Concatenated Polar Codes [J]. IEEE Communications Letters, 2016, 20(12): 2342–2345. DOI:10.1109/lcomm.2016.2607169
- [11] ZHANG Q S, LIU A J, PAN X F, et al. CRC Code Design for List Decoding of Polar Codes[J]. IEEE Communications Letters, 2017, 21(6): 1229–1232. DOI:10.1109/lcomm.2017.2672539
- [12] MORI R, TANAKA T. Performance of Polar Codes with the Construction Using Density Evolution [J]. IEEE Communications Letters, 2009, 13(7): 519–521. DOI:10.1109/lcomm.2009.090428

- [13] TAL I, VARDY A. How to Construct Polar Codes [J]. IEEE Transactions on Information Theory, 2013, 59(10): 6562–6582. DOI: 10.1109/tit.2013.2272694
- [14] LI B, SHEN H, LAND Ingmar, A RM-Polar Codes [EB/OL].[2018–12–20] <https://arxiv.org/abs/1407.5483>
- [15] WANG T, QU D, JIANG T. Cluster Pairwise Error Probability and Construction of Parity-Check-Concatenated Polar Codes [EB/OL]. [2018–12–20]. <https://arxiv.org/abs/1810.04458>
- [16] Samsung. Performance of Short-length Polar Codes: R1–1609072:3GPP TSG RAN WG1 Meeting #86[R]. Portugal, 2016
- [17] PARK J, KIM I, SONG H Y. Construction of Parity-Check-Concatenated Polar Codes Based on Minimum Hamming Weight Codewords [J]. Electronics Letters, 2017, 53(14): 924–926. DOI:10.1049/el.2017.1037
- [18] Huawei HiSilicon. Details of the Polar Code Design: R1–1611254:3GPP TSG RAN WG1 Meeting #87[R]. USA, 2016

### 作者简介



王涛, 华中科技大学电子信息与通信学院在读博士研究生; 主要研究方向为信道编码。



屈代明, 华中科技大学电子信息与通信学院教授; 主要研究方向为多载波通信、大规模多天线等; 已发表论文 100 余篇。



江涛, 华中科技大学电子信息与通信学院教授; 主要研究方向为移动通信系统与理论、多载波通信、天地一体化信息网络等; 已发表论文 300 余篇, 出版著作 6 本。

# 高频通信技术

## High Frequency Communication Technology

**摘要:**毫米波是第5代移动通信系统的重要通信技术,利用大带宽承载大量信息,有效提升系统的吞吐量。第3代合作伙伴计划(3GPP)已经对毫米波频段做了划分,各个国家逐步对频谱进行分配。毫米波因其频谱特性,需要与大规模天线阵列结合使用,获得波束赋形增益。通过对波束设计与波束管理等核心技术的创新,实现热点高速传输。同时,通过高低频无线协作组网的方式,实现整个区域内的基础覆盖,以适应不同的通信场景需求。

**关键词:**毫米波;混合波束赋形;波束管理;高低频协作组网

**Abstract:** Millimeter wave is an important technology for the 5th generation mobile communication system. By utilizing wide frequency band for high data transmission, millimeter wave can greatly increase the experienced data rates of users and the system throughput. The millimeter-wave bands have been standardized by third Generation Partnership Project (3GPP), and the spectrum has been allocated gradually by many countries. Due to its spectral characteristics, millimeter waves need to be combined with large-scale antenna arrays to obtain beamforming gain. The high-speed transmission of hotspots is realized by the innovation of beam design and beam management. In addition, the basic coverage in the entire area and the requirement for different communication scenarios can be satisfied by multi-frequency wireless cooperative networking.

**Key words:** millimeter wave; hybrid beamforming; beam management; multi-frequency cooperative networking



李萍/LI Ping  
魏浩/WEI Hao  
黄静月/HUANG Jingyue

(中兴通讯股份有限公司, 广东 深圳  
518057)  
(ZTE Corporation, Shenzhen 518057, China)

随着社会经济的日益发展,无线通信中的数据业务量持续增长,在近几年来呈现出指数的爆发态势。数以亿计的智慧终端将介入网络,相互连接交互信息,使得业务和应用更加多样化和多元化<sup>[1]</sup>。第5代移动通信系统(5G)也称为IMT-

2020,已经成为当前最为热门的研究方向和研究领域。根据IMT-2020(5G)推进组《5G愿景与需求白皮书》的要求,5G还将大幅提高网络部署和运营效率,与4G相比,频谱效率提升5~15倍,能量效率和成本效率提升百倍以上,5G的容量预计是4G的

DOI: 10.12142/ZTETJ.201901003  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190129.1118.002.html>

收稿日期: 2018-12-28  
网络出版日期: 2019-01-29



1 000 倍<sup>[2]</sup>。

然而,作为构建新一代信息基础设施重要载体,无线电频谱资源愈发稀缺,资源结构性紧缺问题愈发突出<sup>[3]</sup>;因此,在规划和分配 5G 系统频率的时候,基于 6 GHz 以上高频段的频谱划分和使用,就自然成为大势所趋。毫米波具体是指波长在 1~10 mm 的电磁波,其频率大约在 30~300 GHz 之间。毫米波可以采用大载波带宽承载大量数据信息,能够支持 5G 时代 0.1~1 Gbit/s 的用户体验速率要求。

近年来,工业界对 5G 毫米波频段的商用推进愈发加快。Verizon、Vodafone、中国移动、中国电信、中国联通等全球知名电信运营商,相继公布其在 5G 毫米波频段的布局 and 试验进展,毫米波通信的商用已经近在眼前。

## 1 5G 高频频谱分配

### 1.1 协议标准的频谱划分

2018 年 7 月,针对 5G 毫米波频谱规划问题,未来移动通信论坛 5G 微波毫米波特别工作组发布了《5G 毫米波频谱规划建议白皮书》,建议在 2018 年前分别完成 24.75~27.5 GHz 以及 37~42.5 GHz 的 5G 频率规划等,以给产业明确指导方向<sup>[4]</sup>。标准化方面,国际电联无线电通信部门 (ITU-R) 在 2017 年底发布的 IMT-2020 技术评估报告中,室内热点及密集城区增强型移动宽带 (eMBB) 场景中以 4 GHz (代

表以 3.5 GHz 为主的频段区间) 和 30 GHz (代表 26 GHz、39 GHz 等毫米波频段) 作为 5G 技术评估的典型频段配置。第 3 代合作伙伴计划 (3GPP) 5G 标准作为主流的 5G 技术体制,目前获得了最为广泛的国家及移动通信产业的支持。3GPP 新空口 (NR) 毫米波频段的射频标准讨论和制定工作由 3GPP 无线接入网络 (RAN) 4 牵头开展,26 GHz 和 39 GHz 为全球范围内备受关注的毫米波频段,被首先纳入了 3GPP 标准,为相关毫米波频段的产业培育及设备开发制造创造条件并提供保障。

目前 3GPP 定义的频率范围分为频率范围 (FR) 1 和 FR2。FR1 是低于 6 GHz 的部分,频率范围为 450~6 000 MHz,覆盖能力越强穿透能力越好。FR2 范围主要是高频,也就是我们通常说的毫米波频段,频率范围为 24 250~52 600 MHz,穿透能力较弱,但带宽充足,频谱干净,未来的应用十分广泛。在 FR2 频段中,又继续细分了不同的频段,并做了标号,如表 1 所示。

### 1.2 各国频谱分配

各国电信监管机构在分配 5G 频谱时都会通盘考虑已分配频谱和未分配频谱。目前全球主流运营商都已对频谱进行了分配,并制定了对未来 5G 频谱的分配计划,如表 2 所示。

在 2018 的韩国平昌冬奥会上,Intel 和韩国运营商 KT 部署

▼表 1 频率范围 2 中的 NR 工作频段 (3GPP 已标准化的毫米波频段)

NR 频段	UL 和 DL 频段	双工模式
n257	26 500 ~ 29 500 MHz	TDD
n258	24 250 ~ 27 500 MHz	TDD
n260	37 000 ~ 40 000 MHz	TDD
n261	27 500 ~ 28 350 MHz	TDD

DL: 下行链路  
NR: 新空口  
TDD: 时分双工  
UL: 上行链路

了迄今为止规模最大的 5G 网络,在 10 个奥运场馆搭建 22 个 5G 链路,支持交互式网络电视 (IPTV)、虚拟现实、Wi-Fi 等应用,还为观众和游客提供千兆级速度服务。同时,为备战 2020 年东京奥运会,日本加速推进 5G 技术开发。美国在其可持续频谱战略中提出:要确保美国在 5G 中处于领导地位,促进就业增长和经济增长,并保护国家安全。中国在国务院发布的《“十三五”国家信息化规划》中,16 次提到了 5G,提出在 5G 网络技术上要走在全球前列。

5G 不仅是无线通信产业的一次升级换代,更是一次重大的技术变革,与数字化转型技术、人工智能 (AI) 技术一起,成为国民经济转型升级的重要推动力。5G 已经成为国与国之间战略竞争的重要组成部分。

## 2 传输特性与信道建模

高频的信道特性和模型是 5G 高频关键技术和系统设计的关键环节之一。对于室外到室内的穿透损耗,高频中不再区分 Umi/Uma 场景,且模型中增

▼表2 各国频谱分配现状与计划

国家	高频段频谱分配情况
新加坡	重点关注 28 GHz 毫米波频段
日本	计划于 2019 年 3 月分配 27 ~ 29.5 GHz 的 5G 毫米波频段
韩国	2017 年初发布国家宽带/频谱计划, 高频段为 27.5 ~ 28.5 GHz 频段, 计划于 2019 年 3 月开始 5G 商用
俄罗斯	计划于 2020 年开始部署 5G, 2024 年覆盖所有大城市, 目前 5G 测试高频段为 25.25 ~ 29.5 GHz 频段
德国	确认 26 GHz (24.25 ~ 27.5 GHz)、28 GHz (27.5 ~ 29.5 GHz) 和 32 GHz (31.8 ~ 33.4 GHz) 为 5G 毫米波频段, 并认为 26 GHz 最重要
英国	确定 24.25 ~ 27.5 GHz 作为 5G 的高频段
美国	规划 4 大高频段用于 5G 移动网络和固定无线: 27.5 ~ 28.35 GHz (28 GHz)、37 ~ 38.6 GHz (37 GHz)、38.6 ~ 40 GHz (39 GHz) 和 64 ~ 71 GHz。其中, 28 GHz、37 GHz 以及 39 GHz 为授权频谱, 64 ~ 71 GHz 为未授权频谱, 合计约 11 GHz 带宽
中国	2017 年, 工信部批复了 24.75 ~ 27.5 GHz 和 37 ~ 42.5 GHz 高频段用于 5G 技术研发试验, 这样可确保未来每家运营商平均获得至少 2 000 MHz 带宽的 5G 高频段

加标准方差, 其中穿透损耗取决于墙体材料 (如玻璃、水泥、木头等)。对于信道模型的小尺度参数, 如 delay spread、发射端离开角 (AoD)、接收端到达角 (AoA)、散射簇到移动端的俯仰到达角 (ZoA) 等, 高频模型中均与载波频率相关<sup>[5]</sup>。3GPP 于 2016 年 6 月的第 72 次全会上启动了 5G 高频信道模型 (0.5 ~ 100 GHz) 的研究工作。为了凸显高频段与低频传播特性的区别, 5G 高频信道模型中引入了更多与高频相关的新特性, 包括: 氧衰模型、大带宽及天线阵列模型、空间一致性模型、阻挡模型、多频段相关性模型、时变多普勒模型、用户 (UE) 转动模型和确定性地面反射模型、随机簇模型等。图 1 为 IMT-2020 给出的信道模型模块。

混合信道模型以电磁理论物理学为基础构建信道模型, 适合 400 MHz ~ 100 GHz 多种带宽的配置, 对于当前 5G 对信道的

空间相关性、时间相关性、频率相关性等特性自然满足, 具有很强的向后兼容性。图 2 是混合信道模型。

混合信道模型包括确定性模型、统计模型 2 大部分, 确定性模型主要基于确定性场景 3D 模型, 利用射线追踪技术, 考虑直射、透射、反射、绕射等物理现象, 进行确定性计算, 得到发射及接收的主要射线径确定性结果。统计模型主要反映确定

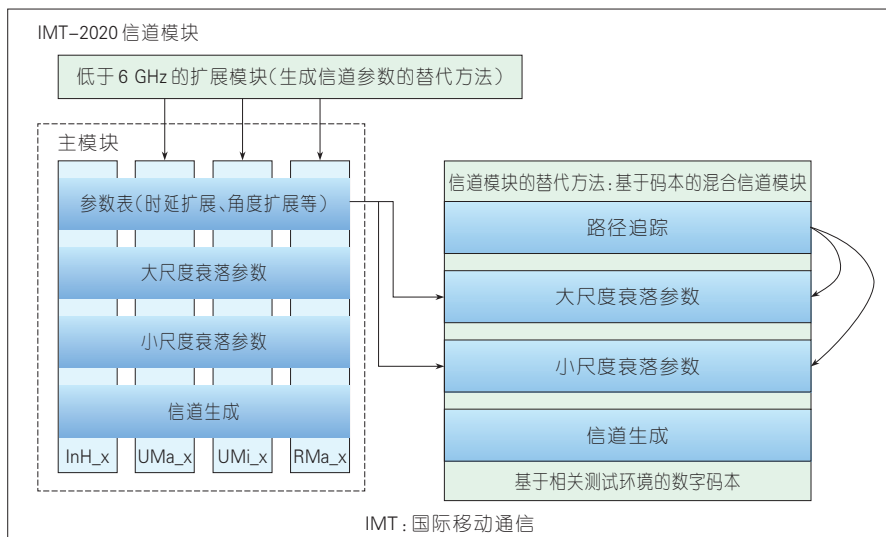
性模型中未进行建模的小物件、粗糙面以及因为人流、车辆、植被等引发的散射、阻挡和闪烁反射现象。

### 3 波束与赋形

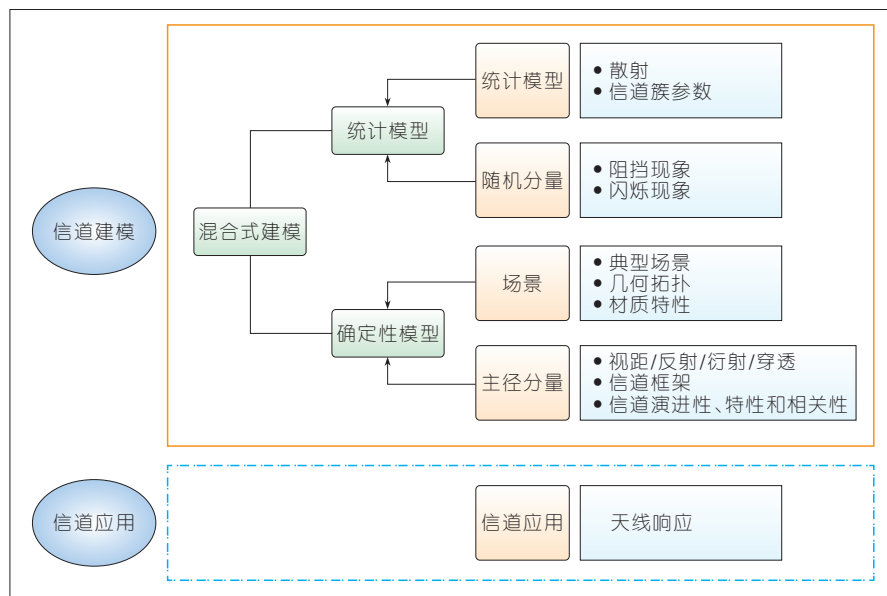
毫米波与低频段电磁波相比, 波长比较短, 面临的一个主要问题是自由空间路损使得接收端信号产生大幅度衰减<sup>[6]</sup>。但同时, 较短的波长使其能够在发射端部署大规模天线阵列以提供显著的波束赋形增益, 以抵抗路径损耗带来的性能损失。波束就成了毫米波的标配属性, 被更多地研究与设计。

#### 3.1 赋形结构

全数字波束赋形的大规模多输入多输出 (MIMO) 系统虽然可以产生最优性能, 然而射频 (RF) 链路个数需要与基站端部署的天线数量相同, 硬件复杂度和成本、信号处理的复杂度和能耗均迅速增加<sup>[7]</sup>。模数混合波束



▲图1 5G信道模型模块图



▲ 图2 混合信道模型总体思路

赋形结构由一个低维度的数字波束和一个高维度的模拟波束组成,不需要如同传统纯数字波束赋形结构一般的RF链路的数量,有效降低实现成本和设备体积,成为毫米波通信的主流结构方式<sup>[8]</sup>,如图3所示。

在图3中, $N_t$ 代表发射端总天线数,而 $N_{RF}^t$ 代表发射端的RF链路个数。每条RF链路都将连接全部的天线,实际RF链路数为 $N_t N_{RF}^t$ 条。在满足一定的条件时,可以得到和纯数字波束赋形结构同等的波束赋形增益,但结构的复杂度和成本都相当高。

因此,可以采用低复杂度和低成本的其他连接结构。

(1)部分连接结构:每条RF链路连接的天线数为 $N_t/N_{RF}^t$ ,该种结构牺牲了一部分的波束赋形增益,但在硬件实现上具有显著优势<sup>[9]</sup>。

(2)重叠子阵结构:每条RF链路连接的天线数在 $N_t/N_{RF}^t$ 到 $N_t$ 之间,是全连接结构和部分连接结构的折中。

(3)双倍移相器结构:每1个等效的模拟波束赋形因子通过2个移相器实现,能够调节的幅度和相位选择更多,实现更好的性能。

(4)开关结构:通过控制对

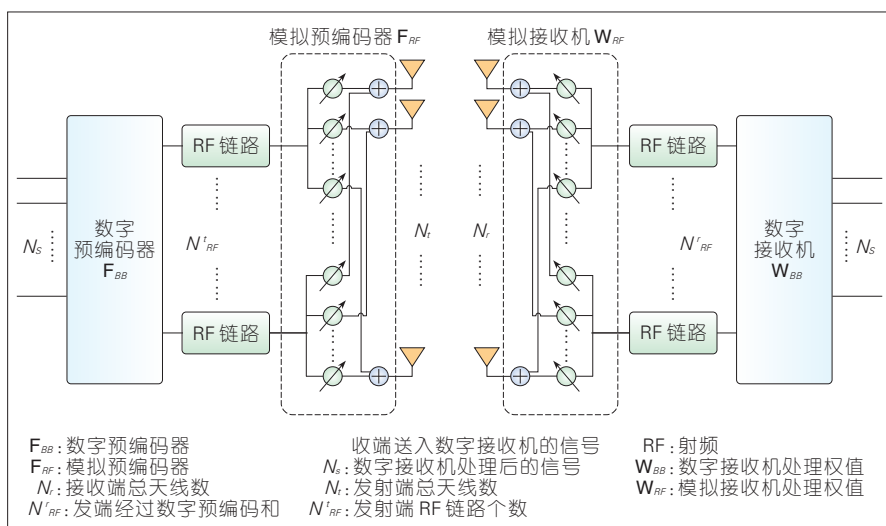
应通路的开关,利用 $N_c$ 个相移器实现幅度和相位的控制,其性能决定于 $N_c$ 的数量。相对于双倍相移器结构来说,开关结构能够达到的精度更高,性能更好。

在实际系统中,模拟赋形增益的效果取决于多种因素,包括移相器的精度、RF链路的长度、器件响应时延、功率放大器的非线性等<sup>[10-12]</sup>;因此,采用何种结构需要在硬件实现复杂度和波束赋形增益上做出权衡。

### 3.2 波束算法

波束是毫米波通信不同于低频段通信的主要特征,基站通过波束赋形的方式来进行毫米波覆盖;因此波束设计与波束管理就必然成为毫米波通信的核心技术<sup>[13]</sup>。

波束设计中对于控制波束和业务波束的管理,可以独立进行也可以统一调整,这2种方式可以通过高层信令配置来实现。对于控制波束,考虑其传输



▲ 图3 模数混合波束赋形结构



的鲁棒性,可以在不同时刻使用不同方向的波束配置给终端控制信道,以解决高频信道突发阻塞的问题。对于业务波束,根据具体场景,相对较窄的波束会有较好的模拟赋形增益和干扰屏蔽效果,更适合于对移动性需求相对较低、用户位置相对较固定的场景。而设计相对较宽的波束,可以减少总波束的个数,降低测量开销,用户移动出波束的时间也更长,有利于更好地跟踪波束<sup>[14]</sup>。

波束管理包括波束扫描、波束跟踪<sup>[15]</sup>和波束恢复。基站和UE通过其最优的波束对实现连接通信,当相邻波束的能量大于当前能量时立即发生切换。基站侧和UE侧分别发起波束轮扫,通过一轮完全轮扫,基站侧与UE侧共同维护了一个最优波束对。在基站侧和UE侧使用最优波束对时,继续进行基站侧和UE侧波束测量与维护;当UE侧发生位置变化或者旋转时,UE侧可以及时切换到最优波束上。同时,基站侧也可以根据UE位置变化来调整基站侧波束。如果当前由于遮挡导致所有候选波束均不可用时,UE侧发起波束失败恢复流程(BFR),向基站发起同步请求,当收到基站侧反馈时则恢复成功,实现快速同步。采用良好的波束管理方案,使得系统在工作过程中,尽可能对波束切换无感知,即波束切换对系统性能没有明显的影响。

当前人工智能(AI)在各个领域掀起了应用热潮,我们还可以将AI与传统通信相结合,基于AI机器学习算法,对实际的通信场景进行分类和识别,以实现波束的自适应设计与管理策略,从而快速跟踪无线信道场景的变化<sup>[16]</sup>。

#### 4 协作组网

高低频无线协作组网是5G网络架构的必然发展趋势,主要是宏基站通过低频段实现整个区域内的基础覆盖,微基站通过高频段承担热点覆盖和高速传输,以满足5G网络更高数据流量、更快用户体验速率、海量终端连接和更低时延的需求<sup>[17-18]</sup>。

在组网方式中,高低频可以各自独立组网,也可以通过双连接部署实现组网,如图4所示。

在独立组网方式中,5G低频/长期演进(LTE)主要用于基础覆盖和移动性保证,而5G高频主要用于提升用户吞吐量,5G低频/LTE与5G高频之间通过小区选择、重选或负载均衡策略接入用户。在双连接部署组

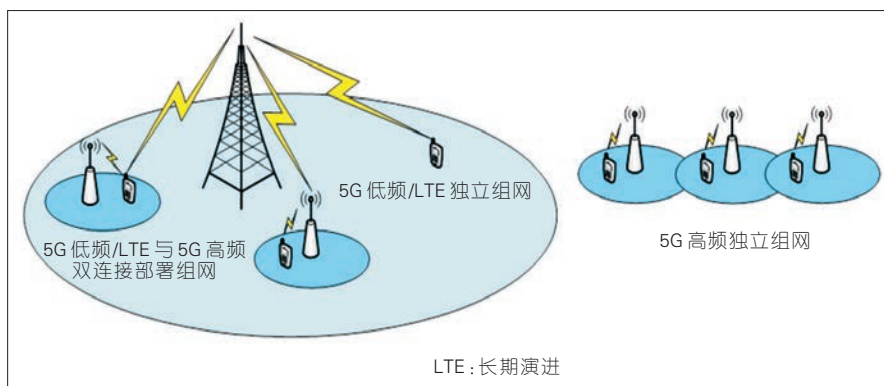
网方式中,5G低频/LTE与5G高频之间通过双连接实现共同组网<sup>[19]</sup>,采用集中单元(CU)和分布单元(DU)分离技术,由CU集中对低频和高频资源做统一管理,并对高低频双连接做数据流的分发。无线资源控制(RRC)和分组数据汇聚协议(PDCP)在CU,无线链路层控制协议(RLC)/媒体访问控制(MAC)/物理层(PHY)在DU。

在实际通信中,具体的组网方式,可以根据具体的组网场景来进行选择。

##### 4.1 组网场景

高频段通信属于5G中的eMBB场景,主要承担热点高速传输,可以细分为如下几种组网场景。

(1)高频孤立小区组网。在热点区域开启单个高频5G小区,小区周围一定区域内没有同频的高频小区。该小区不存在小区间干扰,与低频5G或LTE进行互操作来保证用户的体验。该组网方式的缺点是高频5G存在严重遮挡后的覆盖盲区。



▲ 图4 高低频无线协作组网



(2)高频连续覆盖小区同频组网。这种组网方式就是传统意义上的三叶草同频组网,小区间干扰将降低用户的信号与干扰加噪声比(SINR)和体验,如管理波束规避干扰,也需要重点考虑。

(3)高频连续覆盖小区异频组网。这种组网方式下的小区间干扰将会大大降低,但是频谱资源将有数倍的消耗,取决于运营商是否有足够的频谱。

(4)超密集区域小区组网。如演唱会、运动赛、大规模会议以及球赛等场景,业务需求量很大,高频5G正好能够满足各方面的需求且可以根据覆盖区域、容量等需求灵活调整频率资源。

(5)室内大容量组网。由于室内的房间、隔墙等建筑物构成了天然的隔离,以及室内丰富的多径的存在,使得高频5G尤其适用于室内覆盖,而且可以根据覆盖和容量需求调整频谱资源,选择合适的产品形态。

(6)室外向室内的覆盖。高频穿透玻璃后的效果是非常好的,如大的玻璃幕墙办公区、商区等,但水泥墙体对高频损耗比较大。

## 4.2 关键技术

根据上文所述的组网场景,高低频协作组网的关键技术研究主要涉及以下3个方面。

### (1)对移动终端连续覆盖。

实际组网时,对于每个孤立的高频单站,很容易出现由于固

定或者流动遮挡导致的覆盖空洞。对于高频参与的组网场景,如何在多个站的穿插覆盖区域内保持对移动终端的连续覆盖服务是技术方案首要考虑的问题。该问题包括融合空口设计的考虑、支持数模混合赋型的高频基站的最优设计(射频天线的偏硬件设计、预编码码本等算法设计),以及波束切换、波束搜索等波束管理方面的设计。在支持多用户通信的需求下,还需要对大规模MIMO设备和算法能力提出更高要求。

### (2)移动终端业务连续性。

在超密集组网场景中,高频信号覆盖半径小,多站穿插覆盖部署,导致移动终端在该区域内的业务服务在多个基站间不断快速切换,这需要保证用户业务的连续性,尽可能使得用户对切换无感知。因此,在方案中如何分层调度、具体切换或调度的算法以及包括终端参与的流程等都是需要研究的问题<sup>[20]</sup>。

### (3)协作组网规模能力。

高低频协作组网要针对相应实际可能部署场景,提供整体的解决方案。在组网能力下,关于服务用户数、平均吞吐率,以及覆盖边缘的定义必须明确清晰。同时,高频和低频协作的范围能力、负载均衡等,也是高低频融合组网规模能力的关键问题。这些关键问题的研究,涉及融合空口设计、CU分离和双连接的具体架构设计,以及完整的终端技术方案和移动性管理方

案。此外,还需要结合信道建模基础上的系统仿真,以及验证性样机系统原型开发。

## 5 结束语

中兴通讯从2013年开始对5G无线通信进行研究,在毫米波通信方面有着深厚的技术积累,包括实际应用场景的探究、关键技术的掌握研发、组网方面的策略与方案设计等。目前,已经通过由工信部主导、IMT-2020(5G)推进组负责实施的5G国测第3阶段测试。同时完成5G实验网的建设,与全球知名工业界厂商进行联合测试。2019年将迎来5G预商用建网。

2018年11月,工信部IMT-2020(5G)推进组无线技术工作组表示,启动6G的概念研究工作。未来的6G移动通信系统,是一种泛在融合信息网络,可达到兆比特每秒级别的信息传输速率。为了支持如此高的吞吐量和高速组网,使人们对新频率资源开发的目光自然地从小波频段转向波长更短且从前较少关注的太赫兹频段。毫米波频段已经在学术界和工业界得到广泛认可,而太赫兹无线通信作为一个新兴领域,其学术价值和应用前景更是不可估量。未来是学术界和工业界融合的时代,无线通信也必将迎来崭新的发展。

## 致谢

本文得到了中兴通讯股份

有限公司无线研究院算法部窦建武、柯雅珠、张波、李杰、田力、武艺鸣的大力帮助和支持，谨致谢意！

## 参考文献

- [1] WANG D, ZHANG Y, WEI H, et al. An Overview of Transmission Theory and Techniques of Large-scale Antenna Systems for 5G Wireless Communications [J]. Science China Information Sciences, 2016, 59(8): 081301:1-081301:18. DOI: 10.1007/s11432-016-0278-5
- [2] IMT-2020(5G)推进组. 5G 愿景与需求白皮书 [EB/OL]. (2014-05-28)[2018-12-28]. <http://www.imt-2020.cn>
- [3] 工业和信息化部. 国家无线电管理规划 (2016—2020 年) [EB/OL]. (2016-08-29)[2018-12-28]. <http://www.miit.gov.cn>
- [4] 未来移动通信论坛. 5G 毫米波频谱规划建议白皮书 [EB/OL]. (2018-07-01)[2018-12-28]. <http://www.future-forum.org>
- [5] WANG C, BIAN J, SUN J, et al. A Survey of 5G Channel Measurements and Models [J]. IEEE Communications Surveys & Tutorials, 2018, 20(4):3142-3168. DOI: 10.1109/COMST.2018.2862141
- [6] PI Z, KHAN F. An Introduction to Millimeter-wave Mobile Broadband Systems [J]. IEEE Communications Magazine, 2011, 49(6):101-107. DOI: 10.1109/MCOM.2011.5783993
- [7] MOLISH F, RATNAM V, HAN S, et al. Hybrid Beamforming for Massive MIMO - A Survey [J]. IEEE Communications Magazine, 2017, 55(9):134-141. DOI: 10.1109/MCOM.2017.1600400
- [8] HUANG J, CHENG Z, CHEN E, et al. Low-Complexity Hybrid Analog/Digital Beamforming for Multicast Transmission in mmWave Systems [C]//International Conference on Communications. USA: IEEE, 2017:1-6. DOI: 10.1109/ICC.2017.7997049
- [9] GENG J, WEI Z, WANG X, et al. Multiuser Hybrid Analog/Digital Beamforming for Relatively Large-scale Antenna Arrays [C]//Globecom Workshops. USA: IEEE, 2014: 123-128. DOI: 10.1109/GLOCOMW.2013.6824973
- [10] DEMIR T, TUNCER T. Hybrid Beamforming with Two Bit RF Phase Shifters in Single Group Multicasting [C]//International Conference on Acoustics, Speech and Signal Processing. USA: IEEE, 2016:3271-3275. DOI: 10.1109/ICASSP.2016.7472282.
- [11] AYACH O, RAJAGOPAL S, ABUSURRA S, et al. Spatially Sparse Precoding in Millimeter Wave MIMO Systems [J]. IEEE Transactions on Wireless Communications, 2014, 13(3):1499-1513. DOI: 10.1109/TWC.2014.011714.130846
- [12] RUSU C, MENDEZ-RIAL R, GONZALEZPRELCICY N, et al. Low Complexity Hybrid Sparse Precoding and Combining in Millimeter Wave MIMO Systems[C]// 2015 IEEE International Conference on Communications (ICC). USA: IEEE, 2015:1340-1345. DOI: 10.1109/ICC.2015.7248509
- [13] GIORDANI M, POLESE M, ROY A, et al. A Tutorial on Beam Management for 3GPP NR at mmWave Frequencies [J].IEEE Communications Surveys & Tutorials, 2018: 1-22. DOI: 10.1109/COMST.2018.2869411
- [14] WANG J, ZHU H, DAI L, et al. Low-Complexity Beam Allocation for Switched-Beam Based Multiuser Massive MIMO Systems [J]. IEEE Transactions on Wireless Communications, 2016, 15(12):8236-8248. DOI: 10.1109/TWC.2016.2613517
- [15] XUE Q, FANG X, XIAO M, et al. Beam Management for Millimeter Wave BeamSpace MU-MIMO Systems [J]. IEEE Transactions on Communications, 2018:1-14. DOI: 10.1109/TCOMM.2018.2867487
- [16] WANG J, WANG J, WU Y, et al. A Machine Learning Framework for Resource Allocation Assisted by Cloud Computing [J]. IEEE Network, 2018, 32(2):144-151. DOI: 10.1109/MNET.2018.1700293
- [17] ZHANG Y. Energy Efficiency Analysis of Heterogeneous Cellular Networks with Extra Cell Range Expansion [J]. IEEE ACCESS, 2017, (5):11003-11014. DOI: 10.1109/ACCESS.2017.2713814
- [18] BHUVANESWARI P T V, INDU S, SHIFANAN L, et al. An Analysis on Cell Range Expansion in 4G LTE Networks[C]// International Conference on Signal Processing, Communication and Networking. USA: IEEE, 2015:1-6. DOI: 10.1109/ICSCN.2015.7219824
- [19] MIDDLETON R H, WIGREN T, LAU K, et al. Data Flow Delay Equalization for Feedback Control Applications Using 5G Wireless Dual Connectivity[C]//Vehicular Technology Conference (VTC Spring). USA:IEEE, 2017: 1-7. DOI: 10.1109/VTCSpring.2017.8108608
- [20] NAM W, BAI D, LEE J, et al. Advanced Interference Management for 5G Cellular Networks[J]. IEEE Communications Magazine, 2014, 52(5): 52-60. DOI:10.1109/MCOM.2014.6815893

## 作者简介



**李萍**，中兴通讯股份有限公司无线研究院资深算法系统工程师；从事 5G NR 高频无线技术研究和算法设计工作，主要研究方向为系统架构设计、基带算法方案设计和信号检测解调和网络特性优化等。



**魏浩**，中兴通讯股份有限公司无线研究院算法系统工程师；从事 5G NR 无线通信算法预研与设计工作，主要研究方向为信道估计、信号检测、信道译码、混合预编码等；发表论文 20 余篇。



**黄静月**，中兴通讯股份有限公司无线研究院算法工程师；从事 5G NR 毫米波通信算法研究与开发工作，主要研究方向为信道估计、混合预编码、多播波束赋形等。

# “太极混一”——极化码 原理及 5G 应用

## The Principle and Application of Polar Codes in 5G Systems

**摘要:** 极化码是第一种达到信道容量的构造性编码,已列入 5G 移动通信的控制信道编码标准,这是信道编码领域近年来的重大突破。旨在阐述极化码的基本原理与 5G 中的应用。基于信道极化观点,分析了极化码蝶形编码结构与基本的串行抵消译码算法,提出了级联极化编码结构,并归纳了高性能译码算法的特点。另外,还深入分析了 5G 移动通信中极化码设计的基本思想,概括了极化码实用编码的 3 种方式:凿孔、缩短与重复。最后,指出极化信息处理是未来通信系统优化的新型方法。

**关键词:** 极化码;信道极化;串行抵消译码;串行抵消列表译码;串行抵消堆栈译码;极化信息处理

**Abstract:** Polar code is the first error control code achieving the channel capacity and has been accepted as the coding scheme of the control channels of the 5G wireless communication systems, which is the great breakthrough of the channel coding field in recent years. In this paper, the primary principle of the polar code and the application in 5G systems are surveyed. First, based on the viewpoint of channel polarization, the butterfly structure of polar coding and successive cancellation decoding are analyzed. Then, the concatenated polar code is proposed and the characteristics of high-performance decoding algorithms are summarized. Furthermore, the design of polar codes in 5G systems is addressed and three practical coding schemes of polar codes are overviewed, which is puncturing, shortening and repetition. In the end, it is pointed out that polar coded information processing will become the new diagram of the communication system optimization in the future.

