



中文核心期刊 信息通信领域产学研合作特色期刊  
中国科技核心期刊 第三届国家期刊奖百种重点期刊

ISSN 1009-6868  
CN 34-1228/TN

# 中兴通讯技术

## ZTE TECHNOLOGY JOURNAL

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

2023 年 12 月 · 第 6 期

专题：面向双碳的新一代无线通信网络

ISSN 1009-6868



9 771009 686236

# 《中兴通讯技术》第9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) 麋正琨(南京邮电大学教授)

**主任** 陆建华(中国科学院院士)

**副主任** 李自学(中兴通讯股份有限公司董事长) 李建东(西安电子科技大学教授)

## 编 委 (按姓名拼音排序)

陈建平 上海交通大学教授

陈前斌 重庆邮电大学教授、副校长

段晓东 中国移动研究院副院长

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 伟 东南大学教授

黄宇红 中国移动研究院院长

纪越峰 北京邮电大学教授

江 涛 华中科技大学教授

蒋林涛 中国信息通信研究院科技委主任

金 石 东南大学首席教授、副校长

李尔平 浙江大学教授

李红滨 北京大学教授

李厚强 中国科学技术大学教授

李建东 西安电子科技大学教授

李乐民 中国工程院院士

李融林 华南理工大学教授

李自学 中兴通讯股份有限公司董事长

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

隆克平 北京科技大学教授

陆建华 中国科学院院士

马建国 之江实验室教授

毛军发 中国科学院院士

孟洛明 北京邮电大学教授

石光明 鹏城实验室副主任

孙知信 南京邮电大学教授

谈振辉 北京交通大学教授

唐 宏 中国电信IP领域首席专家

唐雄燕 中国联通研究院副院长

陶小峰 北京邮电大学教授

王文博 北京邮电大学教授、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

王耀南 中国工程院院士

王志勤 中国信息通信研究院副院长

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

向际鹰 中兴通讯股份有限公司首席科学家

肖 甫 南京邮电大学教授、副校长

解冲锋 中国电信研究院教授级高工

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

薛向阳 复旦大学教授

薛一波 清华大学教授

杨义先 北京邮电大学教授

叶 茂 电子科技大学教授

易芝玲 中国移动研究院首席科学家

张宏科 中国工程院院士

张 平 中国工程院院士

张钦宇 哈尔滨工业大学(深圳)教授、副校长

张 卫 复旦大学教授

张云勇 中国联通云南分公司总经理

赵慧玲 工业和信息化部信息通信科技委常委

郑纬民 中国工程院院士

钟章队 北京交通大学教授

周 亮 南京邮电大学教授、副校长

朱近康 中国科学技术大学教授

祝宁华 中国科学院院士

# 目 次

中兴通讯技术 (ZHONGXING TONGXUN JISHU)  
总第 173 期 第 29 卷 第 6 期(卷终) 2023 年 12 月

中文核心期刊 信息通信领域产学研合作特色期刊 中国科技核心期刊 第三届国家期刊奖百种重点期刊 工信部优秀科技期刊 中国五大文献数据库收录期刊 1995 年创刊

## 热点专题 ► 面向双碳的新一代无线通信网络

- 01 专题导读 ..... 葛晓虎, 李建东  
03 基于多点信道海图的低能耗网络覆盖结构优化 ..... 叶远青, 张四海, 朱近康  
11 面向 5G-A 的无线网络节能关键技术 ..... 郭诚, 陈梦竹  
16 面向节能减排的跨制式融合感知通信系统 ..... 周均翼, 周琳, 张舜卿  
23 基于多时隙业务联合整形的低能耗资源调度方法 ..... 李建东, 牛淳隆, 赵晨曦, 刘俊宇  
29 低开销智能反射面辅助无线通信研究综述 ..... 陈天贝, 李娜, 陶小峰  
39 基于视距概率模型的 6G 室内无线通信系统性能分析 ..... 毛曦晨, 王承祥, 杨松江, 黄杰, 黄晨  
46 基于服务覆盖的无蜂窝边缘智能网络部署方法 ..... 姜静, 乔东伟, Worakrin SUTTHIPHAN