**Key words:** polar code; channel polarization; successive cancellation decoding; successive cancellation list decoding; successive cancellation stack decoding; polar coded information processing

1948 年,信息论创始人 C. E. Shannon 在经典文献[1]中,提出了著名的信道编码定理。70 年来,构造逼近信道容量的编码是信道编码理论的中心目标。近 20 年来,虽然以 Turbo 与低密

度校验码(LDPC)为代表的信道编码具有优越的纠错性能,但难以从理论上证明这些码渐近可达信道容量。2009 年,土耳其学者 E. Arıkan 在文献[2]中提出了极化码的设计思想,首次以构



牛凯/NIU Kai

(北京邮电大学泛网无线通信教育部重点实验室,北京 100876)  
(The Key Laboratory of Universal Wireless Communications, Ministry of Education, Beijing University of Posts and Telecommunications (BUPT), Beijing 100876, China)

DOI: 10.12142/ZTETJ.201901004  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190129.1151.006.html>

收稿日期: 2018-12-20  
网络出版日期: 2019-01-29

造性方法证明信道容量渐近可达。由于在编码理论方面的杰出贡献,该论文获得了2010年电子和电气工程师协会(IEEE)信息论分会的最佳论文奖,引起了信息论与编码学术界的极大关注。

极化码发明近10年来,成为信道编码领域的热门研究方向,其理论基础已经初步建立,人们对极化码的渐近性能有了深入理解。特别是2016年底,极化码入选5G移动通信的控制信道编码候选方案,并最终写入5G标准<sup>[3]</sup>,极大推动了极化码的应用研究。

## 1 极化码原理

本节我们将详细介绍极化码的基本原理,包括信道极化原理、极化码构造算法以及极化码的基本译码算法与增强型译码算法。

### 1.1 信道极化与编码

极化码的构造依赖于信道极化现象,我们首先介绍信道极化的基本原理,然后概述极化码的编码过程。

#### (1) 信道极化。

所谓信道极化,最早由 E. Arkan 引入<sup>[2]</sup>,是指将1组可靠性相同的二进制对称输入离散无记忆信道(B-DMC)采用递推编码的方法,变换为1组有相关性的、可靠性各不相同的极化子信道的过程,随着码长(即信道数目)的增加,这些子信道呈现两

极分化现象。图1给出了二元删余信道(BEC)的信道极化演进示例。

令 B-DMC 信道转移概率为  $W(y|x)$ , 则信道互信息与可靠性度量(Bhattacharyya 参数,简称巴氏参数)定义如公式(1):

$$I(W) = \sum_{y \in Y} \sum_{x \in X} \frac{1}{2} W(y|x) \log \frac{W(y|x)}{\frac{1}{2} W(y|0) + \frac{1}{2} W(y|1)}, \quad (1)$$

$$Z(W) = \sum_{y \in Y} \sqrt{W(y|0)W(y|1)}. \quad (2)$$

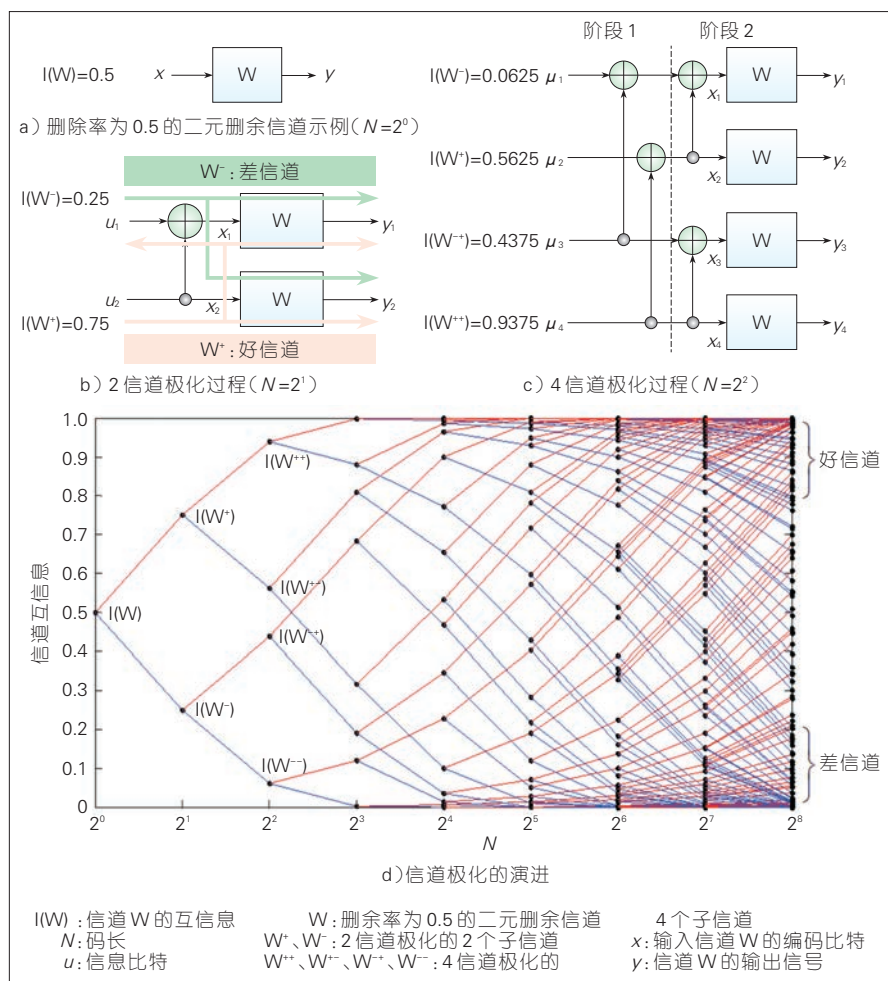
图1 a)给出了删余率为0.5的 BEC 信道的映射关系  $W: X \in \{0,1\} \rightarrow Y$ , 其信道互信息

为  $I(W)=0.5$ , 巴氏参数  $Z(W)=0.5$ 。

图1 b)是2信道极化过程,  $u_1, u_2 \in \{0,1\}$  是输入信道的两比特,  $x_1, x_2 \in \{0,1\}$  是经过模2加编码后的两比特,分别送入信道后得到  $y_1, y_2 \in Y$  2个输出信号。对应的编码过程可以表示为:

$$(x_1, x_2) = (u_1, u_2) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = (u_1, u_2) F. \quad (3)$$

通过矩阵  $F$  的极化操作,将一对独立信道  $(W, W)$  变换为2个相关子信道  $(W^-, W^+)$ 。其中,  $W^-: X \rightarrow Y^2$ ,  $W^+: X \rightarrow Y^2 \times X$ , 其信道输入输出关系分别如图1 b)



▲ 图1 信道极化示例



中绿线和粉线所示。这 2 个子信道的信道互信息与可靠度量满足公式(4)的关系:

$$\begin{cases} I(W^-) \leq I(W) \leq I(W^+) \\ Z(W^-) \geq Z(W) \geq Z(W^+) \end{cases} \quad (4)$$

由于  $I(W^-) = 0.25 < I(W) = 0.75$ , 这 2 个子信道产生了分化,  $W^+$  是好信道,  $W^-$  是差信道。

上述编码过程可以推广到 4 信道极化, 如图 1 c) 所示。此时, 每 2 个  $W^-$  信道极化为  $W^{--}$  与  $W^{-+}$  2 个信道, 每 2 个  $W^+$  信道极化为  $W^{+-}$  与  $W^{++}$  2 个信道。这样原来可靠性相同的 4 个独立信道变换为可靠性差异更大的 4 个极化信道。

信道极化变换可以递推应用到  $N=2^n$  个信道, 给定信源序列  $U_1^N$  与接收序列  $Y_1^N$ , 序列互信息可以分解为多个子信道互信息之和, 即满足公式(5)中的关系:

$$I(U_1^N; Y_1^N) = \sum_{i=1}^N I(U_i; Y_1^N | U_1^{i-1}) = \sum_{i=1}^N I(U_i; Y_1^N U_1^{i-1}), \quad (5)$$

其中,  $I(U_i; Y_1^N U_1^{i-1})$  是第  $i$  个极化子信道的互信息, 相应的信道转移概率为  $W_N^{(i)}(Y_1^N U_1^{i-1} | U_i)$ 。这就是信道极化分解原理, 其本质是通过编码约束关系, 引入信道相关性, 从而导致各个子信道的可靠性或容量差异。图 1 d) 给出了码长  $N=2^0 \sim 2^8$  时, 极化子信道互信息的演进趋势。其中, 每个节点的上分支表示极化变换后相对好的信道(红线标注),

下分支表示相对差的信道(蓝线标注)。显然, 随着码长增长, 好信道集聚到右上角(互信息趋于 1), 差信道集聚到右下角(互信息趋于 0)。

E. Arkan 证明了当信道数目充分大时, 极化信道的互信息完全两极分化为无噪声的好信道(互信息趋于 1)与完全噪声的差信道(互信息趋于 0), 并且好信道占总信道的比例趋于原始 B-DMC 信道  $W$  的容量  $I(W)$ , 而差信道比例趋于  $1-I(W)$  [2]。

#### (2) 极化编码。

极化码有 2 种基本编码结构, 即非系统码与系统码。下面我们简述各自的结构特点。

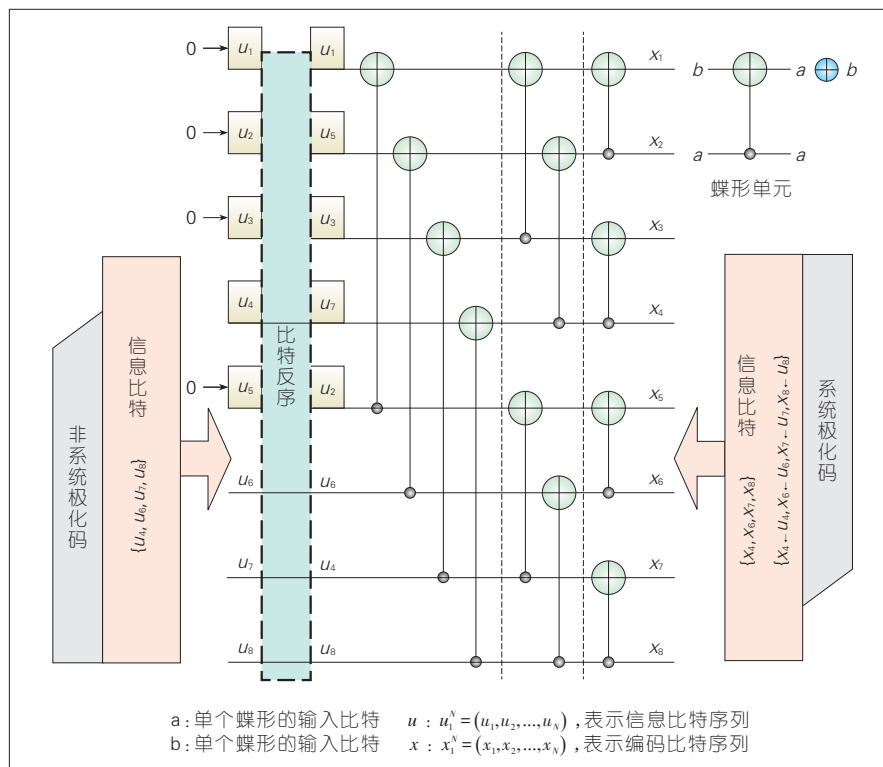
首先, 根据信道极化的递推过程, 可以得到非系统极化码的

编码结构。令  $u_1^N = (u_1, u_2, \dots, u_N)$  表示信息比特序列,  $x_1^N = (x_1, x_2, \dots, x_N)$  表示编码比特序列, E. Arkan 证明 [2] 编码满足公式(6):

$$x_1^N = u_1^N G_N, \quad (6)$$

其中, 编码生成矩阵  $G_N = B_N F^{\otimes n}$ ,  $B_N$  是排序矩阵, 完成比特反序操作,  $F^{\otimes n}$  表示矩阵  $F$  进行  $n$  次 Kronecker 积操作。

图 2 给出了码长  $N=8$ , 码率  $R=0.5$  的极化码编码器的示例。由图 2 可知, 对于非系统极化码, 根据巴氏参数选择可靠性高的  $\{u_4, u_6, u_7, u_8\}$  作为信息比特, 信息位长度为 4, 而可靠性较差的  $\{u_1, u_2, u_3, u_5\}$  作为固定比特, 取值为 0。经过 3 级蝶形运算, 可以得到编码比特序列



▲ 图 2 码长  $N=8$ 、码率  $R=1/2$  的极化码编码器示例

$x_1^8$ 。对于系统极化码,则需要将信息位承载在  $\{x_4, x_6, x_7, x_8\}$ , 对应的编码器左侧输入(信源侧)比特则通过代数运算<sup>[4]</sup>确定取值。由于采用蝶形结构编码,极化码的编码复杂度则可表示为  $O(N \log N)$ <sup>[2]</sup>。

### (3) 实用化极化编码。

笔者在文献[5]中提出了循环冗余校验(CRC)-Polar 级联方案,如图3所示。由  $k$  个信息比特组成的序列首先送入CRC编码器,级联  $m$  个CRC校验比特后送入极化码编码器,产生  $N$  比特码字。这种级联编码方案,以CRC编码作为外码,极化码作为内码,具有显著的性能增益,目前已经成为极化码的主流编码方案。

由于极化码原始码长限定为2的幂次,即  $N=2^n$ ,而实际通信系统往往要求任意码长编码。为了满足这一要求,需要设计极化码的速率适配方案,主要包括凿孔、缩短、重复3种操作。假定速率适配后的码长为

$M < N$ ,则编码器需要删减  $N-M$  个编码比特。对于凿孔操作,这些删减的比特可以任意取值,而译码器并不确定它们的取值,因此相应的对数似然比(LLR)为0。而对于缩短操作,这些删减比特为固定取值(假设为0),译码器也知道其取值,因此相应的LLR取值为 $\infty$ 。对于重复操作,译码器需要将重复发送比特对应的LLR叠加。

笔者在文献[6]中提出了准均匀凿孔(QUP)适配方案,并进一步在文献[7]中提出了反向准均匀缩短(RQUS)适配方案。其中,QUP是凿孔方案,适用于低码率的情况;RQUS是缩短方案,适用于高码率的情况。可以证明,QUP与RQUS方案是理论最优的速率适配方案<sup>[7]</sup>,并且RQUS与文献[8]中提到的缩短方案等价。

## 1.2 极化码构造

极化码构造算法的目的是精确计算各个子信道的互信息

或可靠性,然后从大到小排序,选择其中好的子信道集合承载信息比特;因此,构造算法是极化码编码的关键。

E. Arkan 最早提出基于巴氏参数的构造算法<sup>[2]</sup>。假定初始信道的巴氏参数为  $Z(W)$ ,则从  $N$  扩展到  $2N$  个极化信道的迭代计算过程如公式(7):

$$\begin{cases} Z(W_{2N}^{(2i-1)}) = 2Z(W_N^{(i)}) - Z(W_N^{(i)})^2 \\ Z(W_{2N}^{(2i)}) = Z(W_N^{(i)})^2 \end{cases} \quad (7)$$

这种构造算法复杂度较低,但只适用于BEC信道,对于其他信道,例如二元对称信道(BSC)、加性白噪声信道(AWGN)等,该方法并非最优。

Mori 基于密度进化(DE)方法,得到了BSC、AWGN信道下最优的子信道选择准则<sup>[9]</sup>,但由于涉及到变量与校验节点比特LLR概率分布计算,计算复杂度很高,限制了其应用。更好的方法是I. Tal 与 A. Vardy 提出的迭代算法<sup>[10]</sup>,通过引入极化子信道的上下界近似,该方法能以中等复杂度保证较高的计算精度,但码长很长时,其计算复杂度也会变大。

P. Trifonov 所提出的高斯近似(GA)算法<sup>[11]</sup>是目前较流行的构造方法。给定AWGN信道的接收信号模型为  $y_i = s_i + n_i, i = 1, 2, \dots, N$ , 噪声功率为  $\sigma^2$ ,则接收比特的LLR  $L(y_i) \sim N\left(\frac{2}{\sigma^2}, \frac{4}{\sigma^2}\right)$  服从高斯分布。信道极化的LLR均值迭代

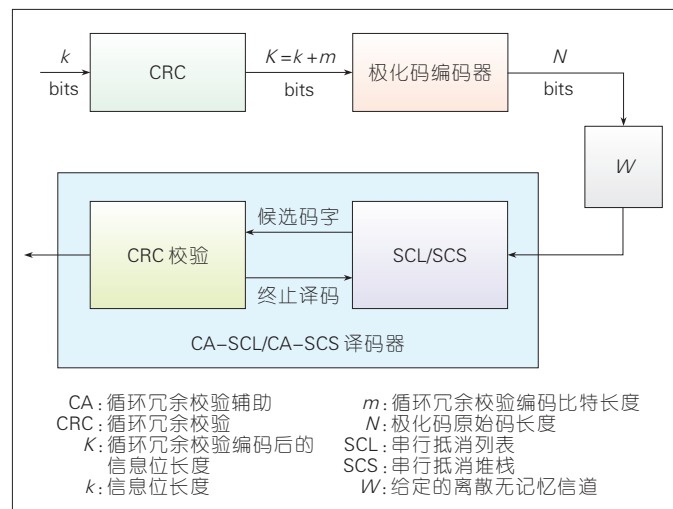


图3 CRC-Polar 级联编译码系统结构

公式为:

$$\begin{cases} E(L_{2N}^{(2i-1)}) = \phi^{-1} \left\{ 1 - \left[ 1 - \phi(E(L_N^{(i)})) \right]^2 \right\} \\ E(L_{2N}^{(2i)}) = 2E(L_N^{(i)}) \end{cases}, \quad (8)$$

其中,  $E(\cdot)$  表示数学期望,  $E(L_1^{(0)}) = \frac{2}{\sigma^2}$ 。公式(8)中的积分函数  $\phi(z)$  定义为:

$$\phi(z) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi z}} \int_{-\infty}^{\infty} \tanh\left(\frac{u}{2}\right) e^{-\frac{(u-z)^2}{4z}} du, & z > 0 \\ 1, & z = 0 \end{cases}, \quad (9)$$

其中,  $\tanh(u/2) = \frac{e^{u/2} - e^{-u/2}}{e^{u/2} + e^{-u/2}}$  表示

双曲正切函数,  $z$  表示 LLR 均值。函数  $\phi(z)$  涉及到复杂的积分运算,一般可以用分段函数近似表示:

$$\varphi_{\text{GA}}(z) = \begin{cases} e^{-0.4527z^{0.86} + 0.0218} & 0 < z < 10 \\ \sqrt{\frac{\pi}{x}} e^{-\frac{z}{x}} \left(1 - \frac{10}{7z}\right) & z \geq 10 \end{cases}, \quad (10)$$

上述 GA 构造算法的计算复杂度为  $O(N \log N)$ , 在中短码长下可以获得较高的计算精度。但这种近似在码长较长时,存在计算误差,我们在文献[12]提出了改进的 GA 算法,满足长码条件下高精度构造的要求。

前述极化码的构造算法,有一个共同的局限,即编码构造依赖于信道条件。最近,不依赖于信道条件的通用构造成为极化码的研究热点。其中,文献[13]提出的部分序构造以及文献[14]提出的极化度量(PW)构造算法具有代表性。假设第  $i$  个子信道序号对应的二进制展开向量为:  $i \rightarrow (b_n, b_{n-1}, \dots, b_1)$ , 则极化度

量计算公式如公式(11):

$$PW_N^{(i)} = \sum_{j=1}^n b_j 2^{j/4}. \quad (11)$$

极化度量越大,说明子信道可靠性越高;因此,将极化度量从大到小排序,选取大度量对应的子信道承载信息比特。基于极化度量构造的极化码,性能与 GA 构造的极化码接近,且度量计算不依赖于信道条件,这种构造方法具有重要的实用价值。

### 1.3 极化码译码算法

#### (1) 串行抵消(SC)译码算法。

对于极化码, E. Arkan 的另一个重要贡献是提出了串行抵消 SC 译码算法<sup>[2]</sup>。SC 译码的基本思想是在 Trellis 上进行软信息与硬判决信息的迭代计算。

给定码长  $N=2^n$  与极化阶数  $n$ , 则 Trellis 由  $n$  级蝶形节点构成。其变量节点的硬判决信息定义为  $s_{i,j}$ , 其中  $1 \leq i \leq n+1$ ,  $1 \leq j \leq N$  分别表示节点在 Trellis 上的行列序号, 而软判决信息定位为相应的 LLR, 即  $L_{i,j} = L(s_{i,j})$ 。图 4 给出了  $N=4$  的

极化码 Trellis 示例。如图 4 所示, Trellis 右侧对应来自于信道的

的 LLR 信息  $L_{n+1,j} = \log \frac{P(y_j|1)}{P(y_j|0)}$ ,

而左侧对应信息比特的 LLR 信息  $L_{1,j} = L(\hat{u}_j)$  以及判决比特信息  $s_{1,j} = \hat{u}_j$ 。这样, 基于蝶形结构中的变量/校验节点约束关系, 软信息从右向左计算与传递, 而硬信息从左向右计算与传递。

软消息迭代计算公式为:

$$L_{i,j} = \begin{cases} 2 \tanh^{-1} \left[ \tanh\left(\frac{L_{i+1,j}}{2}\right) \cdot \tanh\left(\frac{L_{i+1,j+2^{i-1}}}{2}\right) \right] \left\lfloor \frac{j-1}{2^{i-1}} \right\rfloor \bmod 2 = 0 \\ (1 - 2s_{i,j-2^{i-1}})(L_{i+1,j-2^{i-1}}) + L_{i+1,j}, \text{ otherwise} \end{cases}, \quad (12)$$

其中,  $i=1, 2, \dots, n$ ,  $j=1, 2, \dots, N$ ,  $\tanh(\cdot)$  是双曲正切函数,  $\lfloor \cdot \rfloor$  是下取整函数。

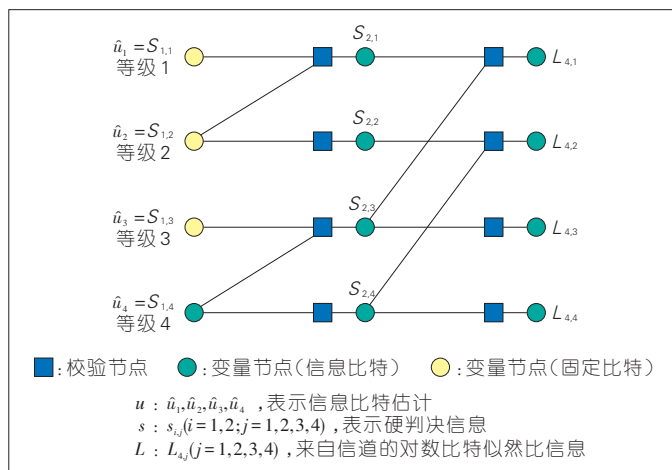
上述计算与 LDPC 码的 BP 迭代译码基本公式类似, 都是在校验与变量节点分别进行软信息计算与更新。

硬消息迭代计算公式为:

$$s_{i+1,j} = \begin{cases} s_{i,j} \oplus s_{i,j+2^{i-1}}, \left\lfloor \frac{j-1}{2^{i-1}} \right\rfloor \bmod 2 = 0 \\ s_{i,j}, \text{ otherwise} \end{cases}, \quad (13)$$

其中,  $\oplus$  是模二加操作。

图 4  
码长  $N=4$  的极化码  
Trellis 示例



当软信息递推到 Trellis 的左侧时, 比特判决准则为:

$$\hat{u}_i = \begin{cases} 1, & L_{1,i} \geq 0 \\ 0, & L_{1,i} < 0 \text{ 或者 } u_i \text{ 是固定比特} \end{cases} \quad (14)$$

SC 算法也可以看作是在码树上进行逐级判决搜索路径的过程。也就是说, 从树根开始, 对发送比特进行逐级判决译码, 先判决的比特作为可靠信息辅助后续比特的判决, 最终得到一条译码路径。文献[2]证明极化码的 SC 译码算法复杂度非常低, 为  $O(N \log N)$ 。

### (2) 增强型 SC 译码算法。

在有限码长下, 基于 SC 译码的极化码性能较差, 远不如 LDPC/Turbo 码。为了提高极化码有限码长的性能, 人们提出了多项高性能的 SC 改进算法。笔者<sup>[15]</sup>与 I. Tal 及 A. Vardy 同时提出了列表 SC 算法<sup>[16]</sup>, 将广度优先搜索策略引入码树搜索机制, 每次译码判决保留一个很小的幸存路径列表, 最终从表中选择似然概率最大的路径作为判决路径。给定列表长度  $L$ , SCL 算法的复杂度为  $O(LN \log N)$ , 其性能可以逼近最大似然 (ML) 译码性能。

图 5 给出了  $L=2$  的 SCL 译码算法示例。由图可知, SCL 算法保留了 2 条幸存路径, 译码器最终从 2 条候选路径中选择译码结果。

另外, 笔者在文献[17]中提出堆栈 SC 算法, 将深度优先搜索策略引入码树搜索中, 由于引

入堆栈存储机制, 可以有效减少译码路径的重复搜索, 极大降低了译码算法复杂度。高信噪比条件下, SCS 算法的复杂度趋近于 SC 算法, 远低于 SCL 算法, 且其性能也能够逼近 ML 译码性能。笔者还在文献[18]中作者提出了混合 SC 算法, 组合串行抵消列表 (SCL) 与串行抵消堆栈 (SCS) 算法优势, 达到时间与空间复杂度更好折中。

图 6 给出了 SCS 译码算法示例。由图可知, 译码器在码树上通过深度优先的方式, 搜索候选路径, 按照从大到小的顺序将候选路径压入堆栈, 每次从栈顶扩展幸存路径直至叶节点, 最终得到译码结果。

进一步地, 笔者在文献[5]中

作者提出 CRC 辅助的 SCL/SCS 译码算法 (CA-SCL/SCS)。如图 3 所示, SCL/SCS 算法输出的候选码字, 送入 CRC 校验模块, 只有通过 CRC 校验的码字才作为最终译码结果。由于有 CRC 校验提供的先验信息, 极大增强了译码性能。进一步地, 文献[19]中作者还提出了自适应 CA-SCL 算法, 可以在算法复杂度与性能之间达到较好折中。目前 SCL 算法已经成为极化码高性能译码的主流算法, 文献[20-21]中作者深入讨论了 SCL 译码器的硬件架构设计。

### (3) 其他译码算法。

极化码也可以采用置信传播 (BP) 译码算法, 文献[22]中作者最早研究了 BP 算法调度机制

图 5  
串行抵消列表译码  
算法示例

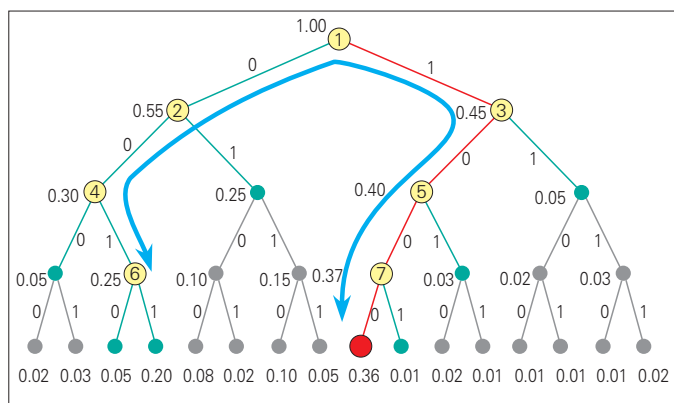
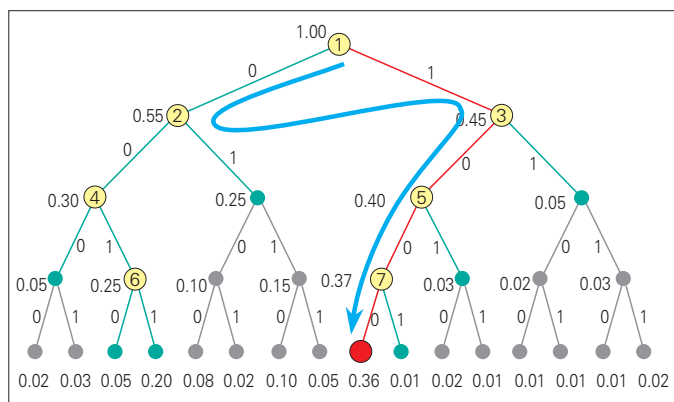


图 6  
串行抵消堆栈译码  
算法示例





的优化。另外,对于短码极化码,笔者在文献[23]中提出了低复杂度的球译码算法,能够达到 ML 译码性能,也具有一定的实用价值。

## 2 5G 中的极化码

本节介绍 5G 标准中的极化码设计的基本流程,并且给出了 Polar、LDPC 与 Turbo 码的性能比较结果。

### 2.1 5G 新空口(NR)标准中的极化编码

5G NR 系统中,极化码作为控制信道的差错编码<sup>[3]</sup>。具体而言,对于下行链路,物理下行控制信道(PDCCH)的下行控制信息(DCI)与物理下行广播信道(PBCH)都采用极化码编码;对于上行链路,物理上行控制信道(PUCCH)与物理上行共享信道(PUSCH)的上行控制信息(UCI)(12 bit 以上)都采用极化码编码。

具体的编码流程如图 7 所

示,包括 6 个步骤:首先进行 CRC 编码,然后经过信息比特交织,进行子信道映射后,送入极化码编码器,编码输出的码字先进行子块交织,再进行速率适配,最后进行信道交织,得到编码码字。

#### (1) CRC 编码器。

5G 标准所采用的用于 Polar 码的 CRC 有 3 种,其生成多项式如公式(15):

$$\begin{cases} g_6(x) = x^6 + x^5 + 1 \\ g_{11}(x) = x^{11} + x^{10} + x^9 + x^5 + 1 \\ g_{24}(x) = x^{24} + x^{23} + x^{21} + x^{20} + x^{17} + x^{15} + x^{13} + x^{12} + x^8 + x^4 + x^2 + x + 1 \end{cases} \quad (15)$$

其中,生成多项式  $g_{24}(x)$  用于 PBCH 信道与 PDCCH 信道,而  $g_6(x)$  与  $g_{11}(x)$  用于 UCI 编码。当  $g_{24}(x)$  应用于 DCI 编码时,产生的 CRC 比特中的最后 16 位,需要用 16 bit 的无线网络临时识别码(RNTI)进行扰码。

#### (2) 交织器。

5G NR 的极化码编码主要有 3 种交织器:信息比特交织、子块交织与信道交织。下面简

述各自的基本功能。

- 信息比特交织主要对 CRC 编码的数据比特进行置乱。这种交织只对下行 PBCH 信道或 PDCCH 信道的 DCI 有效,而上行链路不采用。信息比特交织的设计思想,是将 CRC 比特分布到整个信息比特块中,每个校验比特与其约束信息比特相邻,从而方便 SCL 译码算法提前终止,降低广播信道或 DCI 盲检的算法复杂度。

- 子块交织是将  $N$  比特码块分割为 32 个子块,每块长度为  $B = N/32$  比特,根据 5G NR 协议定义的映射表,得到置乱比特序列  $\{c_j\}$ 。

- 信道交织的目的是对抗 Doppler 效应引起的时变衰落,并且用于提高比特交织编码调制(BIPCM)<sup>[24]</sup>的系统性能。这种交织主要应用于 PUCCH 与 PUSCH 的 UCI,而下行链路不采用。5G NR 中采用了三角形交织结构,既保证了数据读写的高并行度,又具有较好的灵活性。

#### (3) 子信道映射。

5G NR 中的极化码采用了与信道条件无关的子信道映射方案,标准中给出了最大长度为个子信道的可靠性排序表。给定信息长度  $K$ ,可以从排序表中选择可靠性排序高的  $K$  个子信道承载信息比特。为了实现方便,子信道映射满足通用性,即原始码长经过速率适配得到任意码长,剩余子信道的可靠性排序保持不变,即排序关系与码

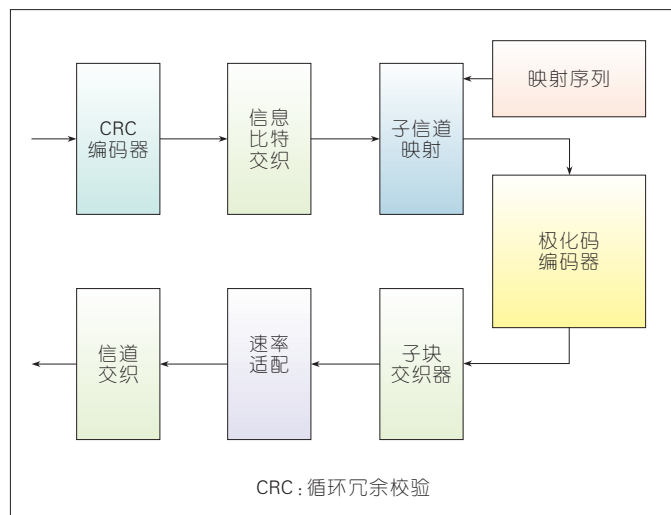


图 7  
5G 新空口控制信道  
编码流程

长无关。

#### (4)极化码编码器。

5G NR 的极化码采用了简化编码方式<sup>[3]</sup>,即:

$$x_i^N = u_i^N F^{N/2^n}, \quad (16)$$

其中,对于下行信道,  $n \leq 9$ ; 对于上行信道,  $n \leq 10$ 。对比公式(6)可知,其编码过程直接进行 Hadamard 变换,不必再进行比特反序操作。E. Arkan 在文献[2]中证明这 2 种编码形式是等价的。公式(16)的编码过程更简单,但在译码端需要调整接收信号的顺序。

#### (5)速率适配。

5G NR 中的速率适配有 3 种模式:凿孔、缩短与重复。下面简述 3 种模式的适用条件。

- 如果最终码长不大于编码码长即  $M \leq N$ , 且编码码率  $R \leq 7/16$ , 即低码率条件下,则采

用凿孔方式。删除子块交织后序列  $\{c_j\}$  的开头  $U=N-M$  个比特, 即只传送序号为  $e_i=c_{j+U}, i=0, \dots, M$  的比特。由于 5G NR 编码不进行比特反序操作,这种凿孔方式本质上就是 QUP 速率适配算法<sup>[6]</sup>。

- 如果  $M \leq N$ , 且编码码率  $R > 7/16$ , 即高码率条件,则采用缩短方式。子块交织后序列  $\{c_j\}$  的末尾  $U=N-M$  个比特不发送, 只传送序号为  $e_i=c_j, i=0, \dots, M$  的比特。类似的,这种凿孔方式等价于 RQUS 速率适配算法<sup>[7-8]</sup>。

- 如果  $M > N$ , 则采用重复方式, 即传送序号为  $e_i=c_{j \bmod N}, i=0, \dots, M$  的比特。

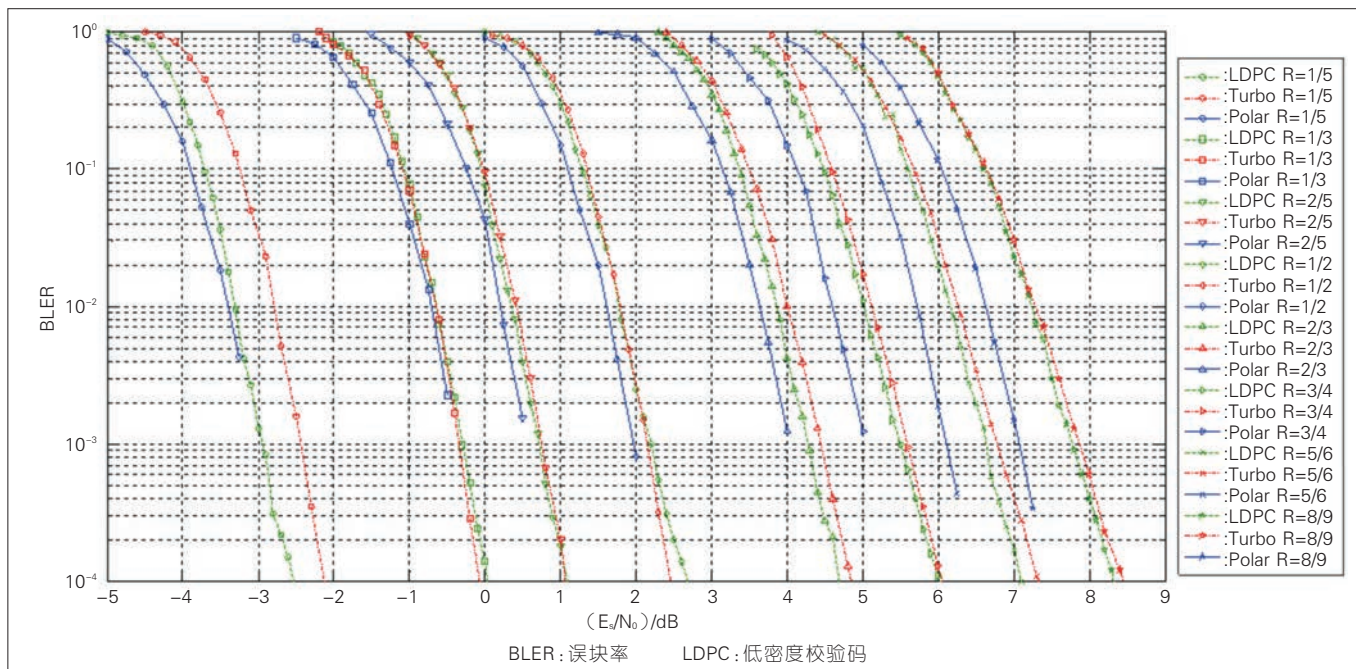
## 2.2 极化码性能比较

图 8 给出了 5G 移动通信系统的 3 种候选编码: Turbo、LDPC

与 Polar 码在 AWGN 信道下的误块率(BLER)性能比较。

其中,3 种编码的信息位长度  $K=400$ , 码率范围  $R=1/5 \sim 8/9$ 。Turbo 码采用 4G 长期演进(LTE)标准配置,采用 Log-MAP 译码算法,8 次迭代。LDPC 码采用 Qualcomm 公司的 5G 编码提案<sup>[25]</sup>,采用 BP 译码算法,50 次迭代。Polar 码采用 5G 标准配置<sup>[3]</sup>,采用 CA-SCL 译码算法,列表大小为 32。由图 8 可见,低码率条件下  $R=1/5 \sim 1/2$ , 极化码与 Turbo/LDPC 码具有类似或稍好的性能;而在高码率条件下,  $R=2/3 \sim 8/9$ , 相对于后 2 种码,极化码则具有显著的性能增益。

笔者在文献[26]指出,在相同的码长码率参数配置下,达到相同的误码率性能,与 Turbo 码对数最大后验概率译码



▲ 图 8 3 种 5G 移动通信候选编码(Turbo、LDPC 与 Polar)的性能比较(信息位长度  $K=400$ )

(Log\_MAP)算法相比, Polar 码 SCL 译码算法复杂度降低 5~10 倍, 并且没有错误平台现象; 与 LDPC 码标准 BP 译码算法相比, Polar 码 SCL 译码算法复杂度降低 3~5 倍。由此可见, 对于中短码长, 极化码具有性能与复杂度的双重优势。

### 3 极化信息处理

#### 3.1 极化码的理论本质

极化码的理论性主要关注信道极化行为的理解与分析, 包括 BLER 与子信道收敛速度。E. Arkan 基于 SC 算法给出了误码块率简洁的上界<sup>[2]</sup>, 并且利用鞅与半鞅理论, 严格证明了子信道的收敛行为, 奠定了信道极化码的基本理论。E. Arkan 在文献[27]首次证明采用  $2 \times 2$  核矩阵  $F$ , 极化码渐近 ( $N \rightarrow \infty$ ) 差错性能  $P_B(N) < 2^{-N^\beta}$ , 其中误差指数  $\beta < 1/2$ , 也就是说, 极化码的差错概率随着码长的平方根指数下降。Korada 等人进一步证明, 如果推广到  $l \times l$  核矩阵, 则渐近性能  $P_B(N) < 2^{-N^{E_c(G)}}$ , 其中  $E_c(G)$  是生成矩阵  $G$  对应的差错指数<sup>[28]</sup>, 极限为 1。

Shannon 在证明信道编码定理<sup>[1]</sup>时, 采用了如下假设:

- (1) 码长充分长, 即  $N \rightarrow \infty$ ;
- (2) 采用随机编码方法;
- (3) 基于信源信道联合渐近等分割 (JAEP) 特性, 采用联合典型序列译码方法。

这 3 条假设对于设计逼近

信道容量的信道编码具有重要的启发性。长期以来, 人们主要关注第 2 个假设, 通过构造方法模拟随机编码, 例如 Turbo 码或 LDPC 码, 都具有一定的随机性, 能够在码长充分长时逼近信道容量。但第 3 个假设更重要, 应用 JAEP 特性, 采用联合典型序列译码是信道编码定理证明的关键步骤。