## 专家论坛 ► 54 6G 绿色网络发展趋势和关键技术 ..... 李福昌, 郭希蕊 60 零信任关键技术与产业发展研究 ..... 张云畅, 柴瑶琳, 穆域博

## 企业视界 ► 66 集成电路产业技术发展趋势探讨 ..... 刘新阳, 晁沛荫, 李婷宇

## 技术广角 ► 71 软件定义广域网络(SD-WAN)2.0 测试方法 ..... 党小东, 柴瑶琳, 穆域博

## 综合信息 ► 45 《中兴通讯技术》2024 年专题计划 I 《中兴通讯技术》第 29 卷总目次 III 《中兴通讯技术》入选《中文核心期刊要目总览》

## 《中兴通讯技术》2023 年热点专题名称及策划人

### 1. 面向云网安全的新型防护技术

中国电信研究院教授级高工 解冲锋  
北京邮电大学教授 杨义先

### 3. 数字孪生技术

重庆邮电大学教授 陈前斌

### 5. 6G 网络技术

北京邮电大学教授 王文东

### 2. 语义通信

中国科学院院士 陆建华  
清华大学教授 陶晓明

### 4. 算力网络和东数西算

工业和信息化部通信科技委  
专职常委 赵慧玲

### 6. 面向双碳的新一代无线通信网络

华中科技大学教授 葛晓虎  
西安电子科技大学教授 李建东

# MAIN CONTENTS

ZTE TECHNOLOGY JOURNAL  
Vol. 29 No. 6 Dec. 2023

## Special Topic ▶

### New Generation Wireless Communication Network for Dual-Carbon

- 01 Editorial ..... GE Xiaohu, LI Jiandong  
03 Optimization of Low-Energy Overlay Network Structure Based on Multi-Channel Charting ..... YE Yuanqing, ZHANG Sihai, ZHU Jinkang  
11 Key Technologies for 5G-A Wireless Network Energy Saving ..... GUO Cheng, CHEN Mengzhu  
16 Energy Conservation and Emission Reduction Oriented Perception and Communication Fused Wi-Fi and Cellular Integrated Communication System ..... ZHOU Junyi, ZHOU Lin, ZHANG Shunqing  
23 A Multi-Slot Traffic Shaping Based Resource Allocation Method for Low Energy Consumption ..... LI Jiandong, NIU Chunlong, ZHAO Chenxi, LIU Junyu  
29 Survey on Low-Overhead Reconfigurable Intelligent Surface Assisted Wireless Communication ..... CHEN Tianbei, LI Na, TAO Xiaofeng  
39 LoS Probability Model Based Performance Analysis for 6G Indoor Wireless Communication Systems ..... MAO Xichen, WANG Chengxiang, YANG Songjiang, HUANG Jie, HUANG Chen  
46 Service Coverage-Oriented Design on Mobile Edge Computing Aided Cell-Free Massive MIMO Networks ..... JIANG Jing, QIAO Dongwei, Worakrin SUTTHIPHAN

## Expert Forum ▶

- 54 Development Trends and Key Technologies of 6G Green Networks ..... LI Fuchang, GUO Xirui  
60 Key Technologies and Industrial Development of Zero Trust ..... ZHANG Yunchang, CHAI Yaolin, MU Yubo

## Enterprise View ▶

- 66 Development Trends of Integrated Circuit Industry Technology ..... LIU Xinyang, CHAO Peiyin, LI Tingyu

## Research Paper ▶

- 71 Test Methodology of Software Defined Wide Area Network (SD-WAN) 2.0 ..... DANG Xiaodong, CHAI Yaolin, MU Yubo