对于信道极化的理论理解, 笔者在文献[26]中指出, 极化变换实际上是 JAEP 特性的构造性示例。Turbo 码与 LDPC 码虽然模拟了随机编码的行为, 但难以模拟 JAEP 特性; 而在极化编码中, 极化变换所得到的好信道可以看作是联合典型映射, 这种方法更加符合 Shannon 原始证明的基本思路。极化码渐近差错率随码长指数下降。这样极化码与随机编码具有一致的渐近差错性能, 相当于给出了信道编码定理<sup>[1]</sup>的构造性证明。

#### 3.2 极化信息处理

作为信道容量可达的新型编码, 极化码的优势集中体现在 3 个方面:

- (1) 高可靠性。

极化码可以严格证明没有错误平台, 这一点是极化码相比于 Turbo/LDPC 码最重要的性能优势。同时, 在中短码长 (100 ~ 2 000 bit) 下, 采用 CA-SCL 译码算法的极化码性能要显著优于 Turbo/LDPC 码。由于这 2 方面的优势, 极化码能够达到更低的差

错概率, 非常适合于高可靠低时延的通信传输需求。

- (2) 高效性。

已有研究表明, 极化编码调制的性能可以超过 Turbo/LDPC 编码调制。针对极化编码调制的联合优化, 可以在高信噪比条件下逼近信道容量极限, 极大提升频谱效率, 非常适合于高频谱效率传输需求。

- (3) 低复杂度。

极化码的代表性译码算法, 如 SC、SCL/SCS、BP 译码算法, 都可以用低复杂度方式实现。如果能够在译码性能与算法复杂度之间优化设计, 将获得复杂度与可靠性的双重增益, 具有重要的工程实用价值。

自从极化码发明以来, 大量的理论分析表明, 信道极化不仅存在于编码系统中, 也是各种通信系统普遍存在的现象, 例如: 多输入多输出 (MIMO) 系统、多址接入信道 (MAC)、中继系统。一般来说, 我们可以把基于信道极化的通信系统称为极化信息处理系统, 图 9 给出了其详细系统框架。

由图 9 可知, 极化信息处理系统的发射机基于信道极化观点, 对 MIMO/MAC/Relay 等传输信道进行分解, 并与多个极化码编码器进行映射匹配。在接收端, 可以将 MIMO/MAC/Relay 检测算法与多个极化码译码器构成串行抵消结构的整体接收机。

在理论性能上, 极化信息处理系统, 能够逼近相应的信道容



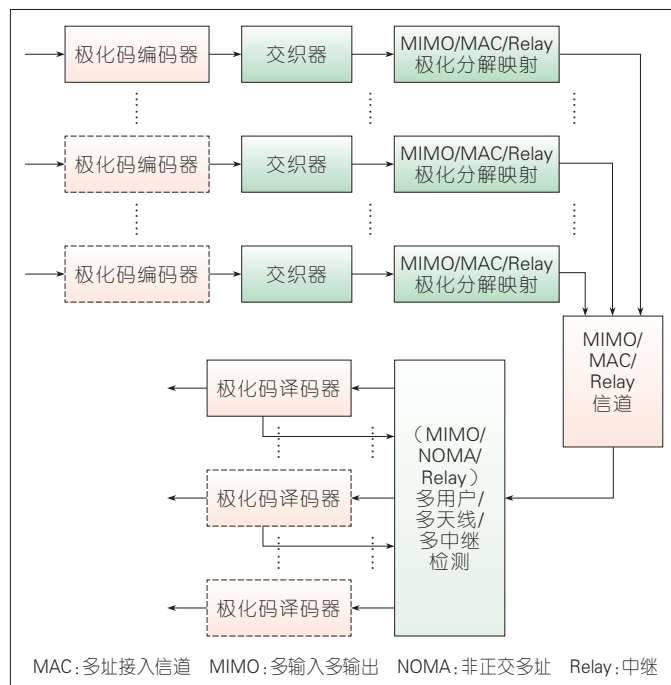


图9  
极化信息处理系统  
框架

量,具有同等或者更好的渐近性能;在实用化方面,极化信息处理系统,由于采用了不需要迭代的类似SC结构的接收机,能够获得复杂度与性能的双重优势。

笔者在文献[29-30]中提出了极化编码的非正交多址(NOMA)与极化编码MIMO系统。与同等条件的Turbo编码系统相比,这些极化编码系统具有显著的性能增益。

#### 4 结束语

本文首先回顾与总结了极化码的基本编译码原理,包括极化编码、构造算法、译码算法,接着介绍了5G移动通信标准中极化码设计的基本思想,最后阐述了极化码的理论优势,并展望未来研究方向。极化码的设计完美体现了JAEP的信息论思

想,可以看作是通信系统整体优化的“大道”与“太极”。极化信息处理系统暗合了中国道家“大道至简、太极混一”的系统论思想,将成为未来通信系统优化的新方向。

#### 参考文献

- [1] SHANNON C E. A Mathematical Theory of Communication [J]. The Bell System Technical Journal, 1948, 27(3):379-423, 623-656. DOI: 10.1002/j.1538-7305.1948.tb01338.x
- [2] ARIKAN E. Channel Polarization: A Method for Constructing Capacity-Achieving Codes for Symmetric Binary-Input Memoryless Channels [J]. IEEE Transaction Information Theory, 2009, 55(7): 3051-3073. DOI: 10.1109/TIT.2009.2021379
- [3] 3GPP. Multiplexing and Channel Coding: 3GPP 38.212 V.15.1.0 [S]. 2018
- [4] ARIKAN E. Systematic Polar Coding [J]. IEEE Communications Letters, 2011, 15(8): 860-862. DOI: 10.1109/LCOMM.2011.061611.110862
- [5] NIU K, CHEN K. CRC-Aided Decoding of Polar Codes [J]. IEEE Communications Letters, 2012, 16(10): 1668-1671. DOI: 10.1109/LCOMM.2012.090312.121501
- [6] NIU K, CHEN K, LIN J R. Beyond Turbo Codes: Rate-Compatible Punctured Polar Codes [C]//IEEE International Conference on Communications (ICC). USA: IEEE, 2013: 3423-3427. DOI: 10.1109/ICC.2013.6655078
- [7] NIU K, DAI J C, et al. Rate-Compatible

Punctured Polar Codes: Optimal Construction Based on Polar Spectra [EB/OL]. [2018-12-26]. <https://arxiv.org/pdf/1612.01352>

- [8] WANG R, LIU R. A Novel Puncturing Scheme for Polar Codes [J]. IEEE Communications Letters, 2014, 18(12): 2081-2084. DOI: 10.1109/LCOMM.2014.2364845
- [9] MORI R, TANAKA T. Performance of Polar Codes with the Construction Using Density Evolution [J]. IEEE Communications Letters, 2009, 13(7): 519-521. DOI: 10.1109/LCOMM.2009.090428
- [10] TAL I, VARDY A. How to Construct Polar Codes [J]. IEEE Transactions on Information Theory, 2013, 59(10):6562-6582. DOI: 10.1109/TIT.2013.2272694
- [11] TRIFONOV P. Efficient Design and Decoding of Polar Codes [J]. IEEE Transactions on Communications. 2012, 60(11): 3221-3227. DOI: 10.1109/TCOMM.2012.081512.110872
- [12] DAI J C, NIU K. Does Gaussian Approximation Work Well for the Long-Length Polar Code Construction? IEEE Access [J]. 2017, 5: 7950-7963. DOI: 10.1109/ACCESS.2017.2692241
- [13] SCHURCH C. A Partial Order for the Synthesized Channels of a Polar Code [C]//IEEE International Symposium on Information Theory (ISIT). USA: IEEE, 2016: 220-224. DOI: 10.1109/ISIT.2016.7541293
- [14] HE G N, BELFIORE J C.  $\beta$ -Expansion: A Theoretical Framework for Fast and Recursive Construction of Polar Codes [C]//IEEE GLOBECOM2017. USA: IEEE, 2017: 1-6. DOI: 10.1109/GLOCOM.2017.8254146
- [15] CHEN K, NIU K, LIN J R. List Successive Cancellation Decoding of Polar Codes, Electronics Letters [J]. 2012, 48(9): 500-501. DOI: 10.1049/el.2011.3334
- [16] TAL I, VARDY A. List Decoding of Polar Codes, IEEE Transactions on Information Theory [J]. 2015 6(5): 2213 - 2226. DOI: 10.1109/TIT.2015.2410251
- [17] NIU K, CHEN K. Stack Decoding of Polar Codes [J]. Electronics Letters, 2012, 48(12): 695-696. DOI: 10.1049/el.2012.1459
- [18] CHEN K, NIU K, LIN J R. Improved Successive Cancellation Decoding of Polar Codes [J]. IEEE Transactions on Communications, 2013, 61(8): 3100-3107. DOI: 10.1109/TCOMM.2013.070213.120789
- [19] LI B, SHEN H, TSE D. An Adaptive Successive Cancellation List Decoder for Polar Codes with Cyclic Redundancy Check [J]. IEEE Communications Letters, 2012, 16(12): 2044-2047. DOI: 10.1109/LCOMM.2012.111612.121898
- [20] LEROUX C, RAYMOND A J, SARKIS G, et al. A Semi-Parallel Successive-Cancellation Decoder for Polar Codes [J]. IEEE Transactions on Signal Processing, 2013, 61(2): 289-299. DOI: 10.1109/tsp.2012.2223693
- [21] ZHANG C, PARHI K K. Low-Latency Sequential and Overlapped Architectures for Successive Cancellation Polar Decoder [J]. IEEE Transactions on Signal Processing, 2013, 61(10): 2429-2441. DOI: 10.1109/tsp.2013.2251339

下转第 62 页



# 未来无线网络下的空中 计算技术

## Over-the-Air Computation for Future Networks

**摘要:** 空中计算提供了一种通信和计算一体化的架构,在不需恢复出每个节点个体数据的前提下,通过所有节点的并发传输,利用无线信道的叠加特性,实现目标函数在空口信道中的直接运算。空中计算可以在通信容量受限的计算场景下,减少由于先传输再计算而造成的巨大传输时延。介绍无线网络中空中计算技术的基本原理,同时针对实现计算过程的关键技术展开讨论。

**关键词:** 计算;多址接入;无线信道;未来网络

**Abstract:** Over-the-air computation provides a novel communication and computation integrated architecture. Without recovering individual node's data, it utilizes the summation property of wireless channel to compute the target function directly with all nodes' concurrent transmission. Thus, over-the-air computation can reduce the transmission delay in communication-constrained computation scenario. This paper introduces the principle of over-the-air computation, and discusses the key technologies of its implementation.

**Key words:** computation; multiple access; wireless channel; future networks



陈力/CHEN Li  
卫国/WEI Guo

(中国科学技术大学,安徽 合肥 230027)  
(University of Science and Technology of  
China, Hefei 230027, China)

### 1 空中计算的研究背景

支持海量节点的接入,是以5G为代表的未来无线网络的主要愿景之一,而接入节点的数量将继续呈现爆炸式的增长。以物联网(IoT)为例,其各种应用场景的实现都将依赖于海量的传感节点部署。在5G移动通信网络中,IoT节点的接入数目将达到亿级,其部署密度将达到每平方千米百万个<sup>[1-3]</sup>。根

据2016年思科的预测,在2021年,全球每月的移动数据流量预计将增至49 EB,比2016年增长7倍,复合年增长率为47%。海量节点的接入需求无疑对传统无线通信中的多址接入方法提出了巨大挑战。

与此同时,未来无线网络将由数据为中心向计算为中心转变。以人工智能为代表的信息处理技术的高速发展,将提供无所不在的计算及智能服务。通

DOI: 10.12142/ZTETJ.201901005  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190129.1606.008.html>

收稿日期: 2018-12-02  
网络出版日期: 2019-01-29

过这些计算和服务,可以实现对海量数据的分析及处理,例如图片识别与分类、智能机器人等泛在的人工智能服务。这意味着我们将来可能更加关注基于数据的计算结果而不是个体数据本身。而传统的“先通信再计算”的分离架构需要先恢复出海量的个体数据再进行计算,无疑会使无线信道面临巨大压力。

为了应对由于海量节点接入而导致的通信受限下的计算问题,以“通信计算一体化”为主要特征的空中计算技术,可能提供了一种解决之道。

## 2 空中计算技术的基本原理

我们首先介绍无线网络中空中计算的基本概念。考虑一个由  $K$  个节点和 1 个接入点组成的无线网络,其中节点  $k \in \{1, \dots, K\}$  的消息为  $s_k$ ,接入点的目标函数为  $f(s_1, \dots, s_K)$ 。

如图 1a) 所示,传统的通信和计算分离的架构,先通过多址接入的无线通信,在接入点端恢复出每个节点的消息  $\{s_k, \forall k\}$ ,再进行目标函数的计算。为了避免节点间干扰,每个节点都需要分配正交的无线资源。以时分多址(TDMA)系统为例,为了完成一个完整的计算,所有参与计算的节点数据必须排队等待分配的传输时隙。只有当最后 1 个节点数据到达计算中心之后,完整的计算才能完成。当节点数目增多且无线资源有限时,计算等待时延将会大大地增加,

从而造成通信受限的计算问题。例如,分别利用传统分离架构和空中计算架构来计算多节点数据平均值,其中接收节点的目标函数为  $\sum_{k=1}^K s_k / K$ 。采用通信和计算分离的架构,需要在接收节点先逐一恢复出所有的发送数据  $\{s_1, \dots, s_K\}$ ,然后再计算平均值;因此需要  $K$  次信道使用。

采用空中计算的架构,假设无线信道满足完美的信号叠加(忽略衰落和噪声),所有发送节点并发传输,接收节点直接空中计算出数据的求和结果  $\sum_{k=1}^K s_k$ ,再通过处理后得到平均值;因此仅需要 1 次信道使用就可以得到计算结果。

考虑一类具有求和结构的目标函数,即满足如下形式:

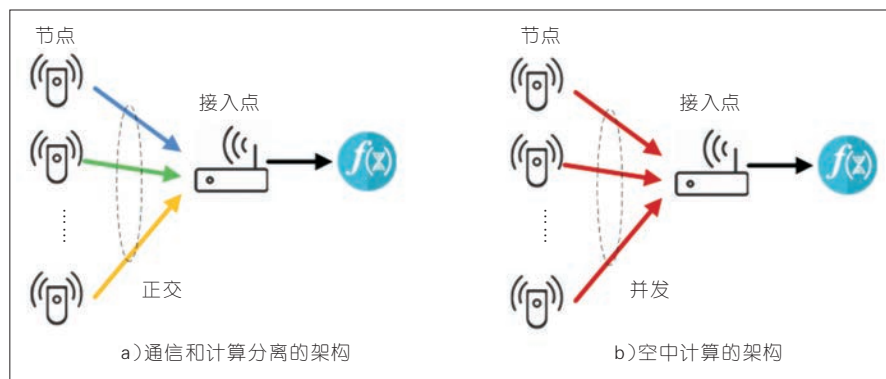
$$f(s_1, \dots, s_K) = \psi \left[ \sum_{k=1}^K \varphi_k(s_k) \right], \quad (1)$$

其中,  $\varphi_k(s_k)$  是每个节点的预处理函数,  $\psi(\cdot)$  是接入点的后处理函数。由于目标函数的求和结构,可以采用空中计算的架构实

现目标函数的计算,如图 1b) 所示。为说明问题起见,这里假设多址接入信道的信道增益都相同,且不考虑噪声的影响。每个节点对数据进行预处理后得到  $\varphi_k(s_k)$ ,通过所有节点的并发,利用无线信道的叠加特性,接入点得到  $\sum_{k=1}^K \varphi_k(s_k)$ ,再通过处理后恢复出目标函数  $f(s_1, \dots, s_K)$ 。可以看出,空中计算可以避免个体数据的收集而直接完成函数的计算,无论多少个节点,1 次信道使用就可以恢复出目标函数,从而大大降低计算延时。

公式(1)代表的是一类具有求和结构可以空中计算的目标函数,当给定不同的  $\varphi_k(\cdot)$  和  $\psi(\cdot)$  可以得到一系列常用的目标函数,如表 1 所示。当目标函数不满足公式(1)的形式时,需要利用函数近似理论,将目标方程近似拆解成满足公式(1)的函数。本文仅讨论具备公式(1)形式的目标函数。

值得注意的是,除了计算上述具有加性结构的目标函数,空中计算还可以用于计算最大最

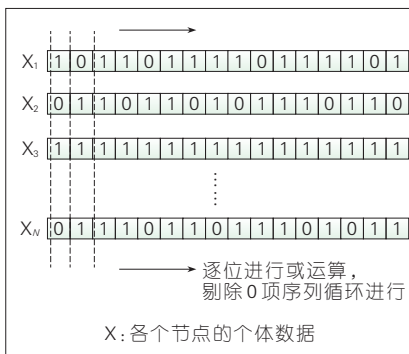


▲ 图1 通信和计算分离的架构对比空中计算的架构

▼表1 不同  $\varphi_k(\cdot)$  和  $\psi(\cdot)$  下的目标函数

目标函数	$\varphi_k(\cdot)$	$\psi(\cdot)$	$f(\cdot)$
算术平均	$\varphi_k(s_k) = s_k$	$\psi(\cdot) = 1/K$	$f = \sum_{k=1}^K s_k / K$
几何平均	$\varphi_k(s_k) = \ln(s_k)$	$\psi(\cdot) = \exp(\cdot)$	$f = (\prod_{k=1}^K s_k)^{\frac{1}{K}}$
多项式	$\varphi_k(s_k) = \omega_k s_k^{\beta_k}$	$\psi(\cdot) = 1$	$f = \sum_{k=1}^K \omega_k s_k^{\beta_k}$
欧式范数	$\varphi_k(s_k) = s_k^2$	$\psi(\cdot) = (\cdot)^{\frac{1}{2}}$	$f = \sqrt{\sum_{k=1}^K s_k^2}$

小值, 即  $f = \max_k(s_k)$  和  $f = \min_k(s_k)$ 。以计算最大值为例, 如图2所示, 通过将  $s_k$  进行二进制量化成  $x_k$ , 利用无线信道的或(OR)特性, 依次从最高位向最低位确认每一位为0还是1。从最高位开始, 最高位为1的节点发送一个脉冲信号, 接收节点如果接收到脉冲信号, 则可以判定最高位为1, 否则最高位为0。如果最高位为1, 则剔除掉最高位为0的点, 再一次进行第2位是0还是1的判决。一直从最高位进行到最低位, 从而确定每一位的0,1值的情况, 从而得到最大值。同理, 通过对  $x_k$  进行反码映射, 则可以求得最小值。通过空中计算可以确定1组节点的最大最小值, 而不需要收集每个节点的个体数据; 因此相应的信道使用次数只和量化



▲图2 利用空中计算来计算一组节点是最大、最小值

序列的长度有关, 而与节点数目无关, 对于海量节点计算时, 可以大大提高通信和计算效率。

空中计算分为模拟和数字2种实现方式, 其基本原理如图3所示。图3 a) 是模拟域中信号叠加的原理图。通过利用M个节点间的干扰, 接收端可以直接获取到所需的求和函数。模拟空中计算中每个节点不对收集

的数据进行任何形式的编码, 而仅对原始数据进行预处理后发射至空中, 最终由汇聚中心收集并且进行后处理。由于在真实环境中存在噪声等干扰, 为了能够方便地评价模拟空中计算的性能, 一般使用失真度作为指标, 其定义为:

$$d(\hat{f}(s), f(s)) = |\hat{f}(s) - f(s)|^2, \quad (2)$$

其中  $\hat{f}(s)$  代表估计所得的目标方程,  $f(s)$  则代表了理想的目标方程。

虽然模拟空中计算实现相对简单, 但是无法有效地对抗噪声, 特别是在发射功率较低的时候, 汇聚中心接收到的数据存在

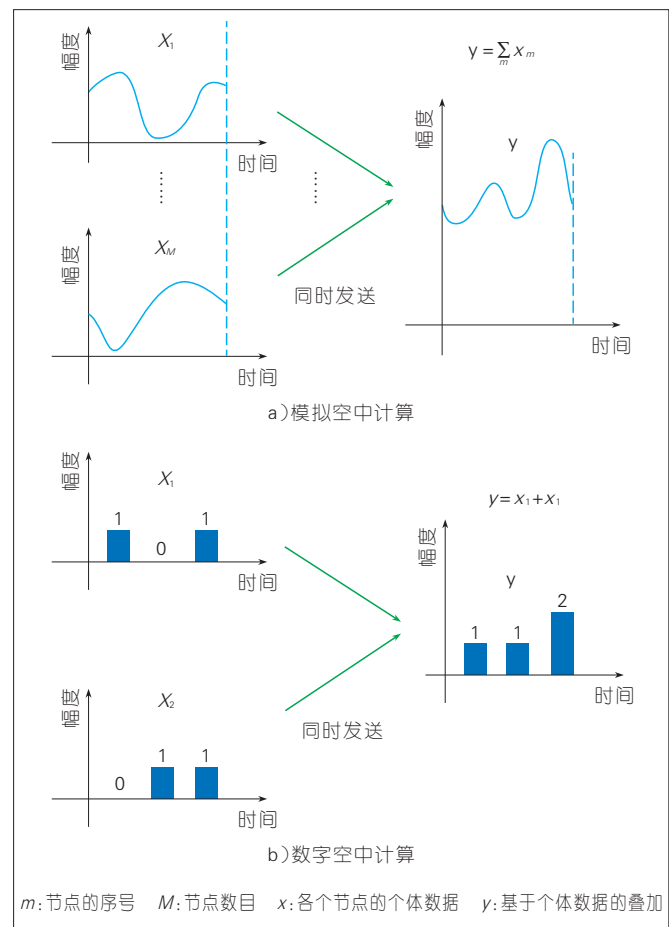


图3 模拟和数字空中计算的基本原理

$m$ : 节点的序号  $M$ : 节点数目  $x$ : 各个节点的个体数据  $y$ : 基于个体数据的叠加

很大干扰;因此,为了改善空中计算的抗噪能力,一类基于信源信道联合编码的数字空中计算被提出。图3 b)是数字域中一个简单的例子。多节点并发使得2个脉冲幅度调制(PAM)信号相互叠加,接收端所获取的新的PAM信号正是待计算的求和函数。节点对采集的数据进行相应的量化编码,再通过空中信道叠加发射,由汇聚节点收集并进行解码。在一个存在噪声和干扰的实际系统中,一般使用可达计算速率来评价数字空中计算的性能,其定义为:

$$R = \frac{k}{n} \quad (\text{每信道使用可以计算的函数数目}), \quad (3)$$

即在  $n$  次信道使用中传输  $k$  次目标方程,且当  $n$  足够大并能够进行无差错传输时,每次信道使用传输的方程数量为可达的计算速率。

### 3 空中计算的相关研究工作

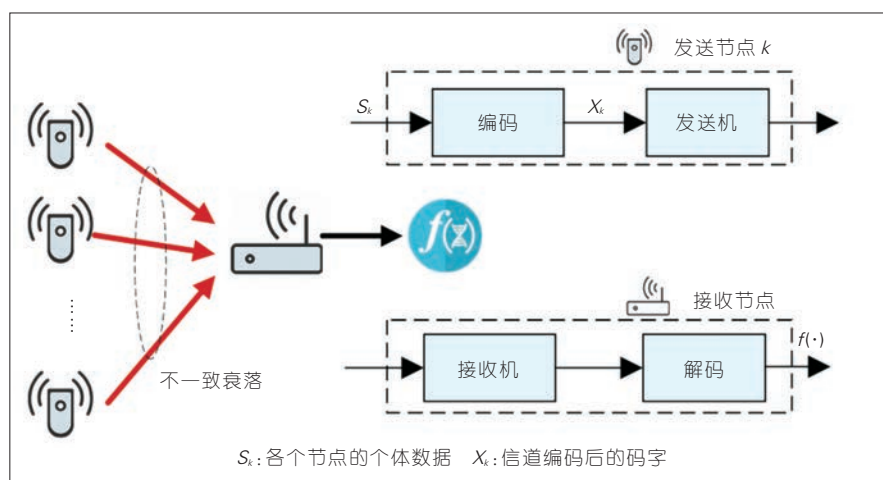
在现有的工作中,文献[4]中的工作使用无线信道的叠加特性实现了一部分简单的模拟函数的计算;文献[5]中的工作则是将模拟空中计算拓展到多簇中进行分析;文献[6]中给出了一种能应用在数字空中计算中且达到标准随机编码性能的数字编码方式;文献[7]中则基于这种编码方式给出了一个简单的数字空中计算架构;文献[8]中进一步提出了一种机会通信的方法来改善数字空中计算中可达计算

速率容易趋近于0的问题。其中,文献[4-7]关于空中计算的研究都非常理想,甚至忽略了无线信道本身具有的衰落特性。虽然文献[7-8]考虑了无线信道的衰落特性,但是所有的分析仅仅是基于窄带通信模型进行的,没有考虑宽带系统的频率选择性信道。除此之外,上述的所有研究都没有对无线网络中的同步或失配进行性能分析或者提出解决方案。

无线通信系统不可避免地受到噪声和干扰的影响,并且无线多址信道的衰落具有随机性和不一致性。这些因素都无法满足空中计算对于信号完美叠加的要求,从而造成相应的计算误差。与传统无线通信系统不同的是,空中计算通信系统中的目标信号不再是个体数据,而是个体数据的叠加,同时等效噪声也不再是节点间干扰和接收噪声,而是非完美叠加造成的计算误差;因此,传统的通信设计方法无法直接用于空中计算通信

系统。现有的空中计算研究对于通信系统的假设往往非常理想,缺乏实际通信参数对计算性能所造成的影响分析,以及针对性的通信设计。

为了对抗多节点无线信道不一致性的衰落,我们提出了一种迫一致性的收发机设计<sup>[9]</sup>,如图4所示,其核心思想在于通过收发机的预失真和后均衡,使得每个节点到接入点的等效信道的衰落都一致,从而满足求和的要求。进一步地,我们还研究了收发节点都具备多天线时相应的收发机设计<sup>[10]</sup>。利用多天线的复用增益,多天线迫一致性收发机不仅仅可以补偿不一致性衰落的信道,同时可以实现多个函数的空中计算。考虑到空中计算依赖于分布式节点的严格同步,我们提出了基于参考信号频谱重构的方法<sup>[11]</sup>,来测量各个节点之间的同步相位的误差。进一步地,基于测量的相位误差值,我们设计了相应的误差均衡器来降低由于同步相位误差而



▲ 图4 对抗不一致衰落的迫一致性收发机



带来的计算误差。

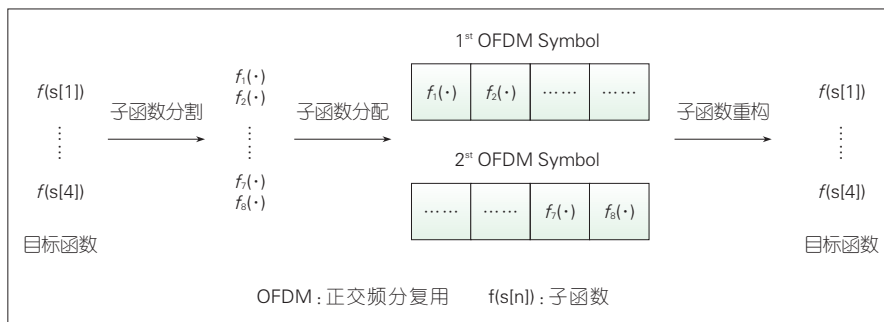
## 4 空中计算工作展望

### 4.1 宽带通信系统中的空中计算

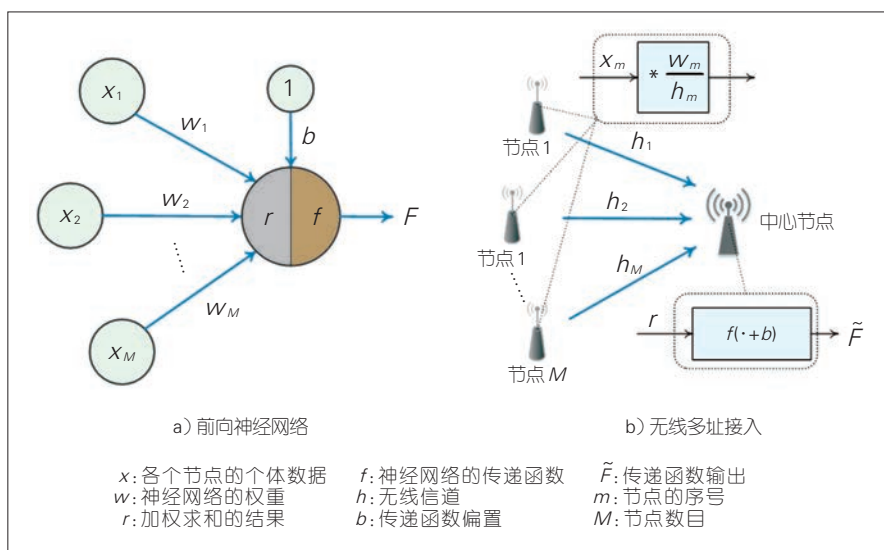
不同于窄带通信系统,宽带通信系统将经历频率选择性信道。即便通过迫一致性收发机设计,也很难完美地补偿无线信道的衰落使其满足计算要求。正交频分复用(OFDM)技术可以将宽带信道划分为多个窄带信道,从而较好地解决了频率选择性衰落的问题。在传统的无线通信系统中,比特级的粒度很小,很容易动态地分配到OFDM系统中的各个资源块。对于空中计算而言,目标方程只有1个,无法匹配OFDM系统中的资源块粒度;因此需要对目标方程进行分割,分配再重构,如图5所示。更小粒度的子方程可以满足资源粒度的匹配,满足空中计算的要求,所以最优的分割、分配和重构设计,是未来研究工作中的一个方向。

### 4.2 利用计算网络来模拟神经网络

如图6所示,前向神经网络的计算依赖于多个信源的加权和,而利用空中计算技术可以在不还原个体数据的前提下获得加权的计算结果;因此我们可以利用无线多址接入的通信网络来模拟前向神经结构的计算网络,从而实现通信和计算一体化,来提高整体效率。如何将通



▲图5 宽带通信系统的空中计算



▲图6 前向神经网络和无线多址接入结构的比较

信网络和计算网络结合在一起进行统一化设计,给出通信和计算的统一化效用并进行最优的设计也是未来研究工作中的一个方向。

## 5 结束语

空中计算的核心思想在于通信计算一体化,它改变了传统体系中通信和计算分离的架构,提示了一种全新的、可以用于直接实现计算结果传输的无线通信新架构。空中计算的技术特别适用于通信受限的计算应用场景,如基于海量物联节点的计

算领域。值得注意的是,空中计算对于无线信道一致性、节点分布式同步的要求比较苛刻,因此如何实现高精度的可靠空中计算是一个亟待探索和研究的重大问题。

### 参考文献

- [1] AKYILDIZ I F, NIE S, LIN S C, et al. 5G Roadmap: 10 Key Enabling Technologies [J]. Computer Networks, 2016, 106: 17-48. DOI:

- 10.1016/j.comnet.2016.06.010
- [2] DEHOS C, GONZALEZ J L, DOMENICO A DE, et al. Millimeter-Wave Access and Backhauling: the Solution to The Exponential Data Traffic Increase in 5G Mobile Communications Systems? [J]. IEEE Communications Magazine, 2014, 52(9): 88–95. DOI: 10.1109/MCOM.2014.6894457
- [3] ANEREWIS J G, BUZZI S, CHOI W, et al. What will 5G be? [J]. IEEE Journal on Selected Areas in Communications, 2014, 32(6): 1065–1082. DOI: 10.1109/JSAC.2014.2328098
- [4] NAZER B, GASTPAR M. Computation over Multiple-Access Channels [J]. IEEE Transactions on Information Theory, 2007, 53(10): 3498–3516. DOI: 10.1109/TIT.2007.904785
- [5] GOLDENBAUM M, BOCHE H, STANCZAK S. Harnessing Interference for Analog Function Computation in Wireless Sensor Networks [J]. IEEE Transactions on Signal Processing, 2013, 61(20): 4893–4906. DOI: 10.1109/TSP.2013.2272921
- [6] NAZER B, GASTPAR M. Compute-and-Forward: Harnessing Interference Through Structured Codes [J]. IEEE Transactions on Information Theory, 2011, 57(10): 6463 – 6486. DOI: 10.1109/TIT.2011.2165816
- [7] JEON S W, WANG C Y, GASTPAR M. Computation over Gaussian Networks with Orthogonal Components [J]. IEEE Transactions on Information Theory, 2014, 60(12): 7841–7861. DOI: 10.1109/TIT.2014.2364572
- [8] JEON S W, JUNG B C. Opportunistic Function Computation for Wireless Sensor Networks [J]. IEEE Transactions on Wireless Communications, 2016, 15(6): 4045–4059. DOI: 10.1109/TWC.2016.2533379
- [9] CHEN L, QIN X, WEI G. A Uniform-Forcing Transceiver Design for Over-the-Air Function Computation [J]. IEEE Wireless Communications Letters, 2018, 7(6): 942–945. DOI: 10.1109/LWVC.2018.2840157
- [10] CHEN L, ZHAO N, CHEN Y, et al. Over-the-Air Computation for IoT Networks: Computing Multiple Functions with Antenna Arrays [J]. IEEE Internet of Things Journal, 2018, 5(6): 5296–5306. DOI: 10.1109/JIOT.2018.2843321
- [11] CHEN L, ZHAO N, CHEN Y, et al. Over-the-Air Computation for Cooperative Wideband Spectrum Sensing and Performance Analysis [J]. IEEE Transactions on Vehicular Technology, 2018, 67(11): 10603–10614. DOI: 10.1109/TVT.2018.2866539

## 作者简介



陈力, 中国科学技术大学副研究员; 主要从事无线通信理论研究及系统实现工作; 作为项目负责人承担国家重大专项、国家自然科学基金项目等; 发表论文 30 余篇, 拥有 5 项国家发明专利。



卫国, 中国科学技术大学教授, 曾任国家“863”计划通信技术主题专家组成员、中国第三代移动通信系统研究开发项目总体组成员、国家“863”计划 B3G 移动通信重大项目总体组成员、“新一代宽带无线移动通信网”国家科技重大专项总体专家组成员; 主要从事无线通信技术、移动通信网络、信号处理等方面的研究; 获国家科技进步二等奖 1 项; 发表论文 100 余篇, 拥有 10 余项国家发明专利。

# 免调度非正交多址接入技术及其系统性能

## The Grant-Free NOMA Technology and Its System Level Performance

**摘要:** 主要探讨免调度非正交多址接入(NOMA)技术及其系统性能。NOMA 技术在实际应用中需要仔细研究并解决用户识别以及信道估计问题,尤其是在有碰撞的情况下。针对该问题,给出了相应的解决方案,以及系统级评估中的建模方法,并对基于预配置方式和随机选择方式的免调度多用户共享接入(MUSA)方案在实际系统中的性能表现进行了系统级仿真评估。仿真结果显示免调度 MUSA 方案具有明显的性能增益。

**关键词:** NOMA; 免调度; MUSA; 预配置; 随机选择; 碰撞

**Abstract:** Grant-free non-orthogonal multiple access (NOMA) technology and the system level performance are discussed in this paper. For grant-free NOMA technology, user identification and channel estimation should be carefully studied, especially for the case with collision. The corresponding solutions are provided in this paper, and modelling methods for system level evaluation are also given. The performances of grant-free multi-user shared access (MUSA) based on pre-configuration or random selection are evaluated by system level simulation. The evaluation results show that grant-free MUSA has obvious performance gain with practical assumptions.

**Key words:** NOMA; grant-free; MUSA; pre-configuration; random selection; collision

在5G新空口(NR)中,非正交多址接入(NOMA)技术受到了业界的广泛关注<sup>[1-2]</sup>。根据目前NOMA技术的研究进展,上行NOMA技术通过多个用户终端(UE)共享使用相同的传输资源,并通过高级接收机实现多用户检测和译码。当上行NOMA技术应用于免调度传输场景时,可以提升系统容量,简化系统流程,节省信令开销,降低终端功耗和传输时延,从而可

以应用于多种业务部署场景,包括大规模机器通信(mMTC)、增强型移动宽带(eMBB)和高可靠低时延通信(uRLLC),以及车到车(V2V)等<sup>[3-4]</sup>。

具体地,对于mMTC,主要业务需求是支持海量低成本、低功耗的终端进行偶发的小数据包传输;对于eMBB,同样关注偶发的小数据包传输场景,传统的接入流程和方案信令开销较大、效率较低;对于uRLLC和V2V,



李卫敏/LI Weimin<sup>1,2</sup>  
唐红/TANG Hong<sup>1</sup>  
李剑/LI Jian<sup>1</sup>  
胡宇洲/HU Yuzhou<sup>1</sup>

(1. 中兴通讯股份有限公司,广东 深圳 518057;

2. 移动网络和移动多媒体技术国家重点实验室,广东 深圳 518057)

(1. ZTE Corporation, Shenzhen 518057, China;

2. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.201901006

网络地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190129.1151.004.html>

收稿日期: 2018-12-26

网络出版日期: 2019-01-29

主要需求之一是低时延。免调度 NOMA 非常适合解决这些场景下的业务需求和问题,而且通过多用户复用可以明显提升系统容量和频谱效率。

关于 NOMA 在不同的无线资源控制(RRC)状态下的应用,首先,在 RRC 连接状态下,UE 通常是上行同步的,按照免调度传输机制,UE 可以自主地进行上行传输,不需要发起调度请求并等待调度,从而可以节省系统流程和信令开销,并降低时延和功耗。多个 UE 可以使用相同的资源进行传输,从而提升系统容量。这种状态下,我们可以采用预配置的方式,令多个 UE 使用相同的资源进行传输,并且令这些 UE 使用的参考信号以及扩展序列等不同,也就是避免这些 UE 的签名序列发生碰撞,以便进行用户识别与检测。但受限于此,系统资源利用率和 UE 传输效率会降低。另外,我们可以考虑的方式包括:通过预配置,令更多的 UE 使用相同的资源,允许 UE 的签名序列发生碰撞;当 UE 有业务到达时,可以随机选择传输资源和签名序列,进行竞争接入,接收机则通过更加复

杂或高级的盲检测算法实现用户识别与检测。

在 RRC 非激活状态下,UE 可能处于上行异步状态或可通过扩展循环前缀(ECP)来保证 UE 尽量上行同步。UE 可以采用免调度 one-shot 传输或采用 2 步随机接入(2-step RACH)机制传输<sup>[5-6]</sup>,从而缩减系统流程,降低时延和功耗。这种方法比较适用于小包、偶发等业务场景。当然,多个 UE 可能会在相同的资源上同时发起传输,结合 NOMA 技术的发射机和接收机设计,有利于支持更多的 UE 数量或业务负载。

目前,在第 3 代合作伙伴计划(3GPP)组织中由中兴通讯牵头的针对 5G NR 的 NOMA 技术研究项目(SI)已经结束,该项目得到了业界的广泛参与和研究,并已输出了技术报告<sup>[2]</sup>。

## 1 免调度 NOMA 方案

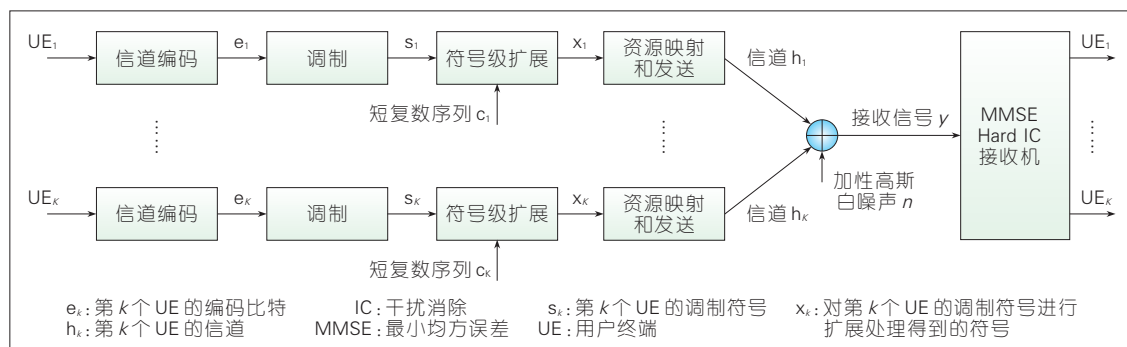
目前业界提出的各种 NOMA 方案中,1 个大类是多个 UE 分别使用长度较短的序列对其调制符号进行扩展,然后在相同的传输资源上发送,基站采用基于最小均方误差(MMSE)和

干扰消除(IC)的高级接收机实现多用户检测与译码。

图 1 为中兴通讯提出的一种 NOMA 方案即多用户共享接入(MUSA)方案。该方案在发射端采用长度较短的复数序列对调制符号进行扩展,序列长度为 2、4、6 等,序列元素来自于集合  $\{1, 1i, -1, -1i\}$ ,序列通常不包括 0 元素,也就是说序列是非稀疏的或可以称为 full length 序列。对于序列长度为 4,按照该序列设计准则,可以得到 16 条互相关绝对值不大于 0.5 的序列、32 条互相关绝对值不大于  $\sqrt{2}/2$  的序列或 64 条互相关绝对值小于 0.8 的序列<sup>[7]</sup>。低互相关序列的数量相对较多,有利于更好地保证免调度 NOMA 传输的性能。另外,序列元素仅实部或虚部有值,而且实部或虚部的取值为 1 或 -1,使得一些涉及序列的运算复杂度很低,仅需要执行加法运算即可。采用非稀疏或 full length 序列扩展,还有利于保持基于离散傅里叶变换扩展的正交频分复用(DFT-S-OFDM)波形的低峰均比(PAPR)属性。

该方案在接收端采用 MMSE Hard IC 接收机,其中

图 1  
多用户共享接入方案的发射机和接收机示意图





“Hard IC”的意思是：当1个UE的数据被成功译码后，根据译码器输出的硬比特，重构该UE发送的符号，用于干扰消除。干扰消除过程可以按照串行干扰消除(SIC)、并行干扰消除(PIC)或并串混合干扰消除(HIC)的方式执行。MMSE Hard IC接收机的实现复杂度和计算复杂度相对较低，而且可复用现有系统中常用的MMSE-IRC(IRC)接收机，复杂度增加不多。为了改善性能，也可考虑采用MMSE Soft IC接收机或期望传递算法(EPA)接收机，不过这些接收机的复杂度相对较高<sup>[2]、[8]</sup>。

将NOMA应用于免调度传输场景，需要考虑的主要问题是用户识别以及信道估计。对于预配置方式，多个UE使用的解调参考信号(DMRS)不同，基站可以通过DMRS进行用户识别及信道估计，漏检率和虚警率比较低，可根据DMRS和扩展序列之间的对应关系获取UE使用的扩展序列，用于多用户检测和译码。对于存在DMRS碰撞的情况，由于2个甚至多个UE可能使用相同的DMRS，基站通过该DMRS仅能识别到1个UE，而且信道估计结果是这些碰撞UE的信道之和；因此，DMRS碰撞将对系统性能产生严重影响。针对该问题，可以考虑通过增加DMRS数量来降低碰撞概率。进一步地，还可以针对DMRS进行干扰消除来改善性能，即当发生DMRS碰撞的多个UE中有1

个UE被成功译码后，重构该UE的DMRS接收信号用于干扰消除。由于原来基于DMRS的信道估计结果并不准确，这里可以利用已经被正确译码的各个UE的译码结果重构其发送符号，然后通过最小二乘(LS)算法重新进行联合信道估计得到更新的信道估计结果，用于DMRS干扰消除。由于同样的原因，数据部分的干扰消除也需要使用LS信道估计更新的信道估计结果<sup>[9]</sup>。

## 2 系统性能评估

本节我们将考虑实际网络部署中的一些因素，包括多小区网络与业务、覆盖情况、小区间干扰、实际信道估计、MMSE检测协方差矩阵估计、DMRS碰撞等，对免调度NOMA方案在实际系统中的性能表现进行系统级仿真评估分析。首先将简要描述系统级仿真(SLS)的链路到系统(L2S)映射方法，以及实际信道估计、DMRS碰撞、MMSE检测协方差矩阵估计的建模方法，然后给出系统级仿真假设与仿真结果。

### 2.1 仿真方法

L2S映射也可以称为物理层抽象，用来建模接收机处理过程，包括信道估计、MMSE检测、译码等。对于免调度NOMA，还需要建模用户识别、干扰消除等过程。

图2是免调度NOMA方案采用的MMSE Hard IC接收机的

L2S映射方法的流程图。如图2所示，该方法包括以下几个主要步骤：

(1)用户识别与信道估计。

对于预配置方式，由于漏检率和虚警率比较低，可以简化假设为理想用户识别。对于实际信道估计，可以通过在理想信道基础上添加信道估计误差来建模，如公式(1)所示：

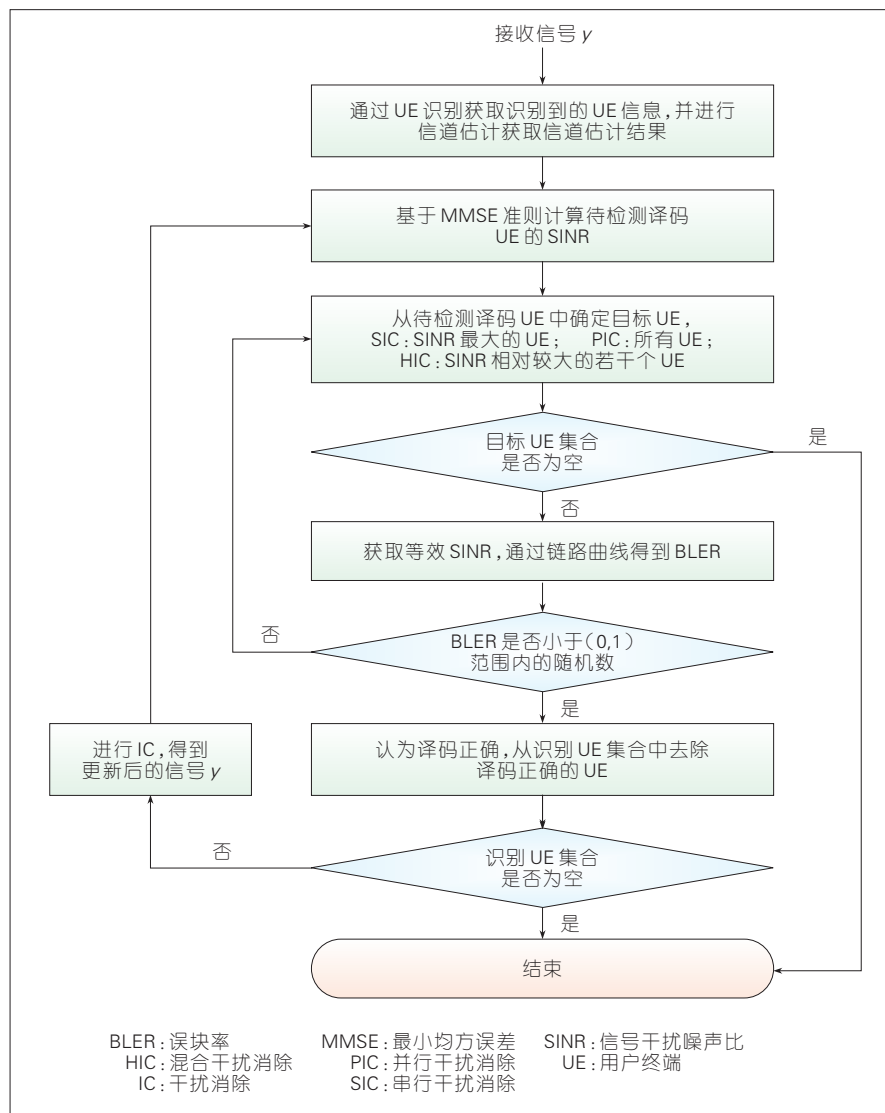
$$H_R = H_I + H_e, \quad (1)$$

其中， $H_R$ 表示实际信道， $H_I$ 表示理想信道， $H_e$ 表示信道估计误差。在实际网络中， $H_e$ 包括其他小区在DMRS上产生的干扰影响和噪声影响，与DMRS设计、信道估计方法、信道估计结果滤波平滑方法有关。可以假设其他小区在DMRS上产生的干扰统计上服从高斯分布，那么，我们则可以将 $H_e$ 建模为均值为0、方差为 $\sigma_e^2$ 的高斯分布随机变量。

对于存在DMRS碰撞的情况，需要考虑DMRS碰撞在用户识别、信道估计等方面的影响。由于2个甚至多个UE可能使用相同的DMRS，基站通过该DMRS仅能识别到1个UE，而且信道估计结果是这些碰撞UE的信道之和，例如，假设UE1和UE2使用了相同的DMRS，则信道估计结果可以按照公式(2)建模：

$$H_R = H_{n1} + H_{n2} + H_e, \quad (2)$$

其中， $H_{n1}$ 和 $H_{n2}$ 分别表示UE1



▲ 图2 免调度非正交多址接入方案采用的 MMSE Hard IC 接收机的链路到系统映射方法的流程图

和 UE2 的理想信道, 信道估计误差  $H_e$  可以按照与上述描述类似的方法建模。

(2) 基于 MMSE 准则计算待检测译码的 UE 的 SINR。

在多小区网络中, 接收信号  $y$  可以按照公式(3)描述:

$$y = \sum_{k=1}^K H_k s_k + \sum_{j=1}^J H_j s_j + n, \quad (3)$$

其中,  $K$  为当前小区中进行

NOMA 传输的 UE 数量,  $s_k$  为当前小区中第  $k$  个 UE 发送的调制符号,  $H_k$  为当前小区中第  $k$  个 UE 的包括空域和码域的联合信道系数。那么,  $H_k$  可以表示为  $NL \times 1$  的矢量, 其中  $N$  为接收天线数量,  $L$  为扩展序列长度;  $J$  表示其他小区中在相同资源上进行传输的小区间干扰 UE 的数量,  $s_j$  为第  $j$  个小区干扰 UE 发送的调制符号,  $H_j$  为第  $j$  个小区

间干扰 UE 的包括空域和码域的联合信道系数,  $H_j$  也可以表示为  $NL \times 1$  的矢量,  $n$  为均值为 0、方差为  $\sigma^2$  的加性高斯白噪声 (AWGN)。

那么, MMSE 检测权重可以按照公式(4)计算:

$$W_k = H_k^H R_{yy}^{-1} = H_k^H \left( \sum_{k=1}^K H_k H_k^H + \sum_{j=1}^J H_j H_j^H + \sigma^2 I \right)^{-1}, \quad (4)$$

其中, 假设当前小区中第  $k$  个 UE 为待检测的目标 UE,  $()^H$  表示共轭转置,  $R_{yy}$  为接收信号  $y$  的自相关矩阵,  $I$  为  $NL \times NL$  的单位矩阵。

进一步, 第  $k$  个 UE 的 MMSE 检测结果可以表示为  $\hat{s}_k = W_k \times y$ , 那么其 SINR 可以按照公式(5)计算:

$$SINR_k = \frac{|W_k H_k|^2}{\sum_{i=1, i \neq k}^K |W_k H_i|^2 + \sum_{j=1}^J |W_k H_j|^2 + W_k (\sigma^2 I) W_k^H} \quad (5)$$

需要注意的是, 在理想情况下, 计算 MMSE 权重和 SINR 时, 可以使用理想信道; 但在实际场景下, 需要进行实际信道估计和  $R_{yy}$  估计。那么, 在计算 MMSE 权重时,  $H_k$  需要使用当前小区中第  $k$  个 UE 的实际信道估计结果,  $R_{yy}$  估计可以根据接收信号  $y$  来估计。具体地, 可将接收信号  $y$  转换为  $NL \times T$  的矩阵  $Y$ , 其中  $T$  为 UE 发送的调制符号的数量, 则有  $R_{yy} = YY^H/T$ ; 而对于 MMSE 检测结果的 SINR, 由于根据  $\hat{s}_k = W_k^* y$  计算, 各个 UE 的理想信道自然的包含在接收信号  $y$  中, 仍然可以使用各个 UE 的

理想信道。

(3)获取等效 SINR,通过链路曲线得到 BLER。

按照基于接收比特互信息率(RBIR)方法<sup>[10]</sup>将当前小区中第  $k$  个 UE 的 SINR 映射为等效 SINR,如公式(6)所示:

$$SINR_k^{eff} = \varphi^{-1}(\frac{1}{M} \sum_{m=1}^M \varphi(SINR_{k,m})), \quad (6)$$

其中,  $M$  为资源单元数量,每个资源单元上有 1 个 SINR 计算结果,  $\varphi(\cdot)$  为用于进行 RBIR 映射的非线性可逆函数。然后,根据等效 SINR 查找 AWGN 信道场景下的 BLER vs. SNR 链路曲线得到 BLER,作为本次传输的 BLER。进一步地,将该 BLER 与 (0, 1)范围内的 1 个随机数进行比较,如果小于则认为本次传输可以译码正确,否则认为本次传输译码失败。

图 2 中当 1 个或 1 组 UE 没有被正确译码时,我们则可以尝试对剩余 UE 继续进行译码,这样做可以改善性能,尤其是对于 SIC 而言<sup>[9]</sup>。

(4)进行 IC,得到更新后的信号  $y$ 。

当 1 个 UE 被正确译码后,可以针对该 UE 进行干扰消除。首先重构该 UE 的发送符号,然后根据其信道,得到该 UE 的接收信号,并将其减去,实现干扰消除。

需要注意的是,在实际场景下需要使用实际信道估计结果进行干扰消除,对于预配置方式,如前文所述,可以根据信道

估计误差  $H_e$  来建模干扰消除残留误差。对于存在 DMRS 碰撞的情况,如前文所述,由于碰撞的影响,原来基于 DMRS 的信道估计结果可能并不准确;因此,可以考虑利用已经被正确译码的各个 UE 的重构发送符号进行 LS 联合信道估计得到各个 UE 的更新的信道估计结果,并用于干扰消除。随着被正确译码的 UE 数量越来越多,LS 联合信道估计结果会越来越准确。这里假设 UE1 被正确译码,经过分析,其 LS 信道估计的归一化误差可以按照公式(7)<sup>[9], [11]</sup>近似建模,其他情况可以以此类推:

$$\frac{|h_e|^2}{|h_1|^2} \approx \frac{\sum SNR_{int} + 1}{(x_1^H x_1) SNR_1}, \quad (7)$$

其中,  $h_1$  为 UE1 在 1 个天线上频域信道系数,  $h_e$  为 UE1 的 LS 信道估计误差,  $x_1$  为 UE1 发送的

数据符号,  $SNR_1$  为 UE1 的理想 SNR,  $SNR_{int}$  则可以包括其他尚未被正确译码的干扰 UE 的理想 SNR。

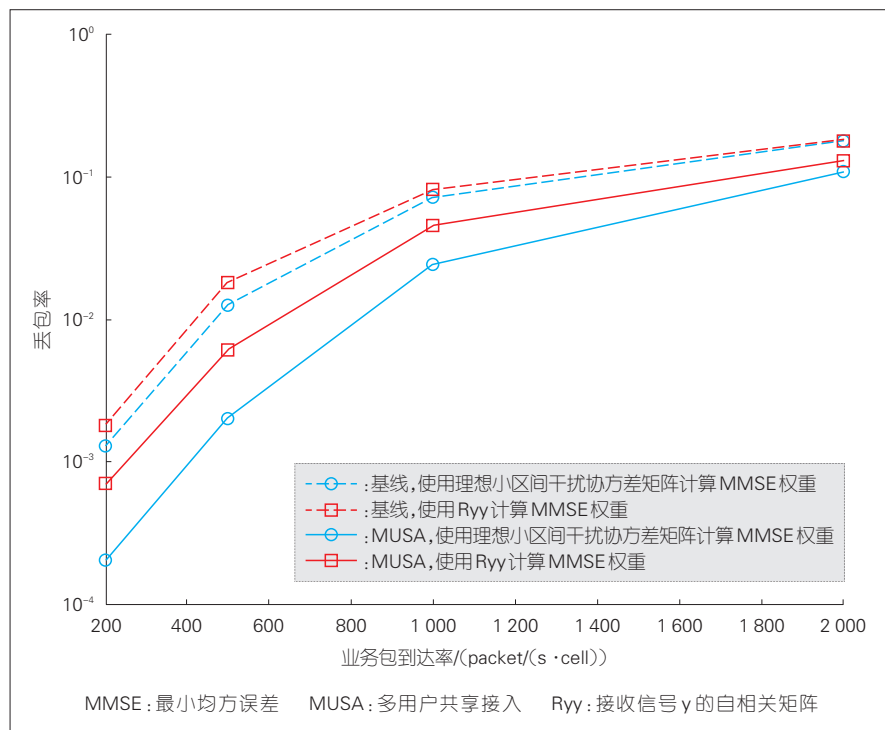
对于存在 DMRS 碰撞的情况,由于 DMRS 碰撞,UE 识别会出现漏检;因此当 1 个 UE 被正确译码并且针对该 UE 进行干扰消除后,可以重新执行如图 2 所示的过程,即重新进行用户识别等步骤。这样有利于发现之前被漏检的用户,从而改善性能。

## 2.2 仿真结果

下面我们将在 mMTC 场景下对免调度 NOMA 方案在实际系统中的性能表现进行系统级仿真评估。仿真假设参考文献[2],另外一些仿真假设如表 1 所示。仿真评估了 2 种方案,包括预配置免调度 MUSA 方案和随机选择免调度 MUSA 方案。

▼表 1 系统级仿真假设

仿真参数	仿真假设
载频与站间距	700 MHz, 1 732 m
仿真带宽	6 个物理资源块
业务包大小	20~200 bytes, 帕累托分布, 成形参数 $\alpha = 2.5$ , 协议开销 29 bytes
传输块大小	将业务包进行分包, 每个传输块 25 bytes, 其中包含 5 bytes 分包开销
扩展序列	长度为 4 的复数扩展序列 <sup>[7]</sup>
HARQ/重复	对于预配置方案, HARQ 传输次数最多为 8; 对于随机选择方案, 则将 UE 分为 3 组, 每组 UE 采用不同的重复次数
丢包判断准则	如果 1 个传输块在 HARQ/重复传输后没有被正确接收, 则认为丢包
每个小区的 UE 数量	100 个
UE 功率控制	开环功率控制: 对于预配置方案, $P_0 = -100$ dBm, $\alpha = 1$ ; 对于随机选择方案, $P_0 = -95$ dBm, $\alpha = 1$
基站天线数量	2 Rx
基站天线倾角	92
信道估计	实际信道估计
基站接收机	MMSE-PIC 接收机
HARQ: 混合自动重传请求    MMSE: 最小均方误差    PIC: 并行干扰消除    UE: 用户终端	



▲ 图3 预配置免调度MUSA方案的仿真结果

图3给出了预配置免调度MUSA方案的仿真结果。按照mMTC场景的仿真假设,UE和基站间的最大耦合损耗达到了144 dB,可以将UE的发射功率集中在1个物理资源块(PRB)上传输,以便改善覆盖的性能;因此,在该方案下,每个UE使用1 PRB + 6 ms的资源单元。另外,在该仿真中,基线方案没有进行扩展,MUSA方案采用长度为4的复数扩展序列进行扩展。从仿真结果可以看到,当MMSE权重计算采用理想的小区干扰协方差矩阵时,MUSA相对于基线在丢包率=1%处大约有70%的性能增益;当MMSE权重计算采用Ryy估计时,MUSA相对于基线在丢包率=1%处有接近50%的性能增益。需

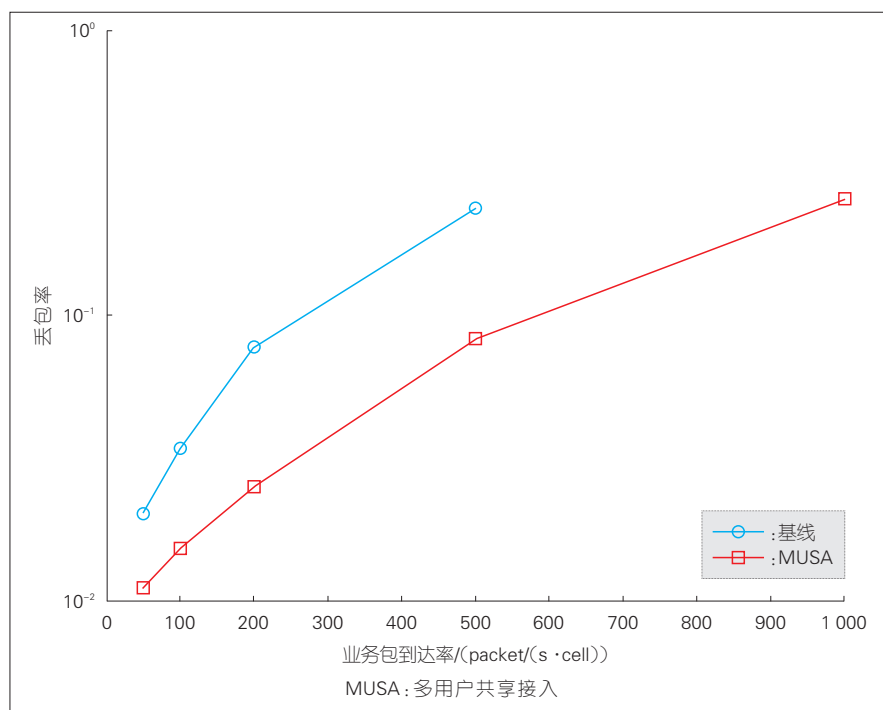
要说明的是,小区内UE均采用如前文所述的实际信道估计。

图4给出了随机选择免调

度MUSA方案的仿真结果。该仿真中,基线方案使用24个DMRS,没有进行扩展,使用的基本资源单元为1 PRB + 1 ms,3组UE的重复次数分别为1、4和16;MUSA方案使用64个DMRS,采用长度为4的复数扩展序列进行扩展,复数扩展序列的总数量同样为64个,DMRS和扩展序列一一对应,使用的基本资源单元为1 PRB + 4 ms,3组UE的重复次数分别为1、1和4,并且将第1组UE的扩展序列的总能量归一化为1,以便公平对比。从仿真结果看,MUSA采用64个DMRS及长度为4的复数扩展序列,碰撞概率更低,性能更好,而且有大约100%的性能增益。

### 3 结束语

免调度NOMA可以提升系



▲ 图4 随机选择免调度MUSA方案的仿真结果



统容量,简化系统流程,节省信令开销,降低终端功耗和传输时延,从而可以应用于多种业务部署场景,包括 mMTC、eMBB、uRLLC 以及 V2V 等,可以应用于 RRC 连接状态以及 RRC 非激活状态。

本文中,我们对免调度 NOMA 方案进行了简要介绍,对免调度 NOMA 在实际应用中需要考虑和解决的用户识别与信道估计等一系列相关问题进行了讨论,给出了对免调度 NOMA 进行系统级评估的接收机建模方法,并对基于预配置方式和随机选择方式的免调度 MUSA 方案在实际系统中的性能表现进行了系统级仿真评估,仿真结果显示免调度 MUSA 方案具有更为明显的性能增益。

## 致谢

本研究得到中兴通讯股份有限公司袁志锋高级工程师、袁

弋非博士、田力博士的帮助,谨致谢意!

## 参考文献

- [1] Study on New Radio Access Technology: Physical Layer Aspects:TR 38.802 V14.2.0 [S]. 3GPP, 2016
- [2] Study on Non-Orthogonal Multiple Access (NOMA) for NR: TR 38.812 V1.0.0 [S]. 3GPP, 2018
- [3] YUAN Z, YU G, LI W, et al. Multi-User Shared Access for Internet of Things [C] // 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring). USA: IEEE, 2016 (1): 1-5. DOI: 10.1109/VTCSpring.2016.7504361
- [4] TIAN L, YAN C, LI W, et al. On Uplink Non-Orthogonal Multiple Access for 5G: Opportunities and Challenges [J]. China Communications, 2017, 14(12): 142-152. DOI: 10.1109/CC.2017.8246331
- [5] ZTE. Procedures Related to NOMA: R1-1812174 RAN1#95 [S]. 3GPP, 2018
- [6] Qualcomm Incorporated. Procedures Related to NOMA: R1-1813407 RAN1#95 [S]. 3GPP, 2018
- [7] ZTE. Link-Level Simulation Results for mMTC Scenario: R1-1808155 RAN1#94[S]. 3GPP, 2018
- [8] ZTE. Complexity Analysis of NOMA Receivers: R1-1812173 RAN1#95 [S]. 3GPP, 2018
- [9] ZTE. Multi-User Receivers for NOMA: R1-1805841 RAN1#93 [S]. 3GPP, 2018
- [10] IEEE 802.16m Evaluation Methodology Document (EMD): IEEE 802.16m-08/004r5 [S]. IEEE, 2009
- [11] ZTE. Physical Layer Abstraction for NOMA SLS: RAN1#94bis, R1-1810208 [S]. 3GPP, 2018

## 作者简介



**李卫敏**,中兴通讯股份有限公司无线研究院算法部技术预研高级工程师;主要负责 5G 物理层技术的研究工作,研究方向包括功率控制、干扰控制、非正交多址接入、海量机器通信等;发表论文 10 余篇。



**唐红**,中兴通讯股份有限公司无线研究院算法部技术预研高级工程师;主要负责 5G 物理层技术的链路级仿真和系统级仿真等工作。



**李剑**,中兴通讯股份有限公司无线研究院算法部技术预研高级工程师;主要从事系统级仿真工作,目前研究方向包括超密集网络、非正交多址接入、高可靠低时延通信等;申请专利 20 余篇。



**胡宇洲**,中兴通讯股份有限公司无线研究院算法部技术预研高级工程师;主要负责 5G 物理层新空口标准、非正交多址接入技术方面的工作。

# 智慧标识网络动态防御机制与应用

## Dynamic Defense Mechanism and Application for Smart Identifier Networks

**摘要:** 介绍了智慧标识网络(SINET)的基本原理,重点分析了 SINET 的资源动态适配机制在安全性方面的解决方案,提出了动态适配安全防护总体架构和具体机制。理论研究和实验表明,智慧标识网络动态防御机制可灵活调度网络资源、智慧服务迁移,从而较好地应对一些网络安全攻击问题。

**关键词:** SINET; 未来网络; 动态安全防护; 资源适配

**Abstract:** In this paper, the system model and design principles of smart identifier networks (SINET) are introduced, and the application of security in dynamic resource adaptation is analyzed. Then the dynamic adaptive defense architecture and mechanism are proposed. Theoretical research and practical deployment both prove that SINET dynamic defense mechanism can flexibly schedule network resources and intelligent service migration, so as to better cope with some network security attacks.

**Key words:** SINET; future networks; dynamic security defense; resource adaptation

现有互联网起源于 20 世纪 60 年代,采用“沙漏模型”的设计思想,具有“三重绑定”的特征,即服务的“资源和位置的绑定”、网络的“控制和数据绑定”以及“身份与位置绑定”。随着用户规模的增长和多媒体应用的增多,传统互联网已经逐渐暴露出来以上各种弊端。这种网络体系和机制是相对静态和僵化的,无法从根本上满足信息网络的高速、高效、海量、泛在等通信需求。

未来网络体系架构已成为

世界各国信息网络领域的研究热点。美国自然科学基金委的全球网络创新环境(GENI)<sup>[1]</sup>、未来互联网设计项目(FIND)<sup>[2]</sup>计划,以及欧盟的未来互联网研究和试验(FIRE)<sup>[3]</sup>计划等。此外,美国自然科学基金委的未来互联网体系架构(FIA)项目资助的命名数据网络(NDN)<sup>[4]</sup>、泛在移动(MobilityFirst)<sup>[5]</sup>、星云(NEBULA)<sup>[6]</sup>、富有表现力的互联网架构(XIA)<sup>[7]</sup>等重大项目,美国开放网络基金会发起的软件定义网络(SDN)研究和欧洲



于成晓/YU Chengxiao  
刘刚/LIU Gang  
张宏科/ZHANG Hongke

(北京交通大学下一代互联网互联设备国家工程实验室,北京 100044)  
(National Engineering Lab on Next Generation Internet, Beijing Jiaotong University, Beijing 100044, China)

DOI: 10.12142/ZTETJ.201901007  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190129.1715.010.html>

收稿日期: 2018-12-28  
网络出版日期: 2019-01-29

电信标准化协会发起的网络功能虚拟化(NFV)研究,都从不同方面研究未来网络架构。中国也非常重视对未来网络体系架构的研究,国家“973”计划先后启动了“一体化可信网络与普适服务体系基础研究”“可信可控的IP网基础研究”“面向服务的未来网络体系结构与机制研究”等项目。然而,未来网络架构需要支持对网络状态的实时感知并智慧地进行网络资源的动态适配,随着网络流量和网络状态自主地从全局角度考虑网络资源,动态提高网络资源利用率。

智慧标识网络(SINET)<sup>[8-12]</sup>作为一种新型网络架构,在解决现有网络存在的诸多弊端上,创造性地引入了“三层”“两域”的概念,纵向适配,横向解耦,并引入行为描述信息来实现内嵌于网络架构中的智慧性,使网络资源能够从根本上支持资源动态适配,实现网络资源高度优化利用。智慧标识网络的核心思想是实现网络资源资源优化适配,文中我们从网络动态适配角度对“智慧标识网络中网络动态防御”这一问题展开研究。

## 1 SINET 研究背景

### 1.1 SINET 体系模型与工作原理

SINET 针对导致互联网诸多弊端的本质原因设计了三层、两域为特征的网络体系架构,如图1所示。三层包括智慧服务层、

资源适配层和网络组建层,具体功能为:智慧服务层主要负责服务的标识和描述,以及服务的智慧查找与动态匹配等;资源适配层通过感知服务需求与网络状态,动态地适配网络资源并构建网络族群,网络族群由一组网络节点或者相似功能的节点组成,负责决策优化、任务分配等,以充分满足服务需求进而提升用户体验,并提高网络资源利用率;网络组件层主要负责数据的存储与传输等实际操作,以及网络组件的行为感知与聚类。两域指实体域和行为域。具体来说,实体域指各种各样的网络对象,是相对静态的,既包括硬件网络资源,又包括软件协议参数,本质上是指实际运行的网络。行为域指对实体域网络对象的处理逻辑和策略,是相对动态的,便于智能、适配、协同和决策等。

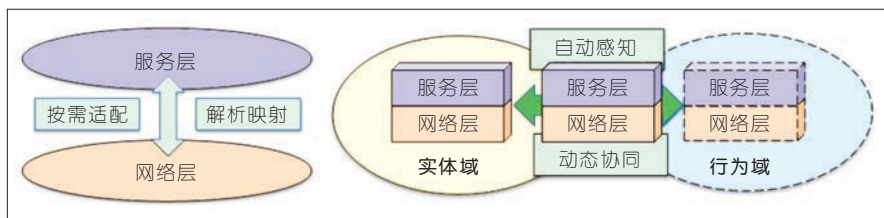
SINET的资源动态适配工作原理如图1所示,在智慧服务层和资源适配层之间,使用行为匹配机制:在行为域中根据服务需求行为描述和族群功能行为描述形成1次映射,为智慧服务寻求最佳的族群功能模块搭配组合,然后根据实体域内族群间的协作机制,控制指定的族群功能

模块进行协同工作,从而实现服务标识到族群标识的映射过程。在资源适配层和网络组件层之间,使用行为聚类机制:在行为域中根据族群行为描述和组件行为描述形成另1次映射,为族群功能模块判定最合理的网络组件构成,然后再根据实体域的族群内联动机制,在族群功能模块内的网络组件之间建立相互联动关系,以完成族群功能模块的整体功能,实现由族群标识到组件标识的映射过程。通过这2次映射,网络资源可以依据服务需求动态适配,从而实现智慧服务。

总之,智慧标识网络的三层、两域体系通过动态感知网络状态并智能匹配服务需求,进而选择合理的网络族群及其内部组件来提供智慧化的服务。同时,通过引入行为匹配、行为聚类、网络复杂行为博弈决策等机制来实现资源的动态适配和协同调度,大幅度提高网络资源利用率,降低网络能耗等,显著提升用户体验。

### 1.2 资源动态适配机制原理

资源适配层通过管理以及编排网络功能族群来满足动态的服务需求,其原理模型如图2



▲ 图1 智慧标识网络总体结构模型

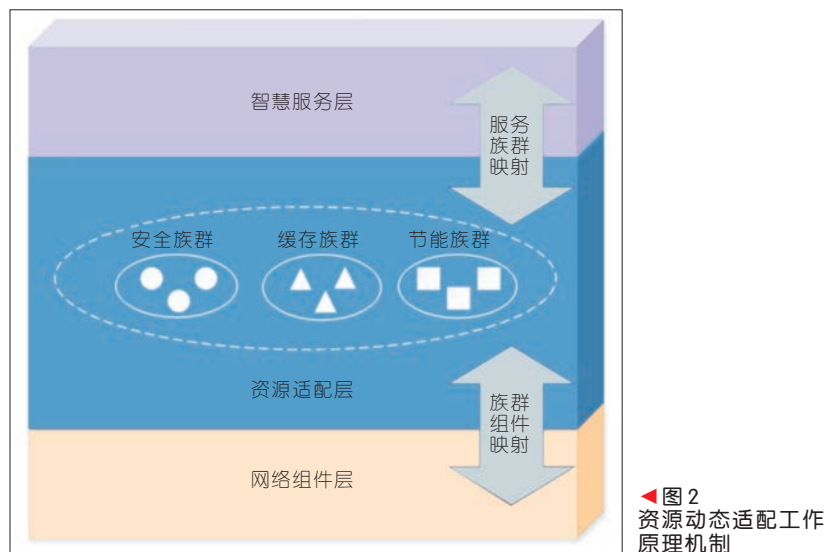


图2  
资源动态适配工作  
原理机制

所示。

资源适配层位于 SINET3 层结构中的中间层,负责感知来自上层服务需求和来自下层的网络状态,进而控制不同的网络组建形成最优化的网络族群,为上层提供特定的优质服务。资源适配层可以根据不同的服务编排不同的网络族群来支持不同类型的服务。基于上述设计,资源适配层可以更好地解决传统网络中存在的性能难题。

特点并结合网络功能服务链,我们提出了 SINET 动态防御管控架构,如图 3 所示。SINET 动态防御管控架构是在 OpenStack 的平台下进行实现的。在短时间内的适配过程中,OpenStack 平台下基于 SINET 中资源适配层族群的特点,与服务功能链(SFC)技术相结合,根据从上层感知到的服务需求,编排好对应的络族群,通过 2 次映射将服务需求与底层网络组件紧密地联

系在一起,来实现不同的网络任务。在长时间内的适配过程中,通过日志数据收集,大数据分析平台以及人工智能技术等功能的引入,让网络更加智能化,提高网络流量的可预测性。基于 SINET 的动态安全防御架构的优势在于:

(1)实现动态资源适配。能够实时监测控制和分析网络中存在的网络资源和流量变化,对不用网络服务适配不同的网络族群来与网络组件对接,动态高效地管理网络,同时保证网络服务质量(QoS),优化网络能耗。

(2)动态网络防御。采用策略驱动的动态防御措施,利用 SFC 技术生成弹性安全功能服务链,快速察觉异常流量侵入,有效地抵御外来攻击,确保网络安全高效运行。

(3)加快网络自适应和动态感知。通过日志数据收集,以及大数据分析平台及人工智能技术等一系列功能的引入,让网络

## 2 SINET 动态安全防御管控

### 2.1 动态防御管控总体架构

SINET 中资源适配层作为整体架构的中间层,是连接上层与下层实现网络的智慧化、自动化的重要组成部分。作为资源适配层的实体部分,族群是构建不同网络服务提供需求,并且根据不同族群的功能高效地管理网络组件。根据 SINET 总体架构以及 SINET 中资源适配层中的

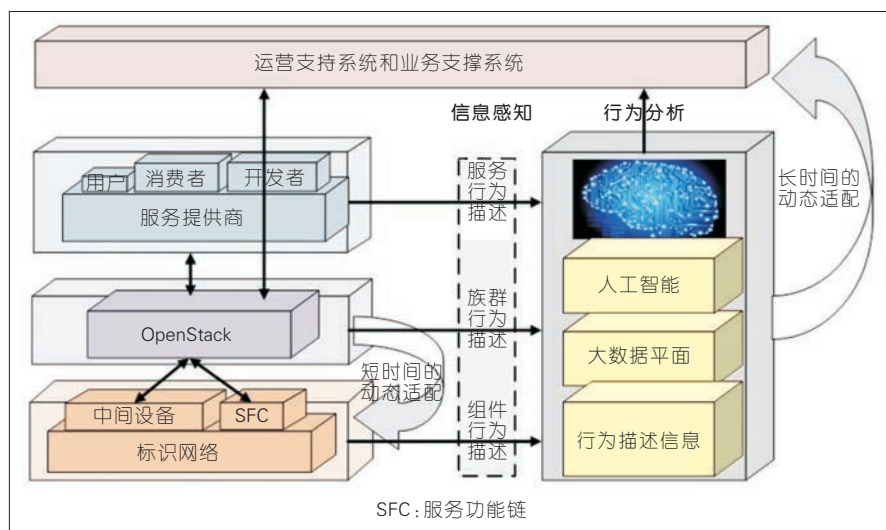


图3 智慧标识网络动态防御管控架构



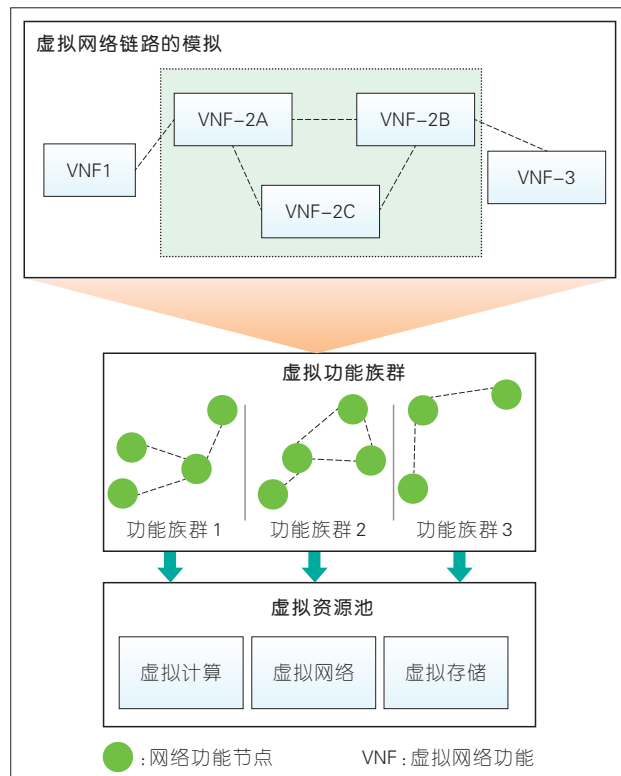
更加智能化,提高网络流量的可预测性。

## 2.2 SINET 动态安全防御云平台

根据 SINET 的动态安全防御架构的特点和优势,基于 OpenStack 技术,将物理机的资源整合到一起形成虚拟资源池进行统一管理。图 4 为 SINET 动态安全防御管控云平台。OpenStack 技术统一管理虚拟组件(包括虚拟计算组件、虚拟网络组件以及虚拟存储组件等),利用 SINET 提出的三层、两域架构,针对不同的动态安全防护措施,动态适配出不同的网络功能族群,然后通过 SFC 技术对每个功能族群里面的虚拟网络功能(VNF)进行编排从而达到对应的服务需求。SINET 动态安全防御管控云平台基于 OpenStack 技术,可以支持多租户,可扩展性强,资源弹性伸缩;融合了容器与内核虚拟技术,根据不同的仿真场景灵活适配不同的技术,并且融合了 Linux 流量控制工具,实现了仿真网络特性的细粒度模拟。

动态安全防御云平台支持大规模拓扑动态变化的网络仿真。拓扑更新速度可达到毫秒级,规模可达 100 倍物理机个数;网络仿真链路信道实时模拟。每条仿真链路的误码率( $10^{-4} \sim 10^{-9}$ )、时延(小于 200 s、精度 1 ms)可单独设置,支持延时抖动、延时与抖动相关性的设置;支持秒级的仿真场景快速复

图 4  
智慧标识网络动态安全  
防御管控云平台



现,支持资源动态适配技术。

图 5 和图 6 为 SINET 动态安全防御功能实例,其中分类器检测异常流量,筛选不同类型的流量,将经过不同 SFC 的数据包进

行不同封装。图 5 是未发起恶意流攻击时网络族群构成的网络组件拓扑,在流量监测控制组件中,正常流量为 2 Mbit/s,异常流量几乎为零。当控制攻击

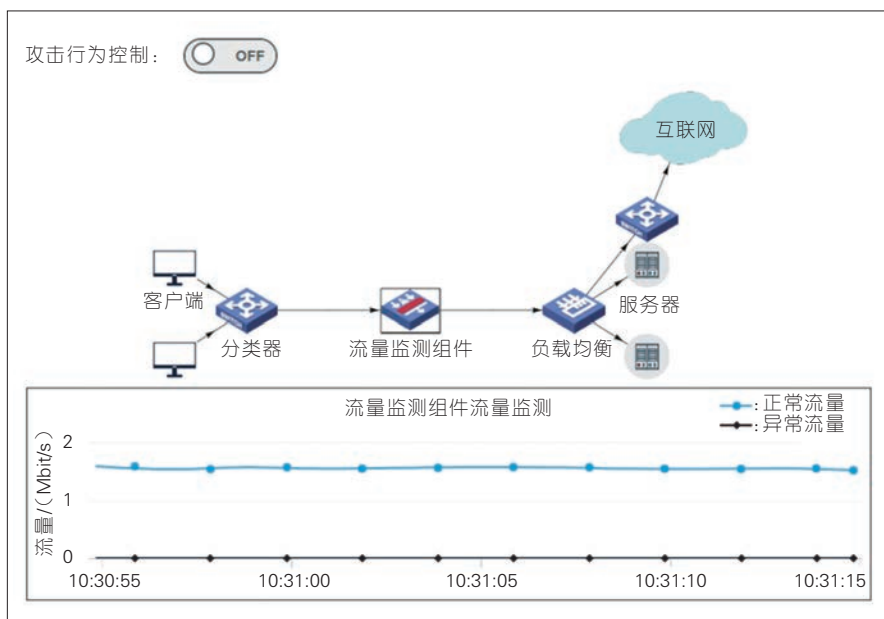
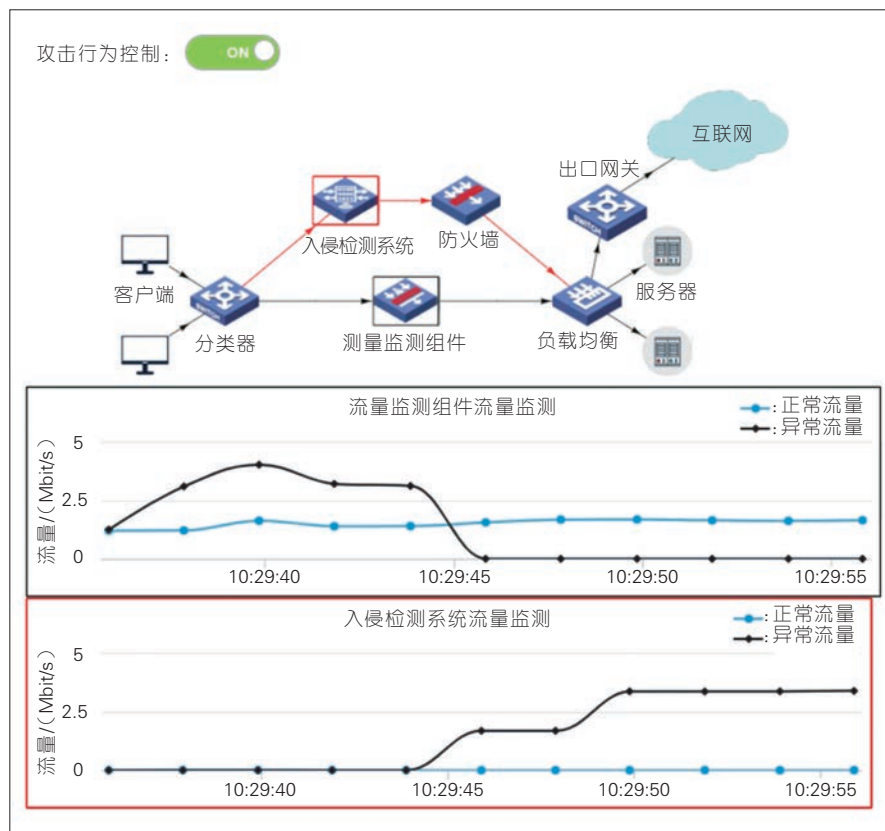


图 5 未开启攻击时网络状态



▲ 图6 开启攻击时网络状态

的行为开启后,网络中存在大量流量的异常流量,如图6所示。流量监测控制组件会感知异常流量的激增,从而智能化地适配新的网络族群,将生成的异常流量导入到新生成的网络族群链中,从而进行进一步的流量检测。当网络中的异常流量消失,又会自动更新为正常情况下的网络族群,实现高效地动态适配网络资源。

### 3 结束语

本文中,我们在简要介绍 SINET 体系结构的基础上,引入 OpenStack 和 SFC 技术,提出了基于 SINET 的动态安全防护总体架构。进一步地,我们还介绍了

SINET 动态安全防护管控云平台基本功能。通过实验,我们展示了智慧标识网络动态安全防护的效果,实现了动态地适配网络资源,从而缓解网络攻击行为造成的网络影响。SINET 作为一种新型网络架构和技术,有望应用于未来的 5G 核心网,使网络更高效、更灵活、更智能、更安全。

#### 参考文献

- [1] NSF. GENI: Global Environment for Network Innovations [EB/OL]. (2018-04-28)[2019-12-28]. <http://www.geni.net>
- [2] NSF. FIND: Future Internet Network Design. [EB/OL]. [2019-12-28]. <http://www.nets-find.net/>
- [3] European Commission. FIRE [EB/OL]. [2019-12-28]. <http://cordis.europa.eu/fp7/ict/fire/>
- [4] NSF. Named Data Networking [EB/OL]. [2019-12-28]. <http://www.named-data.net/>
- [5] NSF. MobilityFirst [EB/OL]. [2019-12-28]. <http://mobilityfirst.winlab.rutgers.edu/>

- [6] NSF. Nebula [EB/OL]. [2019-12-28]. <http://nebula-fia.org/>
- [7] NSF. XIA-eXpressive Internet Architecture [EB/OL]. [2019-12-28]. <http://www.cs.cmu.edu/~xia/>
- [8] 张宏科, 罗洪斌. 智慧协同网络体系基础研究[J]. 电子学报, 2013, 41(7): 1249-1254. DOI: 10.3969/j.issn.0372-2112.2013.07.001
- [9] ZHANG H K, DONG P, YU S, et al. A Scalable and Smart Hierarchical Wireless Communication Architecture Based on Network/User Separation [J]. IEEE Wireless Communications, 2017, 24(1): 18-24. DOI: 10.1109/mwc.2017.1600135wc
- [10] ZHANG H K, LUO H B, CHAO H C. Dealing with Mobility-Caused Outdated Mappings in Networks with Identifier/Locator Separation [J]. IEEE Transactions on Emerging Topics in Computing, 2016, 4(2): 199-213. DOI:10.1109/tetc.2015.2449664
- [11] ZHANG H K, QUAN W, SONG J Y, et al. Link State Prediction-Based Reliable Transmission for High-Speed Railway Networks [J]. IEEE Transactions on Vehicular Technology, 2016, 65(12): 9617-9629. DOI:10.1109/tvt.2016.2598473
- [12] 权伟, 张宏科. 未来互联网体系的研究现状、热点与探索实践[J]. 中国科学:信息科学, 2017, 47(6): 804-810

#### 作者简介



于成晓, 北京交通大学在读博士研究生; 目前主要研究方向为未来网络架构、网络传输协议; 已发表国际会议论文 1 篇。



刘刚, 北京交通大学在读博士研究生; 目前主要研究方向包括未来互联网、网络安全、软件定义网络和可编程数据平面等; 已发表国际会议论文 1 篇。



张宏科, 北京交通大学教授、下一代互联网互联设备国家工程实验室主任, 电子学会理事, 国家“973”项目首席科学家; 目前主要从事未来网络体系理论及工程技术应用相关研究; 曾获得 2014 年国家技术发明奖二等奖、2017 年国家技术发明奖二等奖、2017 年教育部技术发明一等奖等; 已发表论文 100 余篇。

# 上行-下行联合优化的 uRLLC 传输

## Joint Uplink-Downlink Transmission Design for uRLLC

**摘要:** 考虑有界信道状态信息(CSI)误差的影响,解决了上下行联合设计的高可靠低时延通信(uRLLC)传输问题,以保障最差情况下物理层空口传输的端到端可靠性。利用有限码长容量公式近似刻画传输速率、时延和可靠性之间的关系,在信道估计误差、最大功率和传输可靠性的约束条件下,最小化上下行所需的传输时延。通过分析目标函数的单调性和凸性,对优化问题进行等价转化,并提出了有效的交替优化算法来求解该问题。仿真结果分析了信道估计误差、发送功率、可靠性对传输时延的影响,清晰地表明了系统参数之间的折中关系。

**关键词:** uRLLC; 鲁棒波束赋形; 有限码长编码; 资源分配; 交替优化

**Abstract:** A joint uplink and downlink beamforming design and reliability tradeoff is considered to guarantee the end-to-end performance requirements of ultra-reliable and low-latency communication (uRLLC) traffics. Under the bounded channel state information (CSI) error, the worst-case transmission latency is minimized, subject to the over-the-air reliability of all links. The transmission rate, latency and packet error probability are characterized by the approximation equation based on the finite blocklength coding. The optimal design problem is reformulated by utilizing the monotonicity and the convexity of the end-to-end latency with respect to the reliability, and an efficient algorithm is proposed based on the alternating optimization technique. The simulation results validate the impact of the CSI error, power and reliability requirements on the transmission latency. It also shows the tradeoff between these system parameters.

**Key words:** uRLLC; robust beamforming; finite blocklength coding; resource allocation; alternating optimization

**高**可靠低时延通信(uRLLC)是5G的3个典型场景之一,也是未来超5代移动通信系统(B5G)/第6代移动通信系统(6G)需要重点研究的核心场景之一<sup>[1]</sup>,被普遍认为是工厂自动化、自动驾驶、远程手术、增强

现实等应用的基本要求<sup>[2-4]</sup>。在这些应用中,数据业务的传输直接涉及到生产、操作的安全、效率和用户感受,因此对无线通信传输的实时性、可靠性和传输效率都有严格的指标需求,如对于一般的uRLLC传输需要达到的



成晶/CHENG Jing<sup>1</sup>  
沈超/SHEN Chao<sup>1</sup>  
夏树强/XIA Shuqiang<sup>2</sup>

(1. 北京交通大学,北京 100044;  
2. 中兴通讯股份有限公司,广东 深圳 518057)  
(1. Beijing Jiaotong University, Beijing 100044, China;  
2. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.201901008  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190131.0906.002.html>

收稿日期: 2018-12-29  
网络出版日期: 2019-01-31

基金项目: 中兴通讯产学研合作项目; 国家自然科学基金(61871027); 北京市科委重大项目(Z181100003218010)

指标是支持 32 字节数据包的误包率 (PEP) 小于  $10^{-5}$ , 用户平面时延不超过 1 ms<sup>[2]</sup>; 对于增强型车到万物 (eV2X) 场景, 支持 300 字节数据包的 PEP 小于  $10^{-5}$ , 用户平面时延为 3~10 ms。uRLLC 系统优化设计的基本挑战是对短包传输时传输速率、可靠性与传输时延的性能刻画, 进而需要解决的是根据其中的内在机理和关系, 设计有效的传输机制、进行优化的资源分配、控制实际传输条件下的端到端时延和可靠性<sup>[5]</sup>。

一般而言, 对于信噪比为  $\rho$  的加性高斯白噪声 (AWGN) 信道, 我们采用香农公式  $\log(1+\rho)$  来刻画系统传输的信道容量。香农公式需要发送端采用无穷长码长的 Gaussian 编码实现无误码的传输极限, 因此, 对于 uRLLC 场景的短包传输, 香农容量公式不适合刻画系统传输速率、可靠性和时延的关系。为此, Y. Polyanskiy 等人基于有限码长容量的研究结果, 进一步提出了 AWGN 信道下采用有限码长编码时可达容量上界的一个高斯近似。这一近似结果清晰地刻画了传输速率、码长和可靠性之间的关系<sup>[6]</sup>, 因此可以适用于 uRLLC 系统的性能分析和优化设计<sup>[7]</sup>。然而这一近似公式不是一个标准的容量效用函数, 严格地说, 该公式非凸非凹; 因此, 尽管信道编码方面的研究已经表明该公式可以较好地刻画实际系统的性能<sup>[8-9]</sup>, 但基于此进

行系统优化设计时, 非凸性带来了很大的挑战。另一方面, uRLLC 系统要求端到端的性能保障, 因而涉及物理层以及物理层之上的高层协议, 而目前对此类系统尚无较好的跨层建模方法、分析范式和设计工具, 从而使得 uRLLC 系统设计的理论研究仍然是一个较为开放的问题。

为此, 本文中我们考虑一个典型的 uRLLC 传输场景, 如 eV2X 场景。假定传感器节点监测到突发事件后将信息上行传输到路边单元 (RSU), RSU 再将消息转发给附近的移动终端 (MT)。本文中, 我们考虑从传感器上行传输到各终端下行接收的系统设计, 通过上行链路传输与下行链路传输的联合优化设计, 实现物理层空口的端到端系统性能保障。我们将使用有限码长下的容量近似公式来刻画端到端链路的可靠性、传输时延和传输速率之间的内在关系, 并且考虑信道估计误差的影响, 实现最差情况下 uRLLC 传输的性能保障。这一系统优化设计的核心挑战是近似公式的非凸性和上下行传输的联合设计使得问题严格非凸, 信道误差使得优化设计问题等价于存在无穷多的约束条件。为此, 我们通过分析目标函数的单调性和凹凸性, 将问题进行等价的转化, 进而提出了高效的交替优化算法来求解该问题。最后通过数值仿真证明了信道估计误差、传输功率、可靠性指标对传输时延的

影响, 揭示了 uRLLC 系统传输指标之间的折中关系。

## 1 系统模型和问题建模

考虑如图 1 所示的 eV2X 场景: 传感器节点 (SN) 捕捉到路面前方发生车辆碰撞的突发事件后, 通过 RSU 向附近的  $K$  个移动终端发送紧急消息。假设 SN 和各 MT 都配置了单天线, 而 RSU 配置了  $N_t$  根天线。假设从 SN 到 RSU 的上行信道  $h_0 \in \mathbb{C}^{N_t}$  以及从 RSU 到各 MT 的下行信道  $h_k \in \mathbb{C}^{N_r}, k \in K \triangleq \{1, \dots, K\}$  均为准静态块衰落信道, 对应的信道估计分别为  $\hat{h}_0 \in \mathbb{C}^{N_t}, \hat{h}_k \in \mathbb{C}^{N_r}$ , 相应的信道误差为:

$$e_k = h_k - \hat{h}_k, k = 0, 1, \dots, K, \quad (1)$$

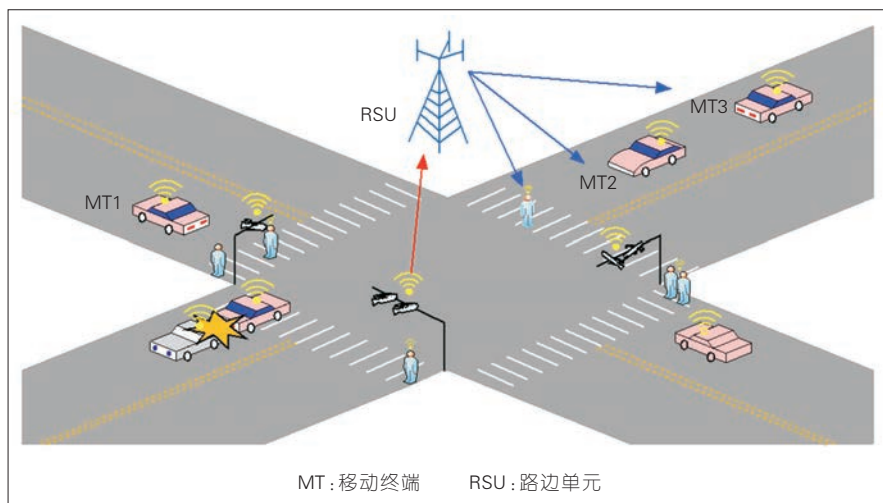
且满足有界误差的假定, 即  $\|e_k\| \leq \delta_k, k = 0, 1, \dots, K$ 。

从 SN 到 RSU 的上行传输数据包为  $N_0$  比特, 被编码成长度为  $m_u$  个符号的码字。RSU 处的接收信号可以表示为:

$$y_0[n] = \sqrt{p_0} h_0 x_0[n] + z_0[n], n = 1, \dots, m_u, \quad (2)$$

其中,  $x_0[n] \sim CN(0, 1)$  表示发送的码字, 并服从零均值、单位方差的高斯分布,  $p_0$  是 SN 的发送功率 ( $p_0 \leq P_0$ ),  $z_0[n] \sim CN(0, \sigma^2 I_{N_t})$  为 RSU 端的 AWGN。RSU 根据信道状态信息 (CSI) 估计结果对  $y_0[n]$  进行接收波束成形, 记接收波束向量为  $w_0$ , 则 RSU 解码  $x_0[n]$  时最差情况下的信噪比 (SNR)  $\rho_0$  可以表示为:





▲ 图1 一个典型的增强型车到万物场景

$$\rho_0 = \min_{\|e_k\| \leq \delta_0} \frac{p_0 E(w_0^H h_0)^2}{N_s \sigma^2 \|w_0\|^2} = \min_{\|e_k\| \leq \delta_0} \frac{p_0 |w_0^H (\hat{h}_0 + e_0)|^2}{N_s \sigma^2 \|w_0\|^2} \quad (3)$$

根据有限码长编码下的容量近似公式,  $N_0, m_u, \rho_0$  以及误包率  $\varepsilon_0 \ll 0.5$  近似服从:

$$\frac{N_0}{m_u} \approx \log_2(1 + \rho_0) - \sqrt{\frac{1}{m_u} \left(1 - \frac{1}{(1 + \rho_0)^2}\right)} \frac{Q^{-1}(\varepsilon_0)}{\ln 2} \triangleq R(\rho_0, \varepsilon_0, m_u), \quad (4)$$

其中,  $Q^{-1}(\cdot)$  表示高斯  $Q$  函数的反函数,  $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-t^2/2} dt$ 。

RSU 以  $\varepsilon_0$  的误包率从  $y_0[n]$  中译码 SN 发送的信息并进行处理后产生  $K$  个信息, 然后分别发送给  $K$  个终端。假设发送给第  $k$  个 MT 的数据包长度为  $N_k$  比特, 这  $N_k$  比特的信息被编码成长度为  $m_d$  个符号的单位功率的码字, 用  $s_k[n], n=1, \dots, m_d$  表示, 则 RSU 发射的下行信号可表示为  $\sum_{k=1}^K w_k s_k[n], n=1, \dots, m_d$ , 其中  $w_k \in \mathbb{C}^{N_t}$  是发送  $s_k[n]$  的波束向量; 因此, 第  $k$  个 MT 的接收信号

则为:

$$y_k[n] = h_k^H w_k s_k[n] + \sum_{j=1, j \neq k}^K h_k^H w_j s_j[n] + z_k[n], n=1, \dots, m_d, \quad (5)$$

其中,  $z_k[n] \sim CN(0, \sigma^2)$  为 AWGN。这样一来, 下行第  $k$  个 MT 的信干噪比 (SINR) 在信道误差的影响下可能的最小 SINR 则可以表示为:

$$\rho_k = \min_{\|e_k\| \leq \delta_k} \frac{|h_k^H w_k|^2}{\sum_{j=1, j \neq k}^K |h_k^H w_j|^2 + \sigma^2} = \min_{\|e_k\| \leq \delta_k} \frac{|(\hat{h}_k + e_k)^H w_k|^2}{\sum_{j=1, j \neq k}^K |(\hat{h}_k + e_k)^H w_j|^2 + \sigma^2}, \quad (6)$$

其中,  $k \in K$ 。类似地, 给定 PEP  $\varepsilon_k \in (0, 0.5]$ , 第  $k$  个 MT 的下行可达速率可近似为:

$$\frac{N_k}{m_d} \approx \log_2(1 + \rho_k) - \sqrt{\frac{1}{m_d} \left(1 - \frac{1}{(1 + \rho_k)^2}\right)} \frac{Q^{-1}(\varepsilon_k)}{\ln 2} \triangleq R(\rho_k, \varepsilon_k, m_d). \quad (7)$$

这里需要注意的是, 我们假定下行  $K$  个数据包均编码成码长为  $m_d$  的符号, 从而便于刻画

系统的性能, 简化系统设计。

称 SN 到 RSU 再到第  $k$  个 MT 的传输为第  $k$  条链路, 那么这一链路的端到端空口传输可靠性为  $(1 - \varepsilon_0)(1 - \varepsilon_k)$ 。本文中我们考虑在 SN 和 RSU 最大发送功率的约束下, 通过上下行传输的联合设计, 使得系统在有界信道误差的影响下仍旧保障所有  $K$  条链路的可靠性, 并最小化系统的空口传输时延  $m_u + m_d$ 。为此, 我们可以将这一上下行传输的联合设计问题建模为:

$$\min_{\rho_0, \{e_k\}_{k=1}^K} m_u + m_d \quad (8a)$$

$$s.t. \rho_0 \leq P_0, \sum_{k=1}^K \|w_k\|^2 \leq P_{\max}, \quad (8b)$$

$$1 - (1 - \varepsilon_0)(1 - \varepsilon_k) \leq \varepsilon_{\max}, \forall k \in K, \quad (8c)$$

$$\frac{N_0}{m_u} = \log_2(1 + \rho_0) - \sqrt{\frac{1}{m_u} \left(1 - \frac{1}{(1 + \rho_0)^2}\right)} \frac{Q^{-1}(\varepsilon_0)}{\ln 2}, \quad (8d)$$

$$\frac{N_k}{m_d} = \log_2(1 + \rho_k) - \sqrt{\frac{1}{m_d} \left(1 - \frac{1}{(1 + \rho_k)^2}\right)} \frac{Q^{-1}(\varepsilon_k)}{\ln 2}, \quad (8e)$$

其中, 公式 (8b) 表示 SNR 和 RSU 的最大发送功率约束; 公式 (8c) 表示第  $k$  条链路的端到端可靠性指标, 不能低于  $1 - \varepsilon_{\max}$ ; 公式 (8d) 和 (8e) 分别刻画了采用有限码长编码时上行、下行传输的性能关系, 其中  $\rho_k, k=0, 1, \dots, K$  的表达式分别见公式 (3) 和公式 (6)。

求解问题 (8) 的挑战主要在于 2 个方面: 首先是有限码长编码下的近似公式 (8d) 和 (8e) 关

于  $\rho_k$  即非凸也非凹;其次是对任意满足有界误差条件的信道误差向量  $e_k$ , 约束公式(8d)和(8e)都要成立,而每一个  $e_k, k \in K$  均影响到所有波束向量  $\{w_k\}_{k=1}^K$  的设计。这些因素在目标函数的作用下又进一步使得上下行传输的优化设计互相耦合,从而使得问题严格非凸,难以处理。

接下来,我们进一步分析问题的结构并提出有效的算法来求解该问题。

### 1.1 单调性和凸性分析

我们首先分析有限码长编码下容量近似公式的性质。该近似公式为:

$$\frac{N}{m} \approx \log_2(1+\rho) - \frac{1}{m} \left(1 - \frac{1}{(1+\rho)^2}\right) \frac{Q^{-1}(\varepsilon)}{\ln 2}, \quad (9)$$

其中,  $N$  是发送信息的比特数,  $m$  是码长,  $\varepsilon \in (0, 0.5]$  表示 PEP,  $\rho$  为信噪比。相应地,我们定义 PEP 函数  $\varepsilon(m, \rho)$  为:

$$\varepsilon(m, \rho) = Q\left(\frac{m \ln(1+\rho) - N \ln 2}{\sqrt{m} \sqrt{1 - (1+\rho)^{-2}}}\right). \quad (10)$$

对于公式(9)和(10)中  $\varepsilon, m, \rho$  之间的相互关系,我们有以下结论。

引理 1<sup>[10]</sup>: 对于任意给定的  $N, \varepsilon$ ,  $m$  则可以为  $\rho$  的严格单调减函数。

引理 2: PEP 函数  $\varepsilon(m, \rho)$  分别是  $m, \rho$  的严格单调减函数,而且对于任意的  $\rho > 0$ , 当  $m$  满足  $\mathcal{M} \triangleq \{m | \varepsilon(m, \rho) < 0.5\}$  时,  $\varepsilon(m, \rho)$  是

$m$  的严格凸函数<sup>[11]</sup>。

下面我们证明引理 2。

首先公式(10)中的 PEP 函数可以写成:

$$\varepsilon(m, \rho) \triangleq Q\left(a\sqrt{m} - \frac{b}{\sqrt{m}}\right), \quad (11)$$

其中,  $a = \frac{\ln(1+\rho)}{\sqrt{1 - (1+\rho)^{-2}}} > 0$ ,

$b = \frac{N \ln 2}{\sqrt{1 - (1+\rho)^{-2}}} > 0$ 。因为

$$Q'(x) = \frac{-1}{\sqrt{2\pi}} e^{-x^2/2}, \text{ 则 } \varepsilon(m, \rho) \text{ 对 } m$$

求偏导得:

$$\frac{\partial \varepsilon(m, \rho)}{\partial m} = \frac{-am - b}{2\sqrt{2\pi} m^{\frac{3}{2}}} e^{-\frac{(a\sqrt{m} - b/\sqrt{m})^2}{2}} < 0, \quad (12)$$

所以 PEP 函数  $\varepsilon(m, \rho)$  是  $m$  的严格单调减函数。

第二, 为证明 PEP 函数  $\varepsilon(m, \rho)$  是  $\rho$  的严格单调减函数, 我们首先证明对于任意给定的  $m$ ,  $f(m, \rho)$  是  $\rho$  的严格单调增函数, 其中:

$$f(m, \rho) \triangleq \frac{m \ln(1+\rho) - N \ln 2}{\sqrt{m} \sqrt{1 - (1+\rho)^{-2}}}. \quad (13)$$

令  $t = 1 + \rho > 1$ , 则:

$$\bar{f}(m, t) \triangleq f(m, \rho) \triangleq \frac{mt \ln t - tN \ln 2}{\sqrt{m(t^2 - 1)}}. \quad (14)$$

$\bar{f}(m, t)$  关于  $t$  的偏导为:

$$\frac{\partial \bar{f}(m, t)}{\partial t} = \frac{m(t^2 - \ln t - 1) + N \ln 2}{(t^2 - 1)\sqrt{m(t^2 - 1)}} \quad (15a)$$

$$\triangleq m(t^2 - \ln t - 1) + N \ln 2 \quad (15b)$$

$$\geq N \ln 2 > 0, \quad (15c)$$

其中,  $x \triangleq y$  表示  $x$  和  $y$  同号。

因为  $t > 1$ , 所以公式(15b)成立。又因为  $t^2 - \ln t$  在  $t > 1$  时是严格增函数, 所以公式(15c)成立。故  $f(m, \rho)$  是  $\rho$  的严格单调增函数, 又基于  $Q(x)$  是  $x$  的严格单调减函数的事实, 所以 PEP 函数  $\varepsilon(m, \rho)$  是  $\rho$  的严格单调减函数。

现在证明  $\varepsilon(m, \rho)$  是关于  $m \in \mathcal{M}$  的严格凸函数。在  $\varepsilon(m, \rho) < 0.5$  的假设下, 根据公式(11)可得:

$$a\sqrt{m} - \frac{b}{\sqrt{m}} > 0 \Leftrightarrow am > b. \quad (16)$$

对  $\varepsilon(m, \rho)$  求关于  $m$  的二阶偏导, 化简后可得:

$$\frac{\partial^2 \varepsilon(m, \rho)}{\partial m^2} \triangleq a^3 m^3 - b^3 + (a^2 b + a)m^2 + (3b - ab^2)m \quad (17a)$$

$$> (a^2 b + a)m^2 + (3b - ab^2)m \quad (17b)$$

$$\triangleq (a^2 b + a)m + 3b - ab^2 \quad (17c)$$

$$> ab^2 + b + 3b - ab^2 = 4b > 0, \quad (17d)$$

其中, 公式(17d)可由公式(16)推导得出。综上所述, 引理 2 证明完毕。

根据上面 2 个引理, 可以得出下面的命题 1。

**命题 1:** 当联合设计问题(8)取得最优解时, 约束条件(8b)与(8c)的等号成立。

证明: 对于约束(8b)中的上行功率约束, 根据公式(3),  $\rho_0$  随  $p_0$  的增大而增大, 而  $m_u$  是  $\rho_0$  的严格减函数, 所以  $p_0$  应尽可能得大使得  $m_u$  尽可能得小, 故最优解时等号成立。对于约束

(8b)的下行功率约束,假设最优解时等号不成立,那么同时线性增大所有  $\{w_k\}_{k \in K}$  的功率从而使使得约束(8b)取等号,根据公式(6),同时扩大  $\{w_k\}_{k \in K}$  以相同的倍数,使得所有  $\rho_k$  增大。根据  $m_d$  是  $\rho_k$  的严格减函数这一性质可知  $m_d$  可以进一步减小,与假设相矛盾;因此最优解时,约束(8b)等号成立。对于约束(8c),假设最优解时存在某个终端,不妨记为 MT  $\bar{k}$ ,取严格小于号。我们可以适当增大  $\varepsilon_{\bar{k}}$  使约束(8c)取等号,那么根据引理2可以在不改变  $m_d$  的前提下通过降低  $w_{\bar{k}}$  的功率来降低  $\rho_{\bar{k}}$ 。所降低的功率又可以进一步分配给所有发射波束从而提高所有移动终端的 SINR。又根据引理2可知,保持  $m_d$  不变、提高  $\{\rho_k\}$  可以降低  $\{\varepsilon_k\}$ ,这就允许我们进一步降低  $m_u$  使得目标函数更小,这与最优性假设相矛盾。因此最优解时,对所有的  $k \in K$  约束(8c)的等号成立。

**命题2:** 当联合设计问题(8)取得最优解时,最差情况下的上行 SNR  $\rho_0 = \frac{P_0}{N_t \sigma^2} (\|\hat{h}_0\| - \delta_0)^2$ 。

证明:根据命题1, SN 发射功率  $p_0 = P_0$ 。我们注意到, RSU 接收波束  $w_0$  的设计仅仅影响约束(8d)中的  $\rho_0$ ; 因此根据引理2,我们总是希望设计  $w_0$  使得  $\rho_0$  最大化,即:

$$w_0^* = \arg \max_{w_0} \frac{|w_0^H (\hat{h}_0 + e_0)|^2}{\|w_0\|^2}, \forall \|e_0\| \leq \delta_0. \quad (18)$$

因为  $g(w_0) \triangleq \frac{|w_0^H (\hat{h}_0 + e_0)|^2}{\|w_0\|^2}$  可以视作瑞利商(Rayleigh quotient),所以上行传输 SNR 的最大值为  $\lambda_{\max} \left( (\hat{h}_0 + e_0)(\hat{h}_0 + e_0)^H \right) = \|\hat{h}_0 + e_0\|^2$ , 其中,  $\lambda_{\max}(X)$  表示矩阵  $X$  的最大特征值,  $w_0^*$  为对应的特征向量。那么在信道误差的影响下,最差情况下的 SNR 为:

$$\rho_0 = \min_{\|e_0\| \leq \delta_0} \frac{P_0}{N_t \sigma^2} \|\hat{h}_0 + e_0\|^2. \quad (19)$$

显然,  $e_0^* = -\frac{\delta_0 \hat{h}_0}{\|\hat{h}_0\|}$ ,  $w_0^* = \hat{h}_0$ , 代入公式(19)可得  $\rho_0 = \frac{P_0}{N_t \sigma^2} (\|\hat{h}_0\| - \delta_0)^2$ , 命题得证。

## 1.2 问题重构

根据命题1和2,我们将  $p_0 = P_0$ ,  $\varepsilon_k = 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \forall k$ ,  $\rho_0 = \frac{P_0}{N_t \sigma^2} (\|\hat{h}_0\| - \delta_0)^2$  代入问题(8)。同时根据引理1和2,码长可以表示为 PEP 和信噪比的隐函数。具体而言,根据(8d)和(8e),上下行传输的码长可以分别表示为  $m_u \triangleq m_u(\varepsilon_0)$  和  $m_d \triangleq m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \rho_k \right)$ , 则联合设计问题(8)可以改写为:

$$\min_{\varepsilon_0, \{\rho_k, w_k\}_{k=1}^K} m_u(\varepsilon_0) + \max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \rho_k \right) \quad (20a)$$

$$s.t. \rho_k \leq \frac{|(\hat{h}_k + e_k)^H w_k|^2}{\sum_{j=1, j \neq k}^K |(\hat{h}_k + e_k)^H w_j|^2 + \sigma^2}, \forall \|e_k\| \leq \delta_k, \forall k, \quad (20b)$$

$$\sum_{k=1}^K \|w_k\|^2 \leq P_{\max}. \quad (20c)$$

**命题3:** 问题(20)与原问题(8)等价。

证明:首先,问题(20)取得最优解时,  $m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \rho_k \right)$  对所有的  $k \in K$  都相同,否则至少存在某个 MT  $\bar{k}$  的码长小于  $\max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \rho_k \right)$ 。那么根据引理2,不妨令  $w_{\bar{k}}$  的功率降低使得  $\rho_{\bar{k}}$  减小到 MT  $\bar{k}$  的码长等于  $\max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}, \rho_k \right)$  而不违背其他任何约束条件。此时, MT  $\bar{k}$  所降低的功率可以用于提高所有终端的 SINR,从而进一步减小下行传输时延,这与假设相矛盾。因此,问题(20)取得最优解时,对所有的  $k \in K$ ,  $m_d$  都相同。

其次,根据引理2,目标函数要求约束问题(20b)在最优解时等号成立,否则可以进一步提高  $\rho_k$  而降低下行传输的时延。

基于上述2个事实,我们证明了问题(20)与原问题(8)是等价的。

**命题4:**  $m_u, m_d$  均是  $\varepsilon_0$  的严格凸函数。

证明:引理2表明  $\varepsilon$  关于  $m$  严格单调减且严格凸,即  $\frac{\partial \varepsilon}{\partial m} < 0$ ,  $\frac{\partial^2 \varepsilon}{\partial m^2} > 0$ 。由反函数求导法则可知,  $\frac{\partial^2 m}{\partial \varepsilon^2} = - \left( \frac{\partial \varepsilon}{\partial m} \right)^{-3} \left( \frac{\partial^2 \varepsilon}{\partial m^2} \right) > 0$ , 所以  $m$  是  $\varepsilon$  的严格凸函数。那么,  $m_u$  是  $\varepsilon_0$  的严格凸函数,  $m_d$  是  $\varepsilon_k$

的严格凸函数。又  $\varepsilon_k = 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0}$  是  $\varepsilon_0$  的严格凸函数, 根据复合函数法则,  $m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0} \rho_k \right)$  是  $\varepsilon_0$  的严格凸函数。命题得证。

**命题 5:** 约束条件 (20b) 等价于  $\begin{pmatrix} \lambda_k I_{N_t} + A_k & A_k \hat{h}_k \\ \hat{h}_k^H A_k & \hat{h}_k^H A_k \hat{h}_k - \lambda_k \delta_k^2 - \sigma^2 \end{pmatrix} \geq 0$ , 其中  $I_{N_t}$  表示  $N_t \times N_t$  的单位矩阵,  $\lambda_k \geq 0$ ,  $A_k = \frac{1}{\rho_k} w_k w_k^H - \sum_{j=1, j \neq k}^K w_j w_j^H, k \in K$ 。

证明: 首先, 问题 (20b) 则等价于:

$$e_k^H A_k e_k + 2\Re(\hat{h}_k^H A_k e_k) + \hat{h}_k^H A_k \hat{h}_k - \sigma^2 \geq 0, \forall e_k, e_k - \delta_k^2 \leq 0. \quad (21)$$

那么根据 S-Procedure<sup>[12]</sup>, 问题 (21) 等价于  $\exists \hat{e}_k$  这使得  $\|\hat{e}_k\| < \delta_k$  (该条件显然成立), 且  $\exists \lambda_k \geq 0$ , 则有:

$$\lambda_k \begin{pmatrix} I_{N_t} & 0 \\ 0 & -\delta_k^2 \end{pmatrix} \geq \begin{pmatrix} -A_k & -A_k \hat{h}_k \\ -\hat{h}_k^H A_k & -\hat{h}_k^H A_k \hat{h}_k + \sigma^2 \end{pmatrix}, \quad k \in K. \quad (22)$$

因此, 命题 5 得证。

根据以上分析, 可进一步将问题 (20) 等价地表示为:

$$\min_{\varepsilon_0, \{\rho_k, w_k, \lambda_k\}_{k=1}^K} m_u(\varepsilon_0) + \max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0} \rho_k \right) \quad (23a)$$

$$s.t. \begin{pmatrix} \lambda_k I_{N_t} + A_k & A_k \hat{h}_k \\ \hat{h}_k^H A_k & \hat{h}_k^H A_k \hat{h}_k - \lambda_k \delta_k^2 - \sigma^2 \end{pmatrix} \geq 0, \forall k, \quad (23b)$$

$$\lambda_k \geq 0, \forall k, \quad (23c)$$

$$\sum_{k=1}^K \|w_k\|^2 \leq P_{\max} \quad (23d)$$

## 2 算法设计

对于问题 (23), 我们可以看

到  $\varepsilon_0$  与  $\{\rho_k, w_k, \lambda_k\}$  分属于独立的约束集, 因此可以利用交替优化的思想来获得问题 (23) 的解, 即通过迭代的方式更新  $\varepsilon_0$  与  $\{\rho_k, w_k, \lambda_k\}$  使得空中系统的空口传输时延最小, 其该解至少是一个局部最优解。

首先, 我们固定  $\varepsilon_0$ , 并且更新  $\{\rho_k, w_k, \lambda_k\}$ 。那么, 对于任意给定的  $\bar{\varepsilon}_0 \in (0, \varepsilon_{\max})$ , 问题 (23) 可以简化为:

$$\min_{\{\rho_k, w_k, \lambda_k\}_{k=1}^K} \max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \bar{\varepsilon}_0} \rho_k \right) \quad (24)$$

$$s.t. \quad (23b) \text{---} (23d). \quad (23b) \text{---} (23d).$$

由命题 3 可知, 该问题取最优解时, 所有用户的  $m_d$  都相同, 因此可以通过二分法确定最优的  $t = m_d(\bar{\varepsilon}_0, \rho_k)$ 。对于给定的  $t$ , 根据  $m_d$  关于  $\rho_k$  的单调性, 我们可以通过二分法得到对应于  $t$  的  $\rho_k$ <sup>[13]</sup>。其中, 对于  $t$  的搜索需要在每次迭代过程中判断下述问题的最小值是否满足 RSU 发射功率的约束:

$$\min_{\{w_k, \lambda_k\}_{k=1}^K} \sum_{k=1}^K \|w_k\|^2 \quad (25a)$$

$$s.t. \begin{pmatrix} \lambda_k I_{N_t} + A_k & A_k \hat{h}_k \\ \hat{h}_k^H A_k & \hat{h}_k^H A_k \hat{h}_k - \lambda_k \delta_k^2 - \sigma^2 \end{pmatrix} \geq 0, \forall k, \quad (25b)$$

$$\lambda_k \geq 0, \forall k. \quad (25c)$$

即如果公式 (25) 的最小值不超过  $P_{\max}$ , 则可以进一步降低  $t$ , 否则增大  $t$ 。而公式 (25) 可以通过半正定松弛 (SDR) 方法近似为一个半正定规划问题。令  $W_k = w_k w_k^H$ , 则公式 (25) 对应的 SDR 问题为:

$$\min_{\{W_k, \lambda_k\}_{k=1}^K} \sum_{k=1}^K \text{Tr}(W_k) \quad (26a)$$

$$s.t. \begin{pmatrix} \lambda_k I_{N_t} + \bar{A}_k & \bar{A}_k \bar{h}_k \\ \bar{h}_k^H \bar{A}_k & \bar{h}_k^H \bar{A}_k \bar{h}_k - \lambda_k \delta_k^2 - \sigma^2 \end{pmatrix} \geq 0, \forall k, \quad (26b)$$

$$\lambda_k \geq 0, W_k \geq 0, \forall k, \quad (26c)$$

其中,  $\bar{A}_k = \frac{1}{\rho_k} W_k - \sum_{j=1, j \neq k}^K W_j$ ,

$\forall k \in K$ 。根据文献<sup>[14]</sup>中的定理 3.2 可知, 公式 (26) 存在一组最优解  $\{W_k^*\}$  满足  $\sum_{k=1}^K \text{Rank}^2(W_k^*) \leq K$ 。因为  $W_k^*$  不能为零矩阵, 所以  $\text{Rank}(W_k^*) = 1, \forall k$ , 即对于公式 (26), SDR 是紧的, 通过求解公式 (26) 可以得到公式 (25) 的全局最优解。

其次, 我们固定  $\{\rho_k, w_k, \lambda_k\}$ , 更新  $\varepsilon_0$ 。对于给定的  $\{\rho_k, w_k, \lambda_k\}$ , 问题 (23) 可简化为:

$$\min_{\varepsilon_0 \in (0, \varepsilon_{\max})} m_u(\varepsilon_0) + \max_{k \in K} m_d \left( 1 - \frac{1 - \varepsilon_{\max}}{1 - \varepsilon_0} \rho_k \right) \quad (27)$$

由命题 4 可知,  $m_u, m_d$  是  $\varepsilon_0$  的严格凸函数, 则问题 (27) 的目标函数也是关于  $\varepsilon_0$  的凸函数; 因此, 可以在  $(0, \varepsilon_{\max})$  上利用黄金分割搜索的方法找到凸优化问题 (27) 的最优解, 使得传输时延最小。

综上所述, 我们提出**算法 1**来求解问题 (23)。

## 3 仿真

本节通过数值仿真分析传输时延  $m_u + m_d$  与 RSU 端天线数  $N_t$ 、下行最大发送功率  $P_{\max}$ 、可靠性指标  $\varepsilon_{\max}$ 、信道误差  $\delta_k = \delta, \forall k = 0, 1, \dots, K$  的关系。另



外我们假定  $K=4$ ,  $\hat{h}_k \sim CN(0,1)$ ,  $N_k=300$  bytes,  $\forall k=0,1,\dots,K$ ,  $\sigma^2=0.1$ ,  $P_0=20$  dBm。

图2刻画了不同误包率  $\varepsilon_{\max}$  下信道误差  $\delta$  对传输时延  $m_u+m_d$  的影响,其中RSU的最大发送功率  $P_{\max}=30$  dBm,天线数  $N_t=4$ 。从图中可以看出,传输时延随着  $\delta$  变大而快速变大,其中  $\delta=0$  对应于理想 CSI 的情况,优化算法可以进一步简化,具体见文献[11]。此外,传输时延随着可靠性指标的提高而严格增加。

图3刻画了RSU最大传输功率对传输时延的影响,其中  $N_t=4$ 。该结果表明,增大传输功率可以显著降低传输时延。与图2类似,相同的信道误差下,不同的可靠性指标导致传输时延存在几乎相似的性能差。

#### 4 结束语

本文中我们研究了存在信道误差条件下,通过上下行联合的优化设计保障 uRLLC 传输的空口端到端性能。为了达到超可靠传输的目的,我们利用有限码长容量公式近似刻画传输速率、时延和可靠性之间的关系;针对有界的信道误差,利用鲁棒优化技术研究并设计了最差情况下的系统波束设计、可靠性分配。研究表明,这一联合的端到端性能优化问题可以解耦为上下行可靠性的平衡问题;仿真结果表明,信道误差、发送功率对系统端到端性能都具有显著的

#### 算法1:基于交替优化的算法求解问题(23)

- 1 设迭代索引  $i=0$ ,  $\Delta=\infty$ , 迭代误差容许值  $\epsilon \geq 0$ ;
- 2 初始化  $\varepsilon_0^{(i)}$ ,  $m_{d,L}, m_{d,U}$ ,  $m_u^{(i)}$  和  $m_d^{(i)}$ ;
- 3 While  $\Delta \geq \epsilon$
- 4 求解问题(24):在  $[m_{d,L}, m_{d,U}]$  二分法搜索获得最优的  $m_d^{(i+1)}$ , 其中  $m_{d,L}, m_{d,U}$  的更新取决于问题(26)的最小值是否小于  $P_{\max}$ ; 并根据(26)的解更新  $\{\rho_k^{(i+1)}, w_k^{(i+1)}, \lambda_k^{(i+1)}\}$ ;
- 5 固定  $\{\rho_k^{(i+1)}, w_k^{(i+1)}, \lambda_k^{(i+1)}\}$ , 求解问题(27)以更新  $\varepsilon_0^{(i+1)}$ ;
- 6 根据  $\varepsilon_0^{(i+1)}$  更新  $m_u^{(i+1)}$ ;
- 7  $\Delta = |(m_u^{(i)} + m_d^{(i)}) - (m_u^{(i+1)} + m_d^{(i+1)})|$ ;
- 8  $i = i + 1$ ;
- 9 end

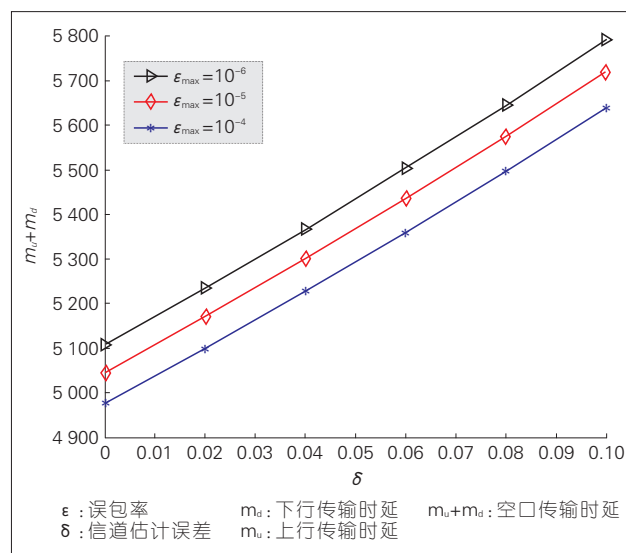


图2  
空口时延与信道估计误差的关系图

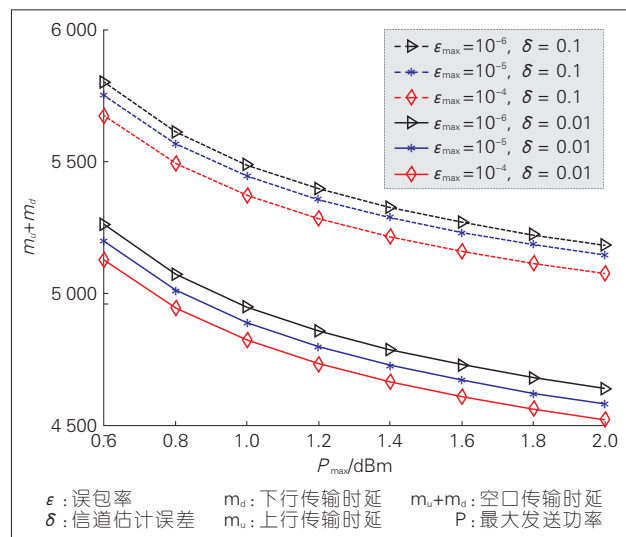


图3  
空口时延与最大发送功率的关系图

影响,因此,针对实际系统进行有效的信道估计和资源分配是 uRLLC 系统优化的核心问题。对于下一步工作,可以从 uRLLC 业务的重传技术,及其与 eMBB 业务的混合传输等方面进一步研究。

#### 参考文献

- [1] BERTENYI B. Summary After TSG-RAN#80 [EB/OL].(2018-06-05)[2019-01-22]. <https://www.brighttalk.com/webcast/15727/328827>
- [2] 3GPP. Study on Scenarios and Requirements for Next Generation Access Technologies: 3GPP TR 38.913 V15.0.0[S]. 2018
- [3] DURISI G, KOCU T, POPOVSKI P. Toward Massive, Ultrareliable, and Low-Latency Wireless Communication with Short Packets [J]. Proceedings of the IEEE, 2016, 104(9): 1711-1726
- [4] BENNIS M, DEBBAH M, POOR H V. Ultra-reliable and Low-Latency Wireless Communication: Tail, Risk and Scale [EB/OL]. [2019-01-22]. Available: <https://arxiv.org/abs/1801.01270>
- [5] SHARIATMADARI H, IRAJI S, LI Z X, et al. Optimized Transmission and Resource Allocation Strategies for Ultra-Reliable Communications[C]//2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC). USA:IEEE, 2016: 1-6. DOI:10.1109/PIMRC.2016.7794801
- [6] POLYANSKIY Y, POOR H V, VERDU S. Channel Coding Rate in the Finite Blocklength Regime [J]. IEEE Transactions on Information Theory, 2010, 56(5): 2307-2359. DOI:10.1109/tit.2010.2043769
- [7] AVRANAS A, KOUNTOURIS M, CIBLAT P. Energy-Latency Tradeoff in Ultra-Reliable Low-Latency Communication with Retransmissions [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(11): 2475-2485. DOI:10.1109/jsac.2018.2874143
- [8] HU Y L, SCHMEINK A, GROSS J. Blocklength-Limited Performance of Relaying under Quasi-Static Rayleigh Channels [J]. IEEE Transactions on Wireless Communications, 2016: 1. DOI:10.1109/twc.2016.2542245
- [9] YANG W, DURISI G, KOCH T, et al. Quasi-Static Multiple-Antenna Fading Channels at Finite Blocklength [J]. IEEE Transactions on Information Theory, 2014, 60(7): 4232-4265. DOI:10.1109/tit.2014.2318726
- [10] XU S F, CHANG T H, LIN S C, et al. Energy-Efficient Packet Scheduling with Finite Blocklength Codes: Convexity Analysis and Efficient Algorithms[J]. IEEE Transactions on Wireless Communications, 2016, 15(8): 5527-5540. DOI:10.1109/twc.2016.2561273
- [11] SHEN C, CHANG T H, XU H Q, et al. Joint Uplink and Downlink Transmission Design for URLLC Using Finite Blocklength Codes [C]//2018 15th International Symposium on Wireless Communication Systems (ISWCS). Portugal: ISWCS, 2018: 1-5. DOI:10.1109/ISWCS.2018.8491069
- [12] CHI C Y, LI W C, LIN C H. Convex Optimization for Signal Processing and Communication [M]. British: Taylor & Francis Group, CRC Press, 2017. DOI: 10.1201/9781315366920
- [13] XU Y Q, SHEN C, CHANG T H, et al. Energy-Efficient Non-Orthogonal Transmission under Reliability and Finite Blocklength Constraints[C]//2017 IEEE Globecom Workshops (GC Wkshps). USA: IEEE, 2017:1-6. DOI:10.1109/GLOCOMW.2017.8269136
- [14] HUANG Y W, PALOMAR D P. Rank-Constrained Separable Semidefinite Programming with Applications to Optimal

Beamforming [J]. IEEE Transactions on Signal Processing, 2010, 58(2): 664-678. DOI:10.1109/tsp.2009.2031732

#### 作者简介



**成晶**, 北京交通大学轨道交通控制与安全国家重点实验室在读博士生; 研究方向为高可靠低时延无线通信系统的优化设计。



**沈超**, 北京交通大学轨道交通控制与安全国家重点实验室副教授; 主要研究方向为高可靠低时延无线通信、面向高速移动场景的无线通信系统设计、无人机通信等; 先后主持 2 项国家自然科学基金项目, 参加多项国家自然科学基金和科技部项目; 已发表论文 30 余篇, 其中 SCI 检索论文 10 余篇。



**夏树强**, 中兴通讯股份有限公司高级工程师; 现主要从事机器通信、高可靠低时延、跨层优化、灵活广播多播等领域的研究; 深圳市国家级领军人才, 并获中国专利金奖、“广东发明人”等荣誉, 主持完成国家重大专项 1 项; 向 3GPP、IEEE 等标准组织输出提案 100 余篇, 已发表论文 10 余篇, 共申请专利 70 余篇。



# B5G: 泛在融合信息网络

## B5G: Ubiquitous Fusion Information Networks

张平/ZHANG Ping

(北京邮电大学, 北京 100876)  
(Beijing University of Posts and Telecommunications, Beijing 100876, China)

DOI: 10.12142/ZTETJ.201901009

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190131.1559.012.html>

收稿日期: 2018-12-24

网络出版日期: 2019-01-31

**摘要:** 延续“线”—“面”—“体”的演进趋势, 超5代移动通信系统(B5G)继续提高通信速率, 拓展通信空间, 完善通信智慧, 演进为泛在融合信息网络。B5G使用更高的频段作为信号载体, 数据速率达到太比特每秒量级。伴随网络性能的增强, B5G的适用空间拓展至陆海空天。与以往移动通信系统不同, 人工智能(AI)成为B5G性能提升的强劲引擎。基于AI的干扰管理、深度学习智能信号处理以及太赫兹技术成为物理层关键技术。基于极化码的中继、多天线、多址技术是传输层关键技术。基于AI的移动网络架构、面向人机物泛在融合的全析网络架构以及认知增强与决策推演的智能定义网络架构等方式的新架构被应用于网络层。

**关键词:** 泛在融合信息网络; B5G; THz; AI

**Abstract:** Following the “line-plane-cube” evolution trend, Beyond Fifth Generation (B5G) in mobile communication system aims at improving communication data rate, extending communication dimensions, implementing communication intelligence and is on the way to evolve into ubiquitous fusion information networks. Higher spectrum bands are used in B5G and the peak data rate could be T bits per second. With the enhancement of network ability, the support scenarios of B5G includes the land, the sea, the sky and the aerospace. Different with the existing generations, artificial intelligence (AI) becomes a powerful engine for B5G. The key technologies in the physical layer include the AI-based interference management, deep learning intelligent signal processing and THz technology. The transport layer uses the polar code-based relay, multiple input multiple output (MIMO) and multiple access technologies. The AI-based mobile network framework, the ubiquitous fusion fully-analyzed network framework for human-machine-things and the intelligent definition network framework for cognitive enhancement and decision deduction are used in the network layer.

**Key words:** ubiquitous fusion information networks; B5G; THz; AI

### 1 信息通信网络发展历程及5G面临的挑战

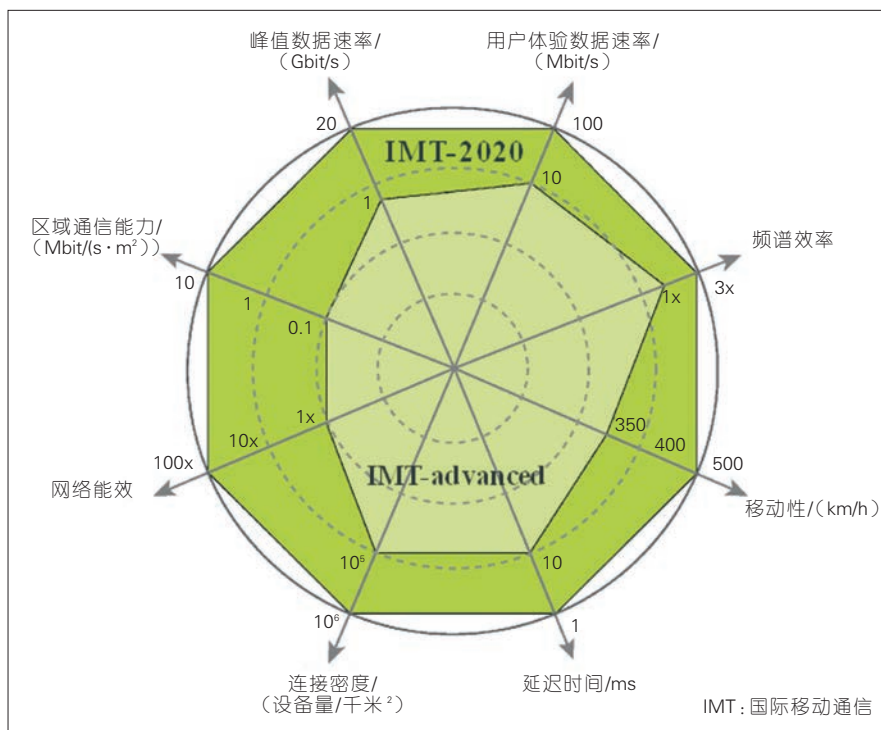
1948年香农发表了《通信的数学原理》, 提出了信息熵的概念, 奠定了信息论和数字通信的理论基础。70年来, 在香农信息论的指导下, 现代无线通信系统从无到有, 不断取得突破性发展, 深刻改变了人们的生活。1978年美国贝尔实验室成功研制出了第1个移动蜂窝电话系统——先进的移动电话系统(AMPS), 它标志着第1代移动

通信系统(1G)正式登上历史舞台<sup>[1]</sup>。20世纪七八十年代, 世界各国纷纷建立起了自己的第1代移动通信系统。由于采用模拟蜂窝和频分多址(FDMA)技术, 1G的容量十分有限, 并且通话质量不高, 不能提供数据业务和漫游服务<sup>[1]</sup>。为解决上述问题, 在20世纪90年代, 以数字技术和时分多址(TDMA)技术为主体的第2代移动通信系统(2G)研制成功<sup>[2]</sup>。与1G相比, 2G具有通话质量高、频谱利用率高和系统容量大等优点; 但是

它对定时和同步精度的要求高, 并且系统带宽有限<sup>[2]</sup>, 无法承载较高数据速率的移动多媒体业务。为了支持和实现较高速率的移动宽带多媒体业务, 以码分多址(CDMA)技术为核心的第3代移动通信系统(3G)<sup>[3-7]</sup>应运而生。相比于前2代移动通信系统, 基于Turbo码和CDMA技术的第3代移动通信系统具有更大的系统容量、更好的通话质量和保密性, 并且能够支持较高数据速率的多媒体业务。然而, 仍受其带宽限制, 3G无法支持超

高清视频等更高质量的多媒体业务<sup>[3-7]</sup>。为了追求更大的系统容量和更高质量的多媒体业务,基于正交频分复用多收发天线(OFD-MIMO)技术和空分多址(SDMA)技术的第4代移动通信系统(4G)应需而来<sup>[8-13]</sup>。与3G通信系统相比,4G通信系统数据传输速率更快,并且它能够更好地对抗无线传输环境中的多径效应,系统容量和频谱效率得到大幅提升。随着硬件工艺的提升和成本的下降,无线设备能力不断增强,数量也持续增加。移动网络承载的数据量呈现爆炸式增长的态势。伴随着“万物互联”的提出,4G急需满足支持超高质量的多媒体业务以及高可靠、低时延、低能耗、大连接等新需求。第5代移动通信系统(5G)研究拉开序幕,并逐步从标准走向实现<sup>[14-22]</sup>。

图1给出了5G相较于与IMT-Advanced通信系统(也即实际满足4G标准的商用系统)能力的增强。从图中可以看出5G考虑了更多的性能维度提升,包括:(1)在峰值数据速率方面,峰值数据速率提升了20倍,由1 Gbit/s提升至20 Gbit/s;(2)在用户体验数据速率方面,就广域覆盖而言,城区和城郊用户有望获得100 Mbit/s的用户体验数据速率,在热点地区,用户体验数据速率值有望提升至1 Gbit/s;(3)在频谱效率方面,频谱效率较4G提升3倍;(4)在移动性方面,支持更高速的移动,专门为



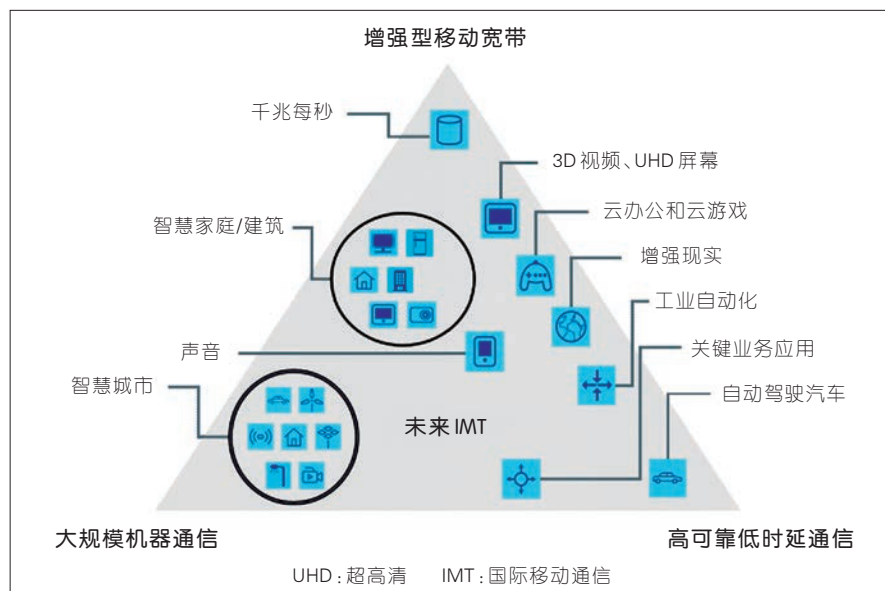
▲ 图1 5G较4G在关键技术方面增强的示意图<sup>[23]</sup>

高速铁路设计服务,由350 km/h提升至500 km/h;(5)在延迟时间方面,支持极低延迟要求的服务,延迟时间降低了10倍,由10 ms降低至1 ms;(6)在连接密度方面,支持更多数量的设备连接,适用于大规模机器类型通信场景,连接密度由10<sup>5</sup>设备量/千米<sup>2</sup>提升至10<sup>6</sup>设备量/千米<sup>2</sup>;(7)在网络能效方面,较上一代提升100倍;(8)在区域通信能力方面,区域通信能力提升了100倍,由0.1 Mbit/(s·m<sup>2</sup>)提升至10 Mbit/(s·m<sup>2</sup>)。基于上述8个方面能力的增强,5G网络开始具备渗透垂直行业的能力,支持的应用场景涵盖增强型移动宽带(eMBB)、超可靠低时延通信(uRLLC)以及大规模机器通信(mMTC)3大场景。图2给出

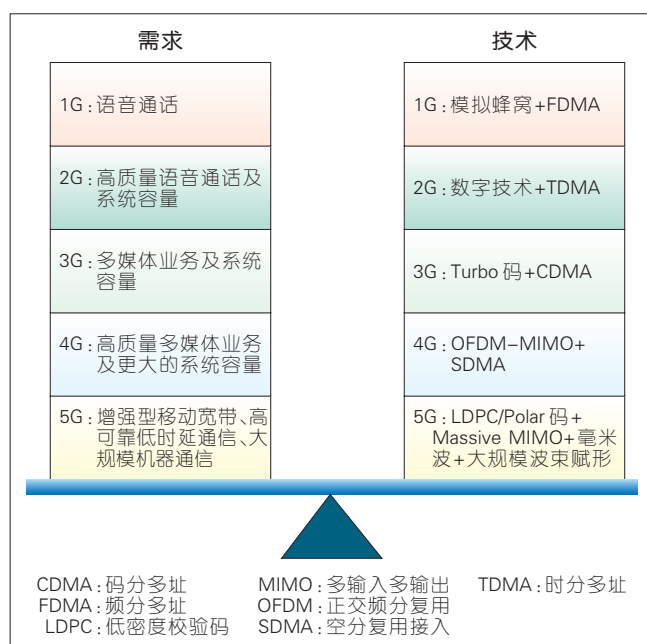
了5G的3大场景典型支持业务,包括4 K/8 K超高清视频、增强现实(AR)/虚拟现实(VR)、全息技术、智能终端、智慧城市、智慧工业、智慧家庭、智慧农业、无人驾驶、车联网、智慧医疗等。为了实现系统性能的增强,5G借助于毫米波频段,并采用大规模MIMO赋形技术弥补毫米波频段的衰减。采用了低密度校验码(LDPC)编码和Polar码分别作为数据和控制信道编码。

如图3所示,每一代通信系统取得成功的原因在于完美平衡了天平两侧的通信需求和技术能力。一旦一方打破这个平衡,就会促使移动通信系统演进到新的平衡状态。1G到5G的演进呈现如下规律:第一,支持





▲ 图2 5G潜在的业务、服务及应用示意图<sup>[23]</sup>



▲ 图3 1G到5G的需求及其关键技术发展示意图

场景逐步多样化,从简单的语音演进至3大场景典型业务;第二,通信速率每代有约1 000倍提升,从2G的千比特每秒量级提升至5G的吉比特每秒量级。按照上述演进规律进行推断,超5代移动通信系统(B5G)将进一步扩展支持的通信场景,同时数

据速率将达到太比特每秒。当前,5G基本上满足了陆地通信系统面向个人终端的基本通信需求。随着国家信息疆域战略扩展部署,5G通信系统尚不能满足全方位、立体化的多域覆盖,尤其在空天通信、空地通信,及海洋通信的能力严重不

足;因此,5G之后的移动通信系统面临如下挑战:

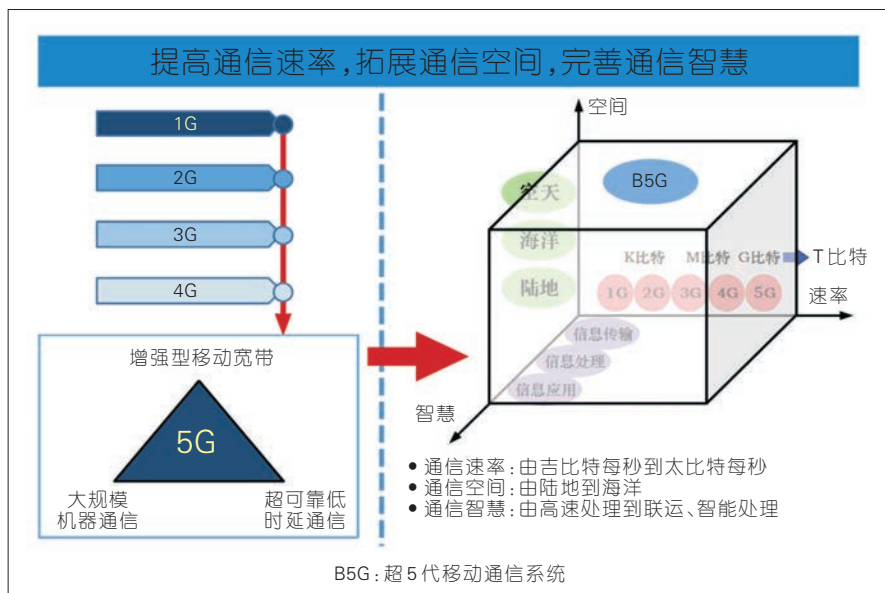
(1)数据的速率将难以达到1 Tbit/s量级以上;

(2)多域网络之间相对独立,没有完整的协同传输框架,难以满足全方位、立体化的多域、跨域传输及覆盖,空天通信、空地通信及海域通信能力严重不足;

(3)随着大数据、互联网、智慧城市、智慧产业和信息物理与社会融合空间的兴起,对网络的创新、智慧、安全融合提出了更高的要求,例如,情景再现与融合、智慧城市神经网络、智能无人网络等。

## 2 B5G 演进: 泛在融合信息网络

目前移动通信系统的演进趋势可以归纳为:由“线”到“面”的演进趋势。线是指演进所围绕的增强移动宽带性能这条主线,即每一代移动通信系统演进的首要目标是大幅提升数据传输速率和网络容量。而面是指从4G到5G的演进逐步开始考虑支持多种业务需求矛盾的场景,而不仅限于增强型移动宽带,例如,5G支持的高可靠低时延通信以及大规模机器通信业务等。基于已有演进规律,B5G通信系统将由一维的线、二维的面拓展演进为三维的体,如图4所示。具体而言,这个三维的“体”包括:速率维度、空间维度以及智慧维度。通过3个维



▲ 图4 “线”—“面”—“体”的演进趋势示意图

度的不断完善, B5G 通信系统最终演进为泛在融合信息网络。泛在融合信息网络旨在进一步提高通信速率, 达到 1 Tbit/s 量级以上; 进一步拓展通信空间, 由目前的陆地覆盖拓展至海洋、天空、太空场景下的多域和广域覆盖; 进一步加强和完善通信智慧, 由目前单一设备的智能处理演进至多设备、多网络之间的协同跨域联动智能处理, 并且从信息传输、处理及应用层面进一步加强和深化通信智慧。

### 3 泛在融合信息网络中的使能关键技术

泛在融合信息网络包括泛在化、社会化、智慧化、情境化等新型应用形态与模式, 蕴含“网络资源按需即用”的核心技术特征。现有的 5G 网络技术难以在信息广度、信息速度及信息深度上支持人、机、物三元空间

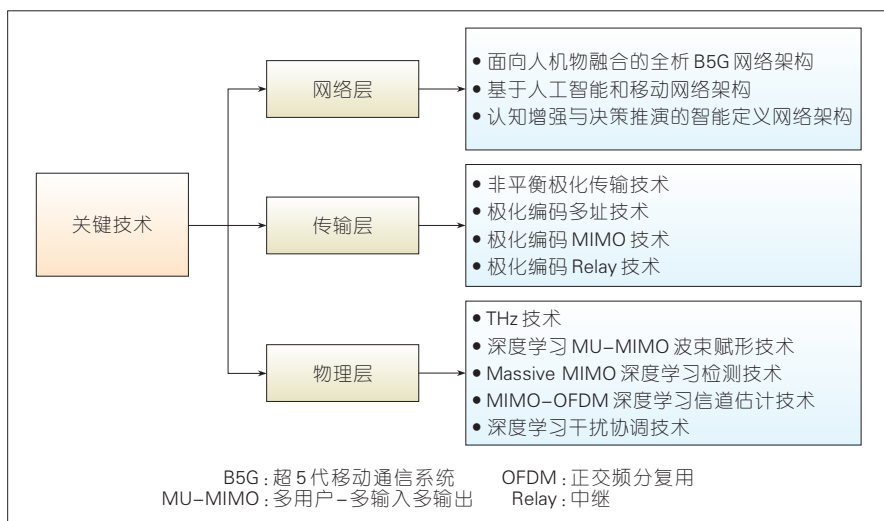
的深度融合与应用, 需要在网络架构和核心技术方面加以突破, 支撑未来应用的业务需求。如图 5 所示, 为了满足 B5G 网络的泛在化、社会化、智能化、情景化、广域覆盖及多域融合的需求, 结合 5G 网络技术的发展以及演进过程, 我们从理论及技术等多个方面探索并研究 B5G 网络可实现的物理层关键技术、传

输层关键技术以及网络层关键技术。

#### 3.1 物理层关键技术

##### (1) 太赫兹技术。

为了支持超高速数据传输, B5G 系统必然采用超宽带体制, 太赫兹通信被认为是未来 B5G 通信的潜在关键技术之一。由于太赫兹频段相比微波频段, 带宽更宽, 可提供数十吉比特每秒甚至更高的无线传输速率。同时, 其波束窄, 方向性更好, 还可采用扩频、调频技术<sup>[24-27]</sup>, 实现更好的通信保密性和抗干扰能力, 因此, 普遍认为太赫兹适合于中、近距离通信或太空无线通信。然而要研究太赫兹频段在移动通信的可行性及相应技术, 首要任务就是要掌握太赫兹频段在多样环境中的无线信道传播特性。由于传播环境中分子共振引起的能量损失可能会引起太赫兹波在传播过程中经历极大的衰减, 因此, 研究雨、雾,



▲ 图5 泛在融合信息网络中的关键技术

甚至空气中的水蒸气对太赫兹传播规律的影响,对其通信的可覆盖范围具有极大的意义。另外,由于太赫兹波段的粒子性强,穿透能力低,在传播环境中的穿透和反射特性与微波频段的规律有较大区别,传统模型难以准确刻画,因此,有必要研究太赫兹在受到不同材质阻挡的情况下反射和穿透的特性。太赫兹信号带宽一般在吉赫兹以上,系统在时延域的分辨率达到纳秒级。同时太赫兹由于波长短,天线尺寸极小,一般会组成超大规模的天线阵列,从而使得多径在时延和空间角度方向的可分离程度极高,而传统信道模型的分辨精度无法支持;因此,对太赫兹频段的超大规模天线和超大带宽的信道特性研究是其在未来 B5G 系统中具体采用何种编码、调制等一系列技术的重要基础。

(2)深度学习智能信号处理技术。

B5G 移动通信是多用户、多小区、多天线、多频段的复杂传输系统,信号接收与检测是高维优化问题。最优的最大似然(ML)或最大后验检测(MAP)是指数复杂度算法,性能优越但难以普遍应用。深度学习理论另辟蹊径,通过大量离线训练,获得高性能的深度神经网络模型,从而逼近 ML/MAP 检测。深度学习为最优信号检测理论提供了新的研究思路。首先,深入分析多用户 MIMO 的波束赋形信

号优化问题<sup>[28]</sup>,以卷积神经网络(CNN)构建优化模型,研究具有普适性与通用性的波束赋形算法,并建立算法的收敛性分析和理论。其次,分析大规模 MIMO 信号特征,设计深度信号检测算法网络,分析算法收敛性,构建高性能、低复杂度的检测算法体系。进一步分析 MIMO-OFDM 信道模型特征,以循环神经网络(RNN)模型为指导,设计高性能的信道估计算法,适应 B5G 移动通信的动态时变特征。最后,利用 CNN 与 RNN 组合模型,对多小区、复杂干扰场景的信号样本进行训练,获取信号的高维度特征,设计通用普适的干扰协调深度学习模型。

### 3.2 传输层关键技术

#### (1)非平衡极化传输。

极化码是信道容量可达的新型编码,已经被接纳为 5G 标准。极化设计思想就是利用编码与信号传输的联合优化,充分放大链路传输中的可靠性差异,最终达到通信系统的非平衡态优化。这种思想是方法论的革新,可以应用于 B5G 移动通信的各种场景:多址接入、广播、中继、MIMO 等。在各种场景下,极化传输都可严格证明达到相应的容量极限。

(2)基于极化码的多址、MIMO 及中继(Relay)技术。

目前,非正交多址接入(NOMA)已成为 5G 通信系统的代表性技术。NOMA 充分利用

多用户自由度,有效提升多址接入系统的容量,可以预见它也将成为 B5G 移动通信的核心技术。将极化编码引入非正交多址系统,需要深入分析 NOMA 的系统结构,从广义极化的观点出发,优化信道极化分解方案。主要研究与现有多址接入方式的组合优化,例如,研究多用户极化码与非正交多址的混合接入技术,简化信号发射机结构,优化用户、子载波的选择准则,设计低复杂度的检测译码算法。此外,MIMO 系统引入了空间维度,为通信系统优化提供了更多的空间自由度,但也对信道极化码的设计带来了诸多挑战,针对 MIMO 系统的 2 种典型结构:空间复用/预编码与空间调制,针对性研究了极化编码 MIMO 传输的优化方案,逼近单小区 MIMO 传输的容量极限。最后,在 Relay 系统中,由于引入了 Relay 节点,为信源到信宿链路的传输速率/可靠性提升提供了更多的优化空间。我们重点研究基于极化码构造的协作编码中继方案:解码转发(DF)与压缩转发(CF)中的极化码实用构造方案。

### 3.3 网络层关键技术

我们提出 3 种网络层架构:

(1)面向人、机、物融合的全析 B5G 网络架构;(2)基于人工智能的移动网络架构;(3)认知增强与决策推演的智能定义网络架构。其中,(1)为总体架构,



是面向人、机、物的全方位、立体化的泛在融合网络架构,而(2)和(3)可以作为(1)在人工智能以及机器学习等多个方面的增强型架构,起到扩展及补充等作用。

(1)面向人、机、物融合的全析B5G网络架构。

B5G网络架构将支持以人为中心的社会空间、以机器连接人的信息空间、以物体为载体的物理空间等三元空间的深度融合,从而实现“物质世界信息化”到“信息世界物质化”的转变。通过原理性抽象建立泛在化的人、机、物融合要素,研究通过人、机、物三元空间的语义衔接、业务适配、协作编排,构建一套面向人、机、物三元空间的信息传输、边缘智能、协同计算等B5G网络架构,支撑人、机、物要素跨界融合,实现陆海空天一体化的全析网络发展。我们需要研究B5G网络在设备功能趋同基础上的去中心化无线接入体制,并重点研究在无线接入网络控制平面与数据平面基础上引入计算平面后的协议结构、计算任务及存储任务的分布式处理,以及各设备计算能力、存储能力、能量信息及位置等信息的高效获取与交互机制等。我们还需要进一步研究不同层级设备在功能趋同基础上的去中心化组网关键技术;研究终端协同通信、终端协同计算、终端协同存储与终端协同供能等关键技术,支持终端对无线网络的全

面协同,实现去中心化的通信、计算、存储及供能的分布式服务,以用户的业务感受质量为核心,灵活利用多连接技术,提供针对连接的专属服务供给及质量保障。此外,与传统基于数据流、控制流和状态空间的经典融合技术不同,面向B5G网络的人、机、物融合技术将构建人、机、物多维、多空间资源的统一表示,人、机、物一体化融合的模式构造和执行。我们需要研究面向人、机、物融合的统一知识本体构建、表示、推理和学习等适配机制;研究基于网络多维可编程的人、机、物融合组装方法;研究人、机、物融合的状态监测控制、同步控制、一致性检查等网络容错机制。将传统集中、局部、层次式的网络管理与调度演进至人、机、物融合的三元网络协同管理,支持边缘智能和适配优化,实现全网资源的多级协同调度。我们需要针对人、机、物融合要素的多样性,研究基于通信协议特征分析的接入融合技术、基于无线信道状态的跨协议融合技术;针对人、机、物融合的资源不确定性,研究资源的需求感知和预测技术,基于边缘智能的资源管理策略,实现协同通信、协同计算、协同存储,提高服务质量。B5G网络需要满足更加多样化的业务需求,同时无线环境、业务环境、网络环境将更加复杂,因此需要研究基于智能推荐的网络适配机理和组网技术。

(2)基于人工智能的移动网络架构。

如果通信网络缺乏对大数据的科学利用,那么它是无法满足智能时代的需求的。目前,业务、网络和服务融合体系缺乏对多维度数据的共享与合理利用,导致业务网络无法自主依据不同类型业务的流量选择路径和调整带宽分配模式,不能满足智能时代用户的需求。人工智能技术的突破是网络智能优化的重要技术基础,同时,软件定义网络(SDN)技术的发展为网络智能优化提供了海量基础数据与架构支持<sup>[29]</sup>。如何利用软件定义的可定制特性,设计一套新的业务网络是未来无线网络亟待解决的关键问题。未来异构无线网络拓扑复杂多变,如何高动态地采集网络信息并大规模合理布置SDN控制器,对于网络整体性能至关重要。目前SDN协议中流量基本特征较少,流量识别困难,有必要研究流量自身的统计特征、标量特征,基于业务特征提取的高级特征,设计新的深度学习识别算法,为实现网络资源智能分配提供依据。网络中存在大量的历史数据,业务网络应自主学习资源分配模式,实时动态地加载最优的业务调度方案,充分发挥SDN动态应用策略和管理网络的能力。基于人工智能技术实现软件定义的业务网络是一种崭新的视角,其研究对于未来机器学习和网络发展具有重大的科学意义。未



来业务网络与物联网的发展急需人工智能技术提升服务能力和效率。

(3) 认知增强与决策推演的智能定义网络架构。

人工智能技术通过自学习状态和特征从而不断迭代优化输出结果,为解决复杂多变的未来 B5G 网络服务提供了新的解决思路;因此,迫切需要针对 B5G 网络面向认知增强与决策推演的智能定义网络基础理论开展研究工作,基于人工智能技术,使未来 B5G 网络自主认知网络环境变化以及服务特性,智能调度网络计算、缓存、带宽等资源,实现资源融合的动态决策推演,不断优化网络资源的适配与管理,持续改善网络服务质量。未来网络下的用户行为个性化,服务需求多样化,网络拓扑频繁变化,而传统网络中服务请求处理缺乏灵活性,数据传输缺乏自适应,导致网络资源浪费。在全面感知网络海量信息的基础上,利用人工智能技术深度挖掘用户、服务及网络之间的关联关系,建立面向服务特征—用户特性—网络环境的质量评估机制,并对网络中动态变化的多维信息进行快速、准确的预测,为资源融合的网络动态决策提供立体化的认知信息。此外,未来网络环境动态复杂,而传统网络中路由、缓存等策略静态僵化,导致网络资源与用户服务适配困难。在海量信息和动态变化的环境下,利用人工智能技术对立

体化的认知信息进行融合,对网络资源分布情况与变化规律进行动态建模分析,并结合集中管控的思想,实现网络中路由、缓存等策略的自适应推演。智能协调不同用户的服务需求,提高用户个性化服务体验质量,深度优化网络资源调配,提高网络资源利用率。同时,依据业务的服务质量、网络能力等需求,基于增强学习等人工智能技术,构建网络功能和服务等资源的编排调度方案;生成网络需求、演进态势、协同策略等相关规则,对网络资源调度有效性进行自主评估,获取正向的资源调度柔性进化方案,以最大化网络的效能和可靠性。

#### 4 结束语

B5G 是一个泛在信息融合网络。与已实现的通信系统相比,B5G 的数据速率达到太比特每秒量级,适用场景涵盖陆海空天,并且其网络的智能化水平进一步提升。得益于软硬件提升带来的大数据及 AI 技术的潜力释放,AI 将为 B5G 关键技术实现提供新思路。在物理层,采用基于学习理论的干扰管理、调制、信道估计波束赋形技术;在传输层,基于极化编码技术将会进一步提升传输效率;而在网络层,采用基于人机物融合、人工智能及机器学习等方式的新架构,进一步提升网络整体效率。B5G 的关键技术储备正日趋完善,实现的时间也将离我们越来越

接近。

#### 参考文献

- [1] 彭小平. 第一代到第五代移动通信的演进[J]. 中国新通信, 2007, 9(4): 90-92. DOI:10.3969/j.issn.1673-4866.2007.04.030
- [2] 曹淑敏. 第二代移动通信向第三代移动通信的演进[J]. 中兴新通讯, 1999, 5(4): 30-32
- [3] 尤肖虎, 曹淑敏, 李建东. 第三代移动通信系统发展现状与展望[J]. 电子学报, 1999, 27(21): 3-8. DOI:10.3321/j.issn:0372-2112.1999.21.002
- [4] 张平. Beyond 3G 移动通信系统关键技术[J]. 北京邮电大学学报, 2002, 25(3): 1-6. DOI:10.3969/j.issn.1007-5321.2002.03.001
- [5] 张平. WCDMA 移动通信系统[M]. 2004
- [6] 谢显中. TD-SCDMA 第三代移动通信系统技术与实现[M]. 北京: 电子工业出版社 2004
- [7] 黄标, 彭木根, 王文博. 第三代移动通信系统干扰共存研究[J]. 电信科学, 2004, 20(7): 34-39. DOI:10.3969/j.issn.1000-0801.2004.07.008
- [8] 刘海文, 李征帆. 第四代移动通讯技术的构架及其关键技术[J]. 微电子学与计算机, 2002, 19(3): 34-39
- [9] SAMPATH H, TALWAR S, TELLADO J, et al. A Fourth-Generation MIMO-OFDM Broadband Wireless System: Design, Performance, and Field Trial Results [J]. IEEE Communications Magazine, 2002, 40(9): 143-149. DOI:10.1109/mcom.2002.1031841
- [10] 任立刚, 宋梅, 乔强国, 等. MIMO+OFDM: 新一代移动通信核心技术[J]. 中国数据通信, 2003, 5(10): 102-105. DOI:10.3969/j.issn.1673-4866.2003.10.023
- [11] 何琳琳, 杨大成. 4G 移动通信系统的主要特点和关键技术[J]. 无线电技术与信息, 2005 (1): 34-36
- [12] 李钊, 韦玮. 第四代移动通信系统中的多天线技术[J]. 电子技术, 2005, 34(12): 39-42
- [13] BOUDREAU G, PANICKER J, GUO N, et al. Interference Coordination and Cancellation for 4G Networks[J]. IEEE Communications Magazine, 2009, 47(4): 74-81. DOI:10.1109/mcom.2009.4907410
- [14] 余莉, 张治中, 程方, 等. 第五代移动通信网络体系架构及其关键技术[J]. 重庆邮电大学学报(自然科学版), 2014, 26(4): 427-433, 560. DOI:10.3979/j.issn.1673-825X.2014.04.001
- [15] ANDREWS J G, BUZZI S, CHOI W, et al. What will 5G Be?[J]. IEEE Journal on Selected Areas in Communications, 2014, 32(6): 1065-1082. DOI:10.1109/JSAC.2014.2328098
- [16] 赵国锋, 陈婧, 韩远兵, 等. 5G 移动通信网络关键技术综述[J]. 重庆邮电大学学报(自然科学版), 2015, 27(4): 441-452. DOI:10.3979/j.issn.1673-825X.2015.04.003
- [17] 雷秋燕, 张治中, 程方, 等. 基于 C-RAN 的 5G 无线接入网架构[J]. 电信科学, 2015, 31(1): 106-115. DOI:10.11959/j.issn.1000-0801.2015028
- [18] 贾亚男, 岳殿武. 面向 5G 的小蜂窝网络研究综述[J]. 电讯技术, 2015, 55(11): 1296-1303. DOI:10.3969/j.issn.1001-893X.2015.11.019
- [19] IMT-2020(5G)推进组发布 5G 技术白皮书 [J]. 中国无线电, 2015(5): 1-15
- [20] GUPTA A, JHA R K. A Survey of 5G Network: Architecture and Emerging Technologies [J]. IEEE Access, 2015, 3: 1206-1232. DOI:10.1109/

- access.2015.2461602
- [21] 张平, 陶运铮, 张治. 5G 若干关键技术评述[J]. 通信学报, 2016, 37(7): 15–29. DOI: 10.11959/j.issn.1000–436x.2016130
- [22] AGIWAL M, ROY A, SAXENA N. Next Generation 5G Wireless Networks: A Comprehensive Survey [J]. IEEE Communications Surveys & Tutorials, 2016, 18(3): 1617–1655. DOI:10.1109/comst.2016.2532458
- [23] ITU–R. ITU–R: IMT Vision–Framework and overall Objectives of the Future Development of IMT for 2020 and Beyond: Recommendation ITU–R M.2083–0[R]
- [24] FRICKE A, KUMER T. Interference Study for THz Intra–Device Communication Systems with Multiple Links[C]//12th European Conference on Antennas and Propagation (EuCAP 2018). London: EuCAP, 2018:1–5
- [25] RAIMUNDO X, HAJJI M, KLEIN A, et al. Channel Characterisation at THz Frequencies for High Data Rate Indoor Communications[C]//12th European Conference on Antennas and Propagation (EuCAP 2018). London: EuCAP, 2018:1–2
- [26] POMETCU L, D'ERRICO R. Characterization of Sub–THz and mmwave Propagation Channel for Indoor Scenarios[C]//12th European Conference on Antennas and Propagation (EuCAP 2018). London: EuCAP, 2018:1–2
- [27] GONZALEZ–GUERRERO L, SHAMS H, FATADIN I, et al. Comparison of Optical Single Sideband Techniques for THz–Over–Fiber Systems [J]. IEEE Transactions on Terahertz Science and Technology, 2019, 9 (1): 98–105. DOI:10.1109/tthz.2018.2884736
- [28] QIN C, NI W, TIAN H, et al. Radio over Cloud (RoC): Cloud–Assisted Distributed Beamforming for Multi–Class Traffic[J]. IEEE Transactions on Mobile Computing, 2018: 1. DOI:10.1109/tmc.2018.2859395
- [29] LYU X, REN C, NI W, et al. Multi–timescale Decentralized Online Orchestration of Software–Defined Networks [J]. IEEE Journal on Selected Areas in Communications, 2018, 36(12):2716–2730. DOI: 10.1109/JSAC.2018.2871310
- 张平, 北京邮电大学教授、博士生导师, 网络与交换技术国家重点实验室主任, IEEE Fellow, 《通信学报》主编; 长期从事先进移动通信系统的研究与实践; 多次获国家技术发明奖、国家科技进步奖、光华工程科技奖、何梁何利基金科学与技术奖等; 出版专著 10 部, 发表论文 352 篇, 获授权发明专利 278 项。
- ←上接第 28 页
- [22] HUSSAMI N., KORADA S. B., URBANKE R. Performance of Polar Codes for Channel and Source Coding [C]//2009 IEEE International Symposium on Information Theory, 2009.USA: IEEE, 2009: 1488–1492. DOI:10.1109/ISIT.2009.5205860
- [23] NIU K, CHEN K, LIN J R. Low–Complexity Sphere Decoding of Polar Codes Based on Optimum Path Metric [J]. IEEE Communications Letters, 2014, 18(2): 332–335. DOI:10.1109/lcomm.2014.010214.131826
- [24] CHEN K, NIU K, LIN J R. An Efficient Design of Bit–Interleaved Polar Coded Modulation [C]//2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC), 2013.USA: IEEE, 2013: 693–697. DOI:10.1109/PIMRC.2013.6666225
- [25] Qualcomm. LDPC Rate Compatible Design Overview: 3GPP TSG R1–1610137[S]. Portugal, 2016
- [26] NIU K, CHEN K, LIN J R, et al. Polar Codes: Primary Concepts and Practical Decoding Algorithms [J]. IEEE Communications Magazine, 2014, 52(7): 192–203. DOI: 10.1109/mcom.2014.6852102
- [27] ARIKAN E, TELATAR E. On the Rate of Channel Polarization [C]//2009 IEEE International Symposium on Information Theory, 2009.USA: IEEE, 2009: 1493–1495. DOI:10.1109/ISIT.2009.5205856
- [28] KORADA S B, SASOGLU E, URBANKE R. Polar Codes: Characterization of Exponent, Bounds, and Constructions [J]. IEEE Transactions on Information Theory, 2010, 56(12): 6253–6264. DOI:10.1109/tit.2010.2080990
- [29] DAI J C, NIU K, SI Z W, et al. Polar–Coded Non–Orthogonal Multiple Access [J]. IEEE Transactions on Signal Processing, 2018, 66 (5): 1374–1389. DOI:10.1109/tsp.2017.2786273
- [30] DAI J C, NIU K, LIN J R. Polar–Coded MIMO Systems [J]. IEEE Transactions on Vehicular Technology, 2018, 67(7): 6170–6184. DOI:10.1109/tvt.2018.2815602
- 牛凯, 北京邮电大学教授、中国电子学会信息论分会副主任委员; 主要研究方向为信息论与信道编码, 在极化码的高性能译码算法以及实用化构造方面, 取得了诸多开创性成果, 设计的极化码高性能编译码算法成为 5G 标准主流方案; 先后主持多项国家自然科学基金项目与“863”项目; 发表论文 200 篇, SCI 检索 45 篇。



# 基于非完备大数据的业务预测

## Traffic Prediction with Incomplete Big Data

李建东/LI Jiandong, 盛敏/SHENG Min, 文娟/WEN Juan

(西安电子科技大学, 陕西 西安 710071)  
(Xidian University, Xi'an 710071, China)

DOI: 10.12142/ZTETJ.201901010

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190131.1032.004.html>

收稿日期: 2018-11-26

网络出版日期: 2019-01-31

**摘要:** 高效、精准预测无线网络业务数据,例如业务的到达率、用户数以及吞吐量等,将为网络提供用户的实时需求,是实现无线网络智能化的关键。然而,由于无线网络传输的不可靠性、采集设备故障、采样率低等原因,使得无线大数据具有不可避免的非完备性。将使系统丢失大量有用信息,从而给无线网络业务预测带来巨大挑战。为了应对该挑战,提出了基于非完备数据集的业务预测架构,从缺失值补充以及空时信息挖掘2个维度高效利用非完备数据集,提升预测精度,助力无线网络的智能化。

**关键词:** 业务预测; 智能无线网络; 非完备数据

**Abstract:** High efficient and accurate wireless traffic prediction, such as arrival rate, user account, and throughput, will provide users' real demand for network providers, which is the key for intelligent wireless networks. However, there exists incomplete nature for wireless big data because of the unreliable wireless transmission, the failure of data acquisition and low sample rate. This unique feature may make wireless networks lose massive useful information and bring great challenge for accurate traffic prediction. To meet this challenge, an incomplete data-based traffic prediction framework is proposed, leveraging the incomplete data set efficiently via filling the missing data and digging the temporal-spatial information.

**Key words:** traffic prediction; intelligent wireless networks; incomplete data set

无线网络正经历着从基于信息论的可靠传输到基于智能的高效通信的巨大变革<sup>[1]</sup>。智能无线资源管理是无线网络智能化的核心,其目标是通过无线资源的动态调配使网络资源与用户需求精准适配。具体来讲,网络将根据用户业务需求的时空分布,在相对较大的时间尺度内,合理地配置各区域网络资源,使网络资源结构最优化;在小时间尺度内,动态地为各用户分配网络资源,使系统资源利用率以及用户体验最大化<sup>[2]</sup>。由此可见,高效、准确预测无线网络业务需求,是无线网络实现智能

化的重要基础。

然而,网络结构的异构化和密集化使网络干扰异常复杂,加剧了无线网络传输的不可靠性,使得无线业务信息在传输过程中产生不可避免的丢失<sup>[3]</sup>。此外,数据采集设备故障或供电不足都将导致业务数据在收集过程中的缺失。最后,由于硬件设备限制导致的低采样率往往无法准确获得业务变化的重要信息。这些原因都将使无线网络业务数据具备如图1所示的非完备特性,即数据缺失或者无法反应业务变化趋势,从而给无线网络业务预测带来巨大挑战。

目前,对于缺失数据处理方法主要是根据已有数据的统计特性,如均值、中位数等,对缺失值进行补充<sup>[4]</sup>。此类方法对于统计规律比较强的数据有很好的作用;但是,当数据统计规律较弱时,其统计特性无法较好地反应数据本身特点。此时用其统计特性进行缺失值填充会引入大量噪声,从而影响数据预测效果。如表1所示,传统业务预测方法主要从时间以及空间2个维度,采用时间序列分析<sup>[5]</sup>、机器学习<sup>[6]</sup>,以及深度学习<sup>[8-9]</sup>等方法,对收集到的业务数据进行预测,但是基本没有考虑数据集

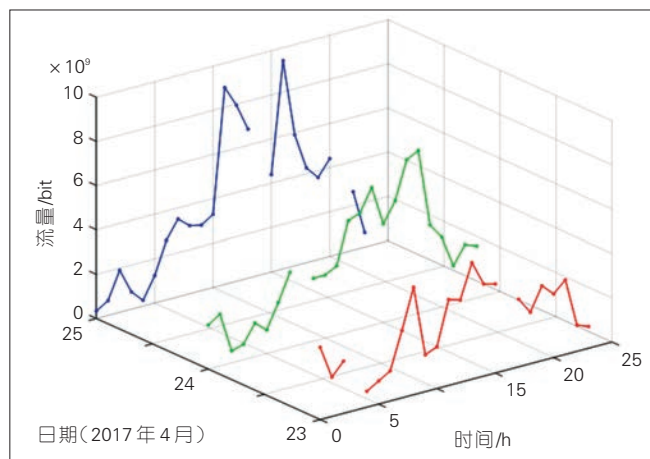


图1  
无线业务非完备特性  
示例

表1 典型业务预测算法分类总结

分类	典型预测算法	特点描述
时间序列分析法	ARIMA、AR、MA等	适用于线性、平稳过程,计算复杂度低
机器学习	Ridge、RF、Light GBM等	不局限于线性、平稳过程,计算复杂度适中
深度学习	LSTM、RNN等	所需训练集样本大,模型可解释性弱,计算复杂度高

AR: 自回归  
ARIMA: 差分整合滑动平均自回归模型  
Light GBM: 轻量级梯度提升机

LSTM: 长短期记忆网络  
MA: 滑动平均  
RF: 随机森林  
RNN: 循环神经网络

的非完备特性对业务预测带来的影响。

## 1 非完备海量数据业务预测

为了应对非完备数据给业务预测带来的挑战,本文中我们提出了如图2所示的基于非完

备无线大数据的业务预测架构,从缺失值填充、时空信息挖掘2个维度,高效利用非完备数据集,助力无线网络智能化。

简单来讲,当预测数据规律性较强时,例如办公楼以及住宅区域的业务数据呈现明显的“潮

汐现象”,我们根据待预测数据的统计特性对缺失值进行补充,然后选取合适的预测算法对待预测数据直接进行预测。当预测数据规律性较弱时,例如交通枢纽区域等业务数据流动性强、规律弱,如果仍根据其统计特性对缺失值补充,将会引入大量噪声;因此我们直接将缺失值丢弃,并充分利用空间维度信息进行数据挖掘,对待预测数据进行间接预测。

首先,我们采用时间序列分解法,将待预测数据分解为规律项和随机项,并根据规律项占业务量比值的大小,将待预测数据分为规律性强或弱2种情况。具体做法为:将待预测数据  $x = \{x_1, x_2, \dots, x_n\}$  (其中  $x_t, 1 \leq t \leq n$  表示第  $t$  时刻待预测的业务量,例如用户数、流量等)分解为周期项  $p_t$ 、趋势项  $m_t$  以及随机项  $r_t$ ,并且将周期项与趋势项的和称为规律项  $y_t = p_t + m_t$ 。当规律项与业务量的比值高于某一门

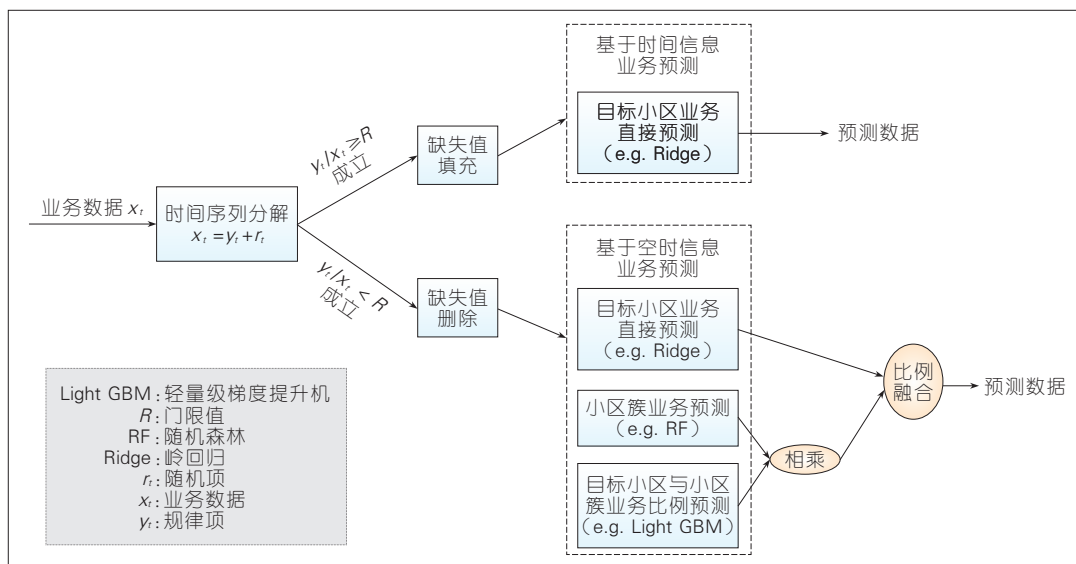


图2  
基于非完备无线大数据的  
业务预测架构



限值  $R$  时,待预测数据规律性较强;反之,当规律项与业务量的比值低于某一门限值  $R$  时,待预测数据规律性较弱。

当待测数据规律性较强时,其历史数据的统计特性例如均值,可较好地反映待测数据规律;因此,我们可利用缺失值历史信息的均值,对其进行填充,扩充训练样本数。我们可以将填充好后的数据直接输入所选择的预测算法,对待预测数据直接进行预测。

当待测数据规律性较弱时,其历史数据的统计特性无法较好地反映待测数据规律。如果对缺失值进行强行填充,将会引入较多噪声,从而影响预测精度。此时,我们将缺失值直接删除,确保使用数据的真实性。经过研究发现,在无线网络中,即使单小区的业务规律性较弱,由多个小区构成的小区簇的业务规律性一般都很强。因此,可以充分挖掘相邻小区的空间信息,先对小区簇的业务总量进行预测,然后再对目标小区业务与小区簇业务比值进行预测,最后将这2部分的预测值相乘,即可得到基于空间信息获得的目标小区待测业务量。为了进一步提升预测精度,我们采用“提升”(boosting)算法的基本思想,即设计多个好而不同的预测方法对同一问题进行预测,并将其结果进行融合,通过模型和数据的分集增益提升预测精度。为此,我们采用与基于空间信息预测

模型不同的预测方法对删除缺失值后的数据直接进行预测。最后,将预测结果与基于空间信息的预测结果进行有机融合,便可利用模型和数据的分集增益提升预测精度。

## 2 仿真设计与分析

为了验证本文提出的基于非完备大数据业务预测架构的有效性,我们采用校园网实测数据对各个区域各时间段的用户数进行预测。

为了判断待预测数据规律性的强弱,我们假设门限值  $R=0.8$ 。对于规律性比较强的业务数据,我们采用均值对相应缺失值进行补充,并使用补充后的数据集作为训练数据集,采用岭回归(Ridge)方法对其进行直接预测。图3对比了对缺失值进行均值填充和缺失值删除后的预测效果。为此,我们将获得的相对完整的数据看做实验中的“完备”数据集,然后在人为随机删除部分数据进行验证。从图3中可以看出,当待测数据规律性

较强时,当缺失值比例不大时,采用均值补充可以有效提升预测精度。此外,删除某些数据时,例如异常值,也可提升预测精度。因此,我们在对数据进行预测前,要先分析数据的特性,并根据数据的特性进行相应的处理。

对于规律性相对较弱的业务数据,我们先将缺失值删除,然后使用删除缺失值后的数据集作为训练数据集,并采用基于时空信息的预测方法对其预测。具体来讲,分别采用Ridge、随机森林(RF)以及轻量级梯度提升机(Light-GBM)方法对目标小区用户数、小区簇用户数以及目标小区和小区簇用户数的比例进行预测,然后将预测出的小区簇用户数和相应比例相乘,所得结果与直接预测的用户数以合适比例融合,得出最后的目标小区用户数预测值。图4对比了仅基于时间信息的业务预测与基于时空信息的业务预测精度。从图中我们可以看出,借助于空间信息可有效提高业务预

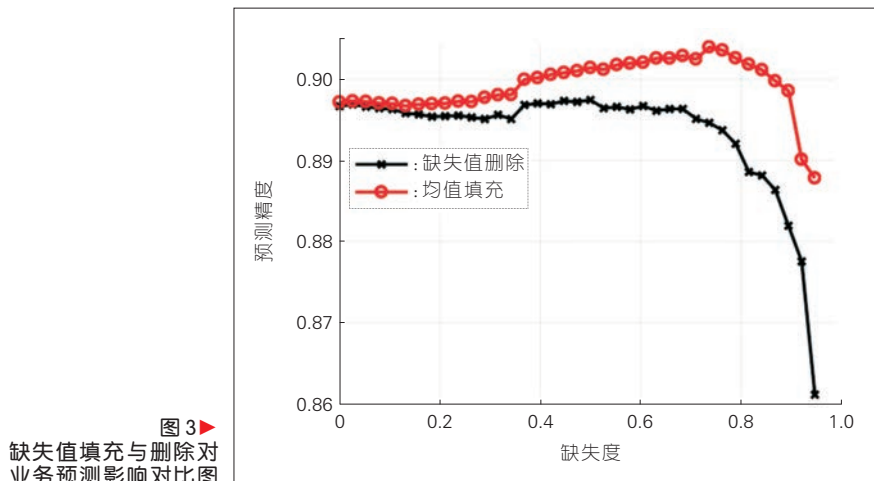


图3 缺失值填充与删除对业务预测影响对比图

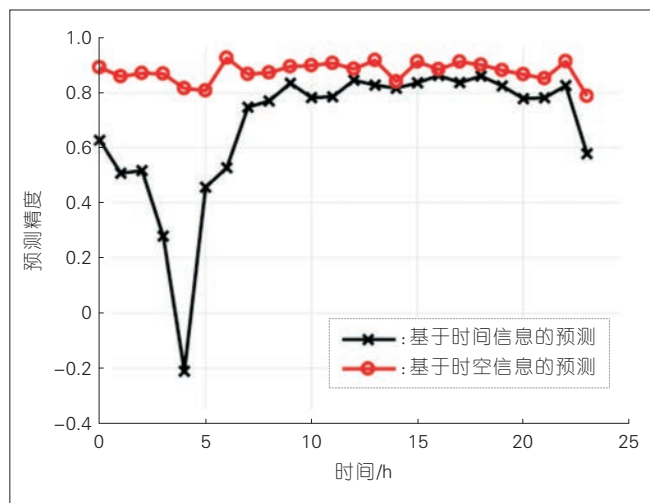


图4  
基于时间与时空信息  
业务预测对比图

测精度。

### 3 结束语

基于无线大数据,结合人工智能算法,将使无线网络的设计、管理与优化更加自动化、智能化与智慧化。然而,无线网络信道传输的不可靠性、业务多样性以及网络结构密集化、异构化等特点使得无线大数据呈现非完备性、空时大尺度变化等特点,为无线大数据挖掘以及人工智能算法应用与结合带来巨大挑战。本文中,我们提出了基于非完备无线大数据业务预测的基本架构,对非完备无线大数据的应用具有重要启发意义。在无线网络智能化的关键时期,仍需要我们不断探索如何针对无

线网络特异性,最大限度地挖掘无线大数据的价值并设计相应的智能算法。

#### 参考文献

- [1] LI R P, ZHAO Z F, ZHOU X, et al. Intelligent 5G: When Cellular Networks Meet Artificial Intelligence [J]. IEEE Wireless Communications, 2017, 24(5): 175–183. DOI: 10.1109/mwc.2017.1600304wc
- [2] 张琰, 盛敏, 李建东. 大数据驱动的“人工智能”无线网络[J]. 中兴通讯技术, 2018, 24(2): 2–5
- [3] LIU J Y, SHENG M, LIU L, et al. Interference Management in Ultra-Dense Networks: Challenges and Approaches [J]. IEEE Network, 2017, 31(6): 70–77. DOI:10.1109/mnet.2017.1700052
- [4] SESSA J, SYED D. Techniques to Deal with Missing Data[C]//2016 5th International Conference on Electronic Devices, Systems and Applications (ICEDSA). United Arab Emirates:ICEDSA, 2016: 1–4. DOI:10.1109/ICEDSA.2016.7818486
- [5] XU F L, LIN Y Y, HUANG J X, et al. Big Data Driven Mobile Traffic Understanding and Forecasting: A Time Series Approach [J]. IEEE Transactions on Services Computing, 2016, 9(5): 796–805. DOI:10.1109/tsc.2016.2599878

- [6] ZARE MOAYEDI H, MASNADI-SHIRAZI M. A. Arima Model for Network Traffic Prediction and Anomaly Detection[C]//2008 International Symposium on Information Technology. Malaysia, 2008: 1–6. DOI: 10.1109/ITSIM.2008.4631947
- [7] WANG X, ZHOU Z M, YANG Z, et al. Spatio-Temporal Analysis and Prediction of Cellular Traffic in Metropolis[C]//2017 IEEE 25th International Conference on Network Protocols (ICNP). Canada: ICNP, 2017: 1–10. DOI:10.1109/ICNP.2017.8117559
- [8] WANG J, TANG J, XU Z, et al. Spatiotemporal Modeling and Prediction in Cellular Networks: A Big Data Enabled Deep Learning Approach[C]//IEEE INFOCOM. USA: IEEE, 2017:1–9

#### 作者简介



**李建东**, 西安电子科技大学教授、博士生导师, 教育部长江学者特聘教授, 国家杰出青年科技基金获得者; 主要研究方向为智能宽带无线通信、认知无线网络、大规模自组织网络以及无线网络的干扰管理等; 先后主持和参加基金项目 20 余项, 获得国家技术发明奖二等奖 2 项; 已发表论文 200 余篇, 其中被 SCI/EI 检索 100 余篇。



**盛敏**, 西安电子科技大学教授、博士生导师, 教育部长江学者特聘教授, 国家杰出青年科技基金获得者; 主要研究方向为智能宽带无线通信、认知无线网络、大规模自组织网络等; 先后主持和参加基金项目 20 余项, 获得国家技术发明奖二等奖 2 项; 已发表论文 100 余篇, 其中被 SCI/EI 检索 80 余篇。



**文娟**, 西安电子科技大学讲师、硕士生导师; 主要研究方向为智能无线网络、信息智能处理与传输技术以及异构无线网络容量研究等; 先后主持和参加国家基金项目 10 余项; 获陕西省科学技术一等奖 1 项; 已发表 SCI/EI 检索论文 10 余篇。



# 5G 在垂直行业中的应用

## Applications of 5G in Vertical Industry

陆平/LU Ping<sup>1,2</sup>李建华/LI Jianhua<sup>1,2</sup>赵维铎/ZHAO Weiduo<sup>1,2</sup>

(1. 中兴通讯股份有限公司, 广东 深圳 518057;

2. 移动网络和移动多媒体技术国家重点实验室, 广东 深圳 518057)

(1. ZTE Corporation, Shenzhen 518057, China;  
2. State Key Laboratory of Mobile Network and  
Mobile Multimedia Technology, Shenzhen  
518057, China)DOI: 10.12142/ZTETJ.201901011  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190130.1603.004.html>收稿日期: 2018-12-28  
网络出版日期: 2019-01-30

**摘要:** 提出 5G 在各行业领域成熟度及应用进程时间表, 包括增强现实(AR)/虚拟现实(VR)/混合现实(MR)/全息现实(HR)、车联网、智能电网及智能制造等。对 5G 在典型垂直行业应用特点、主要应用场景进行分析, 提出相应场景下对 5G 的关键技术指标要求, 并对今后 5G 在垂直行业领域应用的深入开展和落实提出建议。

**关键词:** 5G; AR/VR/MR/HR; 车联网; 智能电网; 智能制造

**Abstract:** In this paper, the maturity and application schedule of 5G in various industries are proposed, including augmented reality(AR)/ virtual reality(VR)/ mixed reality (MR)/ holographic reality (HR), vehicle networking, smart grid, intelligent manufacturing. The application characteristics and main application scenarios of 5G in typical vertical industries are analyzed, and the key technical requirements of 5G in corresponding scenarios are put forward. Finally, some suggestions on the further development and implementation of 5G in vertical industry are proposed.

**Key words:** 5G; AR/VR/MR/HR; vehicle networking; smart grid; intelligent manufacturing

### 1 5G 技术特征及典型应用场景

随着第 5 代移动通信(5G)第 1 阶段标准的冻结及规范发布, 我们已处于 5G 商用的前夜。大规模天线阵列、超密集组网、新型多址、全频谱接入、软件定义网络(SDN)/网络功能虚拟化(NFV)、边缘计算、网络切片、人工智能等全新技术的应用将为我们呈现出一系列令人振奋的网络能力: 1 Gbit/s 的平均体验速率、10 Gbit/s 的峰值速率、每平方千米超过 100 万的连接数、1 ms 的超低空口时延等。

伴随着网络基础能力的提升, 5G 必将为垂直行业的数字化转型带来重大机遇<sup>[1-2]</sup>。

一些主要的标准组织也给出了 5G 的典型应用场景。第 3 代合作伙伴计划(3GPP)将 5G 应用分为 2 大类: 移动互联网应用及物联网应用; 国际移动通信系统(IMT)-2020(5G)推进组定义了 4 个 5G 主要应用场景: 连续广覆盖、热点高容量、低功耗大连接、高可靠低时延; 国际电信联盟(ITU)定义了 3 个主要应用场景: 增强型移动宽带(eMBB)场景、大规模机器通信(mMTC)场景、高可靠低时延通

信(uRLLC)场景<sup>[3-5]</sup>。

以 ITU 定义为例, eMBB 主要实现更大带宽的移动宽带应用。物联网类应用包括 mMTC 及 uRLLC 2 大场景, 依托 5G 的大连接及高可靠低时延特性, 工业控制、车联网、无人机等领域将获得发展, 并进一步带来物联网芯片模组、网络及垂直应用领域的发展机遇。

当前, 全球运营商及相关行业纷纷展开 5G 相关应用的研究及孵化, 覆盖了移动监测控制、超高清视频、增强现实技术(AR)/虚拟现实技术(VR)、游戏、无人机、车联网以及智慧的



城市、电网、工厂、教育和医疗等众多场景,进一步推动5G与垂直行业应用的深度结合,助力整个社会的数字化转型。总体来看,5G新业务将会分阶段逐步成熟。如图1中所示,5G初期核心应用仍然是面向各领域的泛视频和图像的传输处理场景,如AR/VR类、超高清视频及图像类等应用。

## 2 5G 助力垂直行业发展

### 2.1 5G 对 XR 的发展提供机遇

XR是VR、AR、混合现实(MR)、全息现实(HR)多种视频呈现和交互方式的总称。

XR在可视化、交互方式2个核心能力方向上引入了突破性的革命,在我们当前能预见的场景下,在特征及体验方面,XR将

给通信业务带来质的飞越,并使得未来不可预见的场景充满期待和可能性。此外,XR在各个垂直领域也有着非常广阔的行业应用场景和增值空间,如在自动驾驶、移动监测控制、智能制造、家庭娱乐、智慧零售、智慧教育等。

基于XR的通信业务在众多的垂直行业应用中有着广阔的前景。

#### (1) 虚拟原型。

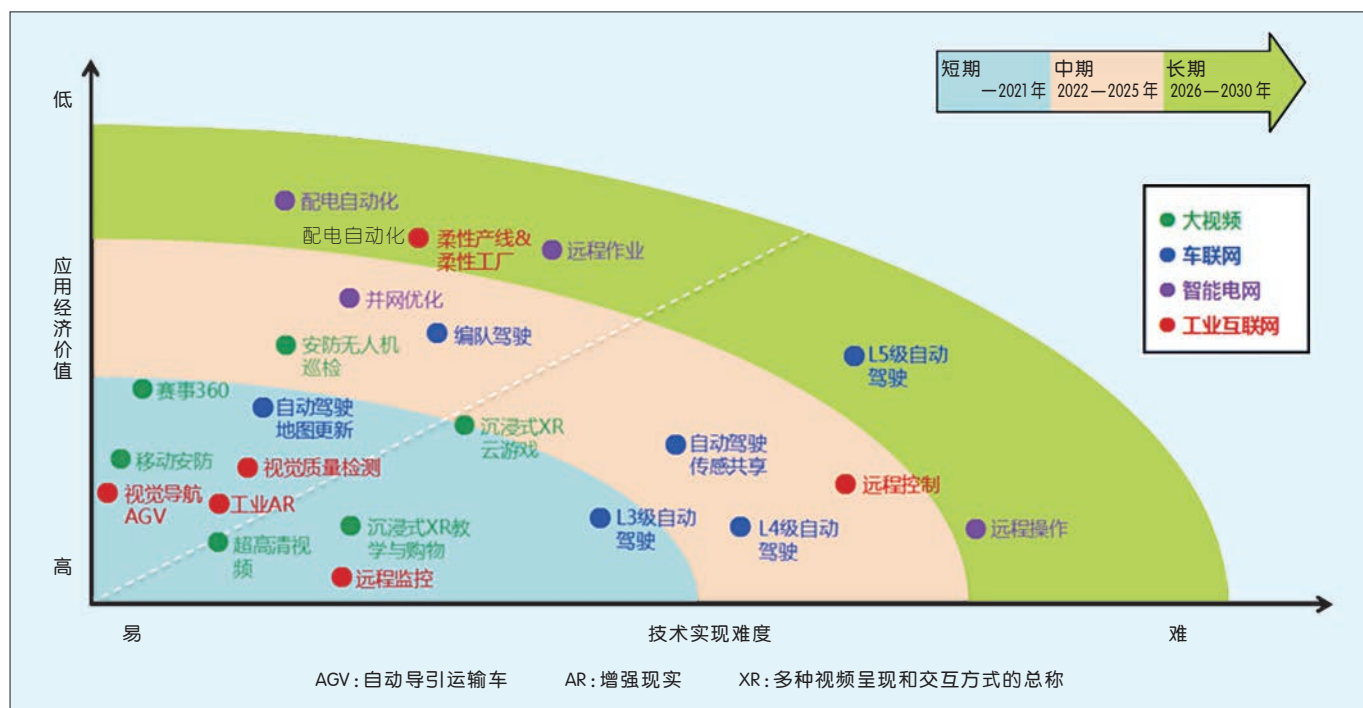
虚拟原型设计使设计师能够与包含原型的虚拟世界快速地进行交互并修改,也能够带客户进入构造的虚拟世界以便交流想法。XR的虚拟原型设计在商业中有重要的用途,包括房地产、建筑业,以及飞机、航天器、汽车、家具、服装和医疗器械的设计。

#### (2) 教育。

对于XR教育来说最大的前景在于人文学科,例如历史学、人类学和外语。用户既可以通过虚拟现实界面观看实体博物馆,又可以直接在虚拟博物馆中扫描和展示文物。

#### (3) 健康医疗。

分布式医疗的发展趋势越来越明显。医生能通过远程呈现提供指导建议,并且运用XR技术来培训。在XR的另一用途中,医生可以让自己沉浸在利用医疗扫描数据生成的3D器官模型中。另外的一类应用是通过XR直接提供理疗以帮助病人,包括通过重复地暴露克服恐惧症和压力失调,改善或维持认知技能以对抗老化,并且改善运动技能以解决身体平衡或者神经系统紊乱失调等问题。



▲ 图1 5G在各行业领域成熟应用进程



#### (4) 增强和混合现实。

在增强和混合现实的应用中,用户可以通过叠加一些附加图形来观看真实世界,以增强其外观。通过在现实世界中放置文本、图标以及其他图形,用户可以利用互联网的力量来辅助许多操作,例如导航、社交互动和机械维护。在工厂的环境中,工人可以在需要组装的零件上看到标识,或可以直接看到机器的内部来确定潜在的、需要替换的零件。

虽然XR有着非常广阔的应用价值和适用场景,但XR对整个通信过程的低延时、高带宽、大并发提出了极严格的要求,具体如图2所示。

XR不同体验下的业务及网络指标如表1所示。

5G下XR视频应用主要包括面向个人(2C)和面向企业(2B)2个场景:2C是现有大视频家庭业务在5G下的延伸;2B主要指5G下的垂直行业应用,例如旅游、教育、医疗、无人机、工业等。

▼表1 XR不同体验下的业务及网络指标

阶段	VR	AR	MR
商用时间预测	2018—2021年	2022—2025年	2026—2030年
视频分辨率/K	4~8(全景)	2~4	2~4
3D业务内容分辨率/K	2~4	2~4	2~4
视场角/度	90~110	40	80
色深/bit	8	8	8
编码标准	H265	H265/H266	H265/H266
帧率/(帧/秒)	30(视频) 50~90(强交互)	50~90	50~90
码率要求/(Mbit/s)	≥40	≥20	≥20
带宽要求/(Mbit/s)	≥60	≥40	≥40
RTT要求/ms	≤20	≤15	≤15
丢包要求	≤9E-5(视频) ≤1.00E-5(强交互)	≤1.00E-5	≤1.00E-5

AR:增强现实 VR:虚拟现实 MR:混合现实 RTT:往返时延

(1) 5G下XR视频业务的2C场景。

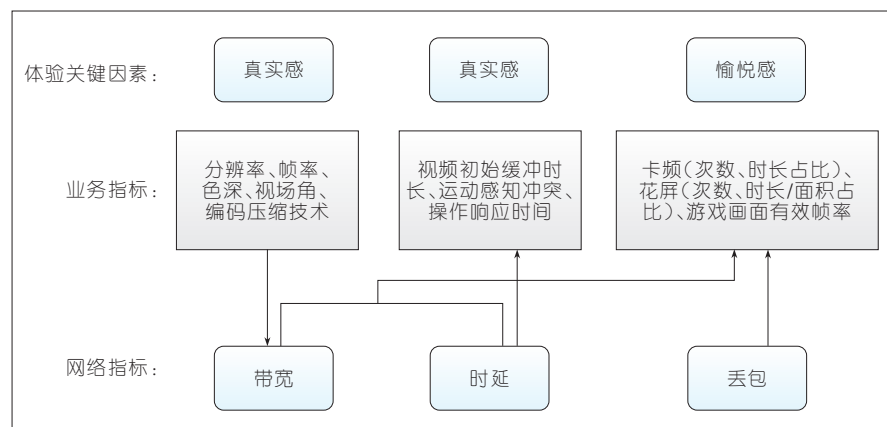
2C也就是通常所说的面向C端用户的超高清分辨率(UHD)和AR/VR视频。目前在这些业务的推广上仍存在一些瓶颈需要通过视频平台协助解决,主要表现为:内容缺乏;终端成本过高,终端体验度不佳(比如目前的VR头显)等。

• 针对内容缺乏的问题。5G下的内容将改变从中心产生分发至边缘的模式,更多的内容

将直接产生于边缘,为此基于移动边缘计算(MEC)实现内容的本地生产将尤为重要。同时,在云端基于人工智能(AI)算法,将现有高清内容转为超高清,也将有助于产生更多UHD内容。

• 针对终端成本过高和体验不佳的问题。需要终端将渲染、计算功能卸载至边缘MEC,由此降低终端的重量、成本,提升接纳度。只有终端做到易于推广,VR业务才能真正推广开。

• 针对UHD/VR内容的体验不佳的问题。考虑通过视频业务专有切片网络,保障优质内容或用户群的观影体验/服务质量(QoS)。这个过程需要云端、边缘甚至终端针对不同的业务和用户类型以及优先级做切片,形成专属保障。基于我们目前的研究,5G网络携带的QoS流识别(QFI)参数可以用作此类切片需求(优先级)识别。除此之外,低时延、视角(FoV)、越过运



▲图2 多种视频呈现和交互方式的应用场景客户体验与移动通信关键指标

营商提供开放互联网服务(OTT)组播等技术的应用有助于QoS的提升。

除了入户的家庭媒体业务外,5G下的场馆场景(如针对场馆内大型赛事、演唱会)也存在以上类似需求;但场馆场景下在MEC平台上会加载更多的视频业务功能模块,如场内多视角的切换、UGC内容的实时剪辑&上传、元数据的加载&呈现、定向广告的实现等。

(2)5G下XR视频业务的2B场景。

2B场景下需要聚焦的问题与2C略有不同。B端需要我们的视频平台有对外开放的能力,供垂直行业灵活调用,所以需要在MEC及视频的核心平台上形成面向垂直行业的业务能力切片。此外,垂直应用当前存在过于离散的问题,需要在云端对离散应用进行汇聚。故在2B场景下需要:

- 对外开放视频核心能力,在MEC上形成能力切片,供垂直应用按需调用。

- 云端对垂直应用形成有效汇聚,将同质化需求整合,并能够由单一应用复制到同类应用,再扩展至同行业应用,最终形成标准。

## 2.2 5G满足车联网统一链接

### 需求

车联网是以车内网、车际网和车载移动互联网为基础,按照约定的通信协议和数据交互标

准,在车辆到任何事物(V2X)(X是车、路、行人及互联网等)之间,进行无线通信和信息交换的大系统网络。近年来,因汽车数量持续增长而引起的交通安全、出行效率、环境保护等问题日益突出,车联网相关应用需求受到了广泛关注<sup>[6]</sup>。

5G移动通信技术在低时延、高带宽、高移动性车联网场景中,以其更加灵活的体系结构解决了车联网多样化的应用场景中差异化性能指标带来的问题,使车辆和车载终端在高速移动下获得更好的性能。另外,随着MEC、网络切片等5G新技术的不断完善,移动通信技术将与汽车行业深度融合,将会改变汽车使用、保有和交通运输本身的传统模式。

借助于人、车、路、云之间的全方位连接和高效信息交互,车联网正从信息服务类应用向交通安全和效率类应用发展,并逐步向支持实现自动驾驶的协同服务类应用演进。未来,5G很有可能作为统一的连接技术,满足远程操作、自动和协作驾驶等连接需求,并在车辆自动驾驶及安全、道路交通管理等方面得到广泛应用。以下我们将介绍几个基于5G技术的典型车联网应用场景。

### (1)信息服务类场景。

提供信息服务能够有效提高车主驾车体验。典型的应用场景包括高精度地图下载/更新、在线导航服务、交通信息获

取等。未来,满足高带宽、移动性需求的5G网络可以为车联网用户提供实时、准确的路况、天气等服务。车辆通过网络与MEC云端服务器连接,请求信息服务,应用服务器通过多渠道收集的数据进行综合分析,将最优信息发送给车辆,随时在有网络覆盖的地方为用户提供服务。

### (2)交通安全类场景。

提升交通安全对于避免交通事故,降低事故带来的生命财产损失有十分重要的意义。未来自动驾驶车辆的感知能力建立在多种信息来源(例如雷达、激光传感器、车载摄像机、高精度地图等)的交换共享和数据融合的基础之上,5G网络可以满足其接近100%的高可靠性、毫秒级端到端时延和每秒上百兆传输速率的需求。通过车载终端将本地传感器和远程服务信息进行智能融合,合作式感知能够有效提升车辆的安全机动性能,在车辆超车、换道、躲避障碍物等方面提供可靠保障。

### (3)交通效率类场景。

优化交通效率对于缓解城市交通拥堵、节能减排具有十分重要的意义。未来可以通过无线通信技术在区域内车辆和基础设施中共享车辆的驾驶意图,在5G网络MEC节点汇合车辆、路况和天气等信息。一方面利用5G网络低时延、高可靠的特点为车辆提供各时间段的路况以及各个路段的实时状况;另一方面可以在MEC节点优化车辆

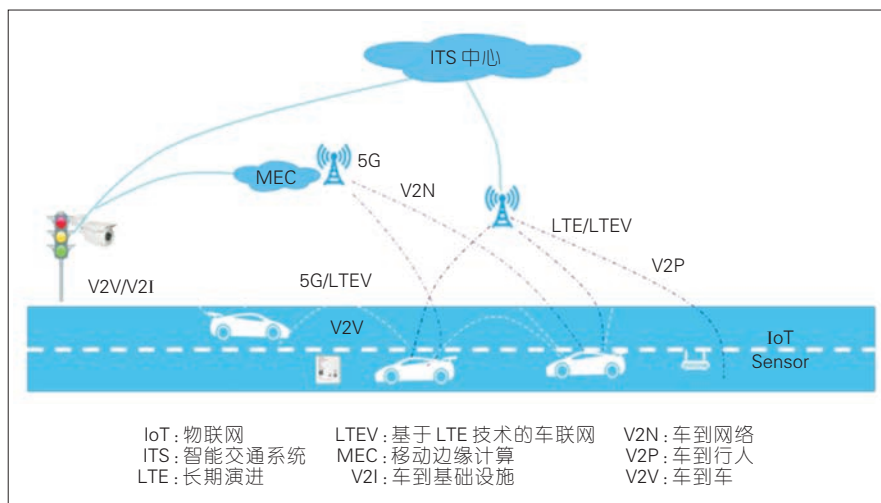
行驶计划,协调车辆的运行策略,从而改善道路通行状况,加快道路通行效率,提升交通安全,并在一定程度上减少排放和降低燃料的消耗。

#### (4)高度协同类场景。

V2X 是一种不容易受到天气、障碍物以及距离等因素影响的人、车、路等信息交互的手段。典型的应用场景包括车辆编队行驶、远程遥控驾驶等。远程遥控驾驶是指驾驶员通过驾驶操控台远程操作车辆行驶。搭载在车辆上的摄像头、雷达等,利用 5G 网络大带宽能力将多路感知信息实时传达到远程驾驶操控台;驾驶员对于车辆方向盘、油门和刹车的操控信号,通过 5G 网络的高可靠低时延特性,实时传达到车辆上,轻松准确地对车辆进行前进、加速、刹车、转弯、后退等驾驶操作。

未来,5G 将以其多异构网络融合以及灵活的终端通信模式,使 5G 车联网在安全通信方面的保障不同于当前车联网。同时,5G 车联网应实行多方面的安全认证,包括车内用户移动终端与 5G 车载单元的安全认证,车与车之间、车与行人之间、车与 5G 移动终端之间以及车与 5G 基站之间的安全认证,并加快推进可商用技术、产品的研发,以此保障车联网的通信安全,如图 3 所示。

车联网正在改变人类交通和通信方式,促使车辆向网络化、智能化发展。可以预见,在



▲ 图 3 未来异构车联网示意图

5G 网络全面商用后,5G 车联网的研究可以促进社会的巨大演进,使人类社会更加方便、安全、快捷、高效。

### 2.3 5G 使能智能电网

在智能电网中,无线通信总体包含控制和采集 2 大场景类(如图 4 所示),对网络有着不同的要求。

智能电网在以下应用场景需要 5G 技术满足其需求:

(1)高精度大范围无线网络授时。

智能配电网分布式终端广泛分布安装在配电网的各种节点上,新能源并网、广域测控保护等业务为授时带来挑战。授时精度要求在  $10\ \mu\text{s}$  内,时延小于  $10\ \text{ms}$ ,可靠性 99.999%。传统的电网授时采用卫星定位等技术,需要单独的天馈系统,受限于部署位置、天气等环境的影响;而 5G 网络授时模块和通信模块合一,直接复用通信天馈,

部署位置灵活,精度可达  $1\ \mu\text{s}$  级别,是电网授时最有潜力的技术之一。

(2)配网同步相量测量(PMU)。

配网同步相量由于测量点多,通信频次高,需要做到实时控制和测量,时延要求  $10\ \text{ms}$  内,可靠性 99.999%。在通信上传统的方式以有线为主、无线及电力线载波通信为辅,其成本高、运维难、扩展性不强。5G 网完全满足配电网 PMU 的数据传输速率、时延和可靠性需求,同时 5G 网络具备的高密度连接能力可满足 PMU 的大量部署需求,使得配电网 PMU 应用具有极大的可扩展空间。同时采用网络切片技术,可为 PMU 提供专用通信通道,与其他配电自动化业务隔离,保证 PMU 业务的安全性和可靠性。同时可在靠近配电网系统本地的 MEC 平台中部署 PMU 部分主站功能,实现高实时性业务快速响应。



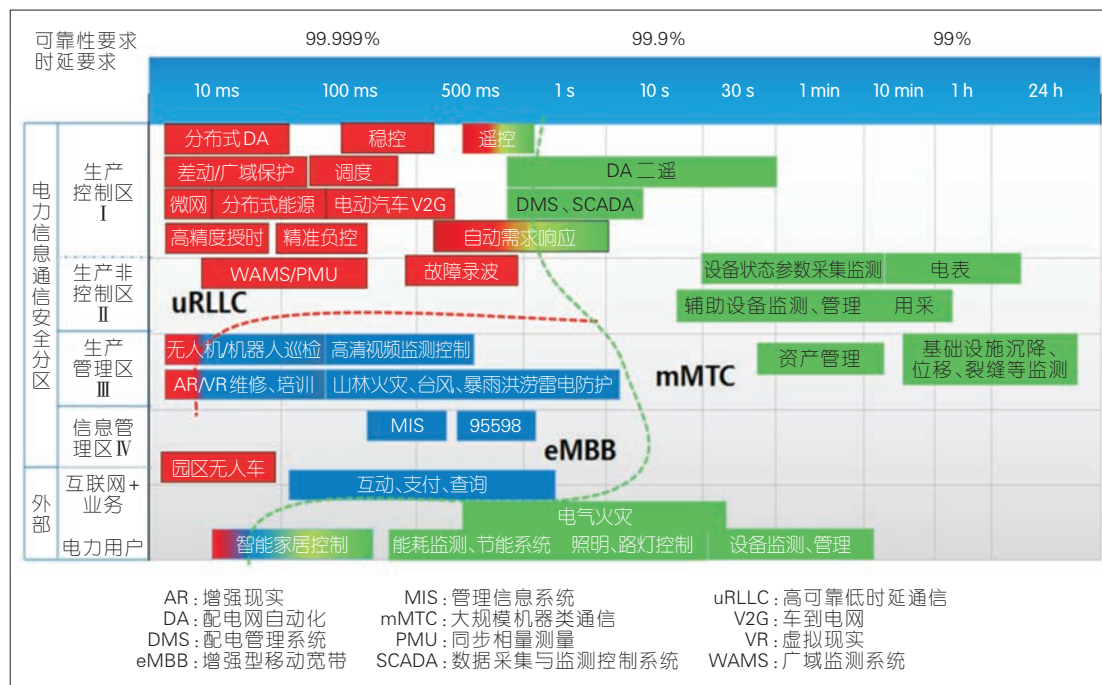


图4  
智能电网应用及对5G  
网络的需求

### (3) 智能配电自动化。

智能配电自动化通过继电保护自动装置监测配电网线路和设备状态信息，快速实现配电网或设备故障的判断及精准定位，迅速隔离故障区段或故障设备。而传统的配网保护采用简单的过流、过压保护，不依赖通信，不能实现分段隔离，停电影响范围大。为了实现故障的精准隔离，可以利用5G切片技术在配网领域推广应用差动保护，要求通信时延小于10 ms，带宽大于2 Mbit/s，可靠性99.999%，信息交互携带10 μs内的高精度时间戳。

### (4) 分布式能源调控。

为适应太阳能、风能、燃料电池、充电桩等分散电力需求和资源分布特点，需要进行数据采集处理、有功功率调节、电压无功功率控制、孤岛监测、调度与

协调控制等。为了实现这些功能，需要达到百万到千万级别的连接数。通过5G面向大连接的能力，可以实现对接入设备进行控制和数据采集，从而有效保障多电源并网的有效运行。

### (5) 用电负荷需求侧响应。

当电力批发市场价格升高或系统可靠性受威胁时，通过引导用户改变用电模式，减少某时间段用电负荷，来保障电网稳定。这需要负荷管理控制时延小于50 ms，带宽小于2 Mbit/s，可靠性99.999%。而当前需求侧响应对负荷的控制指令在终端与主站间交互，终端横向无数据交互，对负荷的控制通常只能切除整条配电网。利用5G低时延高可靠特点，通过用户、售电商、运营商等多方参与，通过灵活多样的市场化需求侧响应模式，优先切除可中断非重要负

荷，实现精准负荷控制。

### (6) 智能巡检和应急指挥。

利用机器人、无人机等搭载高清摄像头或环境传感器，实现对变电站、配电房、输电线路进行巡检的过程中，需要将现场高清视频、图片等数据传回远程监测中心。此外，针对重大电力故障进行现场应急抢险时，也需要通过高清视频等方式将现场情况实时反馈到指挥中心。为满足这些通信要求，需要单路带宽达到4~10 Mbit/s，时延小于100 ms。针对这类业务场景，5G网络提供的高带宽，有效地保障了智能巡检和应急指挥的进行。

### (7) 设施运行状态监测及远程操作。

电网运行中，需要通过对电网设备运行环境、运行状态、故障信息进行全面监测来掌握设施运行情况，这涉及到千万级的



终端设备接入。此外,通过智能电表实现远程抄表,也存在大量的设备接入。基于5G面向大连接的能力,为末端设备的全面接入提供支撑,进一步保障设施安全稳定运行。

## 2.4 5G 助力制造走向智造

当前,工业领域无线技术主要应用于设备及产品信息的采集、非实时控制和实现工厂内部信息化等。由于在可靠性、数据传输速率、覆盖距离、移动性等方面的不足,导致当前工业领域无线技术未能广泛应用,无线通信占比仅约6%左右。随着5G技术的不断发展成熟,特别是其特有的低时延、高可靠,以及大带宽等特性,使得无线技术应用于现场设备实时控制、远程维护及操控、工业高清图像处理等工

业应用新领域成为可能,同时也为未来柔性产线、柔性车间奠定了基础。5G开启工业无线未来发展。

如图5所示,工厂车间中将出现更多的无线连接,这将促使工厂车间网络架构不断优化,有效提升网络化协同制造与管理水平,促进工厂车间提质增效,保持对整个产品生命周期的全连接。以上这些主要表现在以下4个场景:

(1)实时控制。5G特有的低时延、高可靠特性,使得无线技术应用于工业实时控制领域成为可能,5G将成为未来工业有线控制网络的补充或替代。该技术参数要求为:时延为1~20 ms,带宽为千比特每秒;可靠性可达99.9999%。

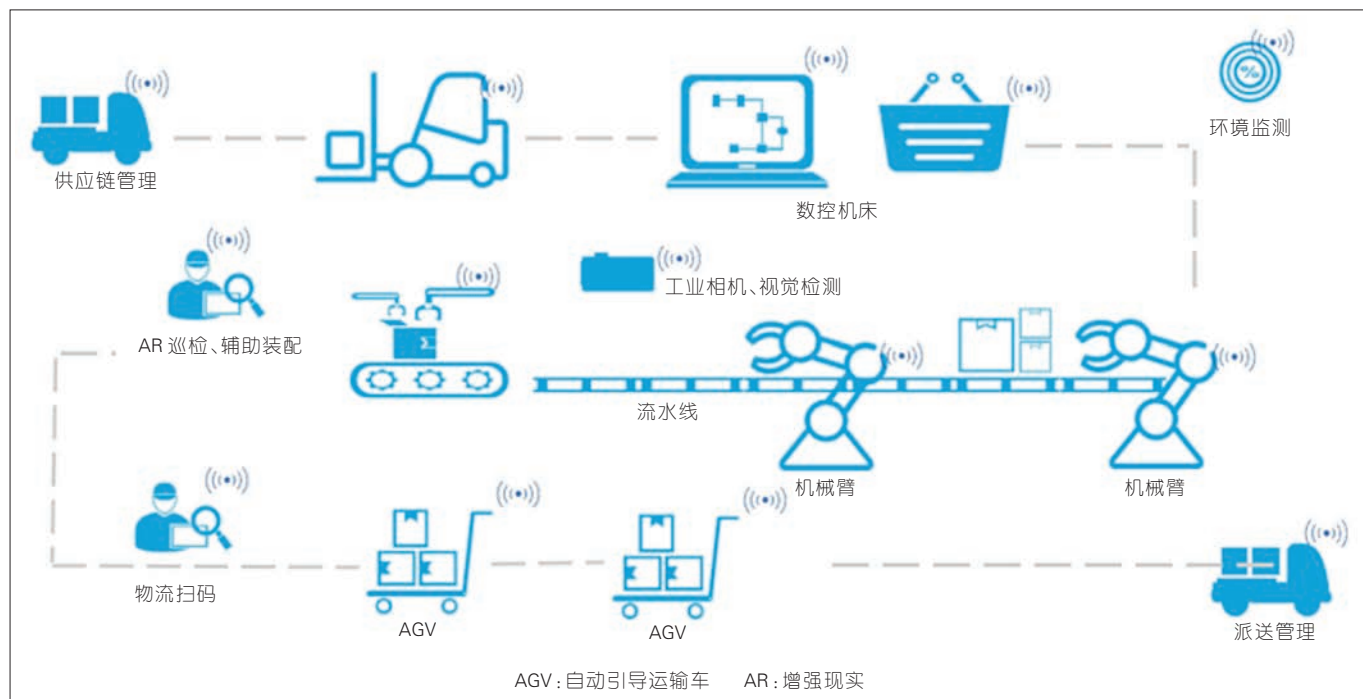
(2)工业穿戴。发挥5G大

带宽及移动性优势,实现工业可穿戴设备以及工业增强现实等较为稳定的实时无线传输。该技术参数要求为:时延<100 ms,带宽为20 Mbit/s~1 Gbit/s。

(3)调度与导航。这可以包括基于5G对移动机器人或移动设备的调度控制、基于视觉的路径导航规划等等。该技术参数要求为:时延为50 ms~1 s,带宽为1 kbit/s~100 Mbit/s。

(4)高密度接入。包括海量工业传感器/变送器、仪器仪表数据采集以及各类人员、装备、产品与物料的连接。该技术参数要求为:连接可达几百/平方米~几千/平方米,带宽为千比特每秒到兆比特每秒。

伴随中国加快实施制造强国战略,推进智能制造发展,5G将广泛深入应用于工业领域。



▲图5 基于第5代移动通信的未来无线工厂

工厂内部和外部之间的界限越来越模糊,工厂不再是独立的封闭实体,而是庞大的价值链和生态系统的一部分。同时 5G 也为制造业开辟新的商业模式提供可能。

### 3 结束语

5G 技术以及人工智能、边缘计算、计算机视觉、传感等基础技术有机高效地进行整合,必将为各垂直行业数字化转型提供切切实实的手段。当前,为促进 5G 产业生态的不断发展壮大,仍需充分发挥政府引导作用,如依托 IMT-2020(5G)推进组、工业互联网产业联盟等行业平台,统筹科研机构、高校、电信运营商、设备制造商、终端厂商、互联网企业和行业企业等产学研用力量,协同开展 5G 技术研究、标准研制、设备开发与行业应用,加快推动 5G 与工业互

联网、车联网、智能电网、远程医疗等垂直行业应用融合发展。

中兴通讯是 5G 业务赋能者、创新应用实践者、生态建设积极参与者。中兴通讯一直积极参与并打造业内产业联盟,与各行业龙头企业合作,共同探索 5G 垂直行业应用及相关标准,打造 5G 行业试点示范应用,为 5G 的产业发展繁荣不断贡献自己的力量。

### 致谢

本研究得到中兴通讯股份有限公司叶郁文、郭雪峰、陈永波、张慧、邓芳伟等行业专家的帮助,谨致谢意!

### 参考文献

- [1] 中兴通讯.中兴通讯 5G 技术白皮书[R].
- [2] IMT-2020(5G).IMT-2020(5G)推进组-5G 愿景与需求白皮书[EB/OL].(2015-02-11)[2018-12-28].<http://www.imt-2020.org.cn/zh/documents/1?currentPage=2&content=>
- [3] 5G 经济社会影响白皮书[EB/OL].(2018-12-28).<http://www.imt-2020.org.cn/zh/documents/1>
- [4] 3GPP.Service Requirements for the 5G

System: 3GPP TS 26.261[R].2017  
[5] C-V2X 白皮书[EB/OL]. [2018-12-28].<http://www.imt-2020.org.cn/zh/documents/1>

### 作者简介



陆平,中兴通讯股份有限公司副总裁、技术规划部部长,中国计算机学会大数据专家委员会委员;主要从事云计算、大数据、人工智能等方面的研究工作;支持和参与了国家科技重大专项、国家科技支撑项目等;发表多篇论文,撰写了《物联网能力开发与应用》《云计算中的大数据技术与应用》等多部著作。



李建华,中兴通讯股份有限公司资深战略规划师、技术委员会专家,大数据产业生态联盟专家;主要研究领域为大数据、人工智能、5G 业务、创新商业模式等;发表论文多篇,获国家发明专利 1 项、实用新型专利 1 项。



赵维铎,中兴通讯股份有限公司行业规划总工程师、技术委员会专家;主要研究领域为 5G 网络、工业互联网、智能电网等;发表论文 10 余篇,共获得专利 7 项。

# 《中兴通讯技术》杂志(双月刊)投稿须知

## 一、杂志定位

《中兴通讯技术》杂志为通信技术类学术期刊。通过介绍、探讨通信热点技术,以展现通信技术最新发展动态,并促进产学研合作,发掘和培养优秀人才,为振兴民族通信产业做贡献。

## 二、稿件基本要求

### 1. 投稿约定

- (1)作者需登录《中兴通讯技术》投稿平台:tech.zte.com.cn/submission,并上传稿件。第一次投稿需完成新用户注册。
- (2)编辑部将按照审稿流程聘请专家审稿,并根据审稿意见,公平、公正地录用稿件。审稿过程需要1个月左右。

### 2. 内容和格式要求

- (1)稿件须具有创新性、学术性、规范性和可读性。
- (2)稿件需采用WORD文档格式。
- (3)稿件篇幅一般不超过6000字(包括文、图),内容包括:中、英文题名,作者姓名及汉语拼音,作者中、英文单位,中文摘要、关键词(3~8个),英文摘要、关键词,正文,参考文献,作者简介。
- (4)中文题名一般不超过20个汉字,中、英文题名含义应一致。
- (5)摘要尽量写成报道性摘要,包括研究的目的、方法、结果/结论,以150~200字为宜。摘要应具有独立性和自明性。中英文摘要应一致。
- (6)文稿中的量和单位应符合国家标准。外文字母的正斜体、大小写等须写清楚,上下角的字母、数据和符号的位置皆应明显区别。
- (7)图、表力求少而精(以8幅为上限),应随文出现,切忌与文字重复。图、表应保持自明性,图中缩略词和英文均要在图中加中文解释。表应采用三线表,表中缩略词和英文均要在表内加中文解释。
- (8)所有文献必须在正文中引用,文献序号按其在文中出现的先后次序编排。常用参考文献的书写格式为:
  - 期刊[序号]作者.题名[J].刊名,出版年,卷号(期号):引文页码.数字对象唯一标识符
  - 书籍[序号]作者.书名[M].出版地:出版者,出版年:引文页码.数字对象唯一标识符
  - 论文集中析出文献[序号]作者.题名[C]/论文集编者.论文集名(会议名).出版地:出版者,出版年(开会年):引文页码.数字对象唯一标识符
  - 学位论文[序号]作者.题名[D].学位授予单位所在城市名:学位授予单位,授予年份.数字对象唯一标识符
  - 专利[序号]专利所有者.专利题名:专利号[P].出版日期.数字对象唯一标识符
  - 国际、国家标准[序号]标准名称:标准编号[S].出版地:出版者,出版年.数字对象唯一标识符
- (9)作者超过3人时,可以感谢形式在文中提及。作者简介包括:姓名、工作单位、职务或职称、学历、毕业于何校、现从事的工作、专业特长、科研成果、已发表的论文数量等。
- (10)提供正面、免冠、彩色标准照片一张,最好采用JPG格式(文件大小超过100kB)。
- (11)应标注出研究课题的资助基金或资助项目名称及编号。
- (12)提供联系方式,如:通讯地址、电话(含手机)、Email等。

### 3. 其他事项

- (1)请勿一稿两投。凡在2个月(自来稿之日算起)以内未接到录用通知者,可致电编辑部询问。
- (2)为了促进信息传播,加强学术交流,在论文发表后,本刊享有文章的转摘权(包括英文版、电子版、网络版)。作者获得的稿费包括转摘酬金。如作者不同意转摘,请在投稿时说明。

编辑部地址:安徽省合肥市金寨路329号凯旋大厦1201室,邮政编码:230061

联系电话:0551-65533356,联系邮箱:magazine@zte.com.cn

本刊只接受在线投稿,欢迎访问本刊投稿平台:tech.zte.com.cn/submission

办刊宗旨:

以人为本,荟萃通信技术领域精英  
迎接挑战,把握世界通信技术动态  
立即行动,求解通信发展疑难课题  
励精图治,促进民族信息产业崛起

---

双月刊 1995年创刊 总第144期  
2019年2月 第25卷 第1期

主管:安徽出版集团有限责任公司  
主办:时代出版传媒股份有限公司  
深圳航天广宇工业有限公司  
出版:安徽科学技术出版社  
编辑、发行:中兴通讯技术杂志社

总编:王喜瑜  
副总编:蒋贤骏  
常务副总编:黄新明  
责任编辑:徐烨  
编辑:卢丹、朱莉  
排版制作:余刚  
发行:王萍萍  
编务:王坤

---

《中兴通讯技术》编辑部  
地址:合肥市金寨路329号凯旋大厦1201室  
邮编:230061  
网址:tech.zte.com.cn  
投稿平台:tech.zte.com.cn/submission  
电子信箱:magazine@zte.com.cn  
电话:(0551)65533356

传真:(0551)65850139  
发行范围:公开发行  
印刷:合肥添彩包装有限公司  
出版日期:2019年2月10日  
中国标准连续出版物号:ISSN 1009-6868  
CN 34-1228/TN  
定价:每册 20.00 元