期刊基本参数：CN 34-1228/TN\*1995\*b\*16\*75\*zh\*p\*¥20.00\*6500\*12\*2023-12

## 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。

# 面向双碳的新一代无线通信网络 专题导读



## 专题策划人



葛晓虎，华中科技大学电子信息与通信学院副院长、国家绿色通信与网络联合研究中心主任、国家高层次人才特殊支持计划科技创新领军人才、中国科协海智计划特聘专家、爱思唯尔高被引学者、华中卓越学者特聘教授、澳大利亚悉尼科技大学兼职教授，还担任国际信息处理联合会（IFIP）理事会中国代表和联合国咨询信息与通信技术专委会委员、国家科技部和国家自然基金委员会评专家、湖北省科技厅评审专家、IET Fellow、中国通信学会会士；近年来主持承担了20余项国家级科研课题和多项国际高水平科技合作项目；2005年获意大利伽利略奖，多篇代表性论文被评为IEEE通信学会领域最佳论文，研究成果荣获世界互联网大会国际组织2023年领先科技奖；已发表论文200余篇。



李建东，西安电子科技大学教授、“长江学者”特聘教授、IEEE Fellow、中国电子学会会士、中国通信学会会士、西安电子科技大学空天地一体化综合业务网全国重点实验室学术委员会副主任、国家新一代宽带无线移动通信网重大专项总体组专家、中国首批“新世纪百千万人才工程”国家级人选、第3届教育部“青年教师奖”获得者、2007年国家杰出青年科学基金获得者、科技部国家“863”计划15周年先进个人等；主要研究领域为宽带无线通信、智能无线网络、大规模自组织网等；主持“863”计划重大课题、国家自然科学基金重大项目等，获得国家技术发明二等奖2项、省部级科技进步奖6项；已发表论文80余篇，出版教材和专著5部。

由于温室效应不断增强，在第75届联合国大会上中国正式提出“双碳”目标。中国在“十四五”规划中明确提出前瞻布局通信领域，旨在引领新一代通信网络技术的发展，实现碳中和的目标。新一代通信技术不仅需要追求高效、快速、可靠，还要以碳减排和环境友好为导向，为数字化社会的可持续发展贡献力量。目前，关于新一代无线通信网络的研究工作主要集中在通信与新能源结合的新能源蜂窝网络、智能反射面辅助网络等方面。本期专题和专家论坛栏目以面向双碳的新一代无线通信网络为主题，邀请该领域的专家学者撰写了多篇文章。这些文章对新一代低碳、低能耗无线通信网络技术的愿景、主要技术挑战、研究内容进行了介绍、分析。

《基于多点信道海图的低能耗网络覆盖结构优化》以

资源小区覆盖为切入点，利用多点信道海图的方法开展优化设计，提出了一种基于多点信道海图的资源小区优化方法，该方法能够有效提高系统网络容量。《面向5G-A的无线网络节能关键技术》一文面向5G-A，介绍了网络节能的关键技术点，从时域、空域、频域和功率域等方面，使能更动态、高效、精准的无线传输，并辅助以终端反馈和基站间交互信息等方式提升节能效果。《面向节能减排的跨制式融合感知通信系统》提出了一套面向节能减排的跨制式融合感知通信系统，包括通信融合、感知融合以及边缘协同管控平台3个部分，该系统有望对未来面向节能减排的跨制式网络设计提供可靠验证与指导。《基于多时隙业务联合整形的低能耗资源调度方法》一文设计了多时隙业务联合整形的低能耗资源调度方法，通过感知用户业务流量和时延要求等需求侧的数据特征，提出了多时隙业务联合整形算法，降低了用户业务请求的非平稳性，并利用深度强化学习算法给出多

个时隙内基站的动态资源分配策略，从而减少了基站的非线性传输特性产生的额外能耗。《低开销智能反射面辅助无线通信研究综述》对低开销智能反射面（RIS）辅助无线通信进行了系统综述，分析了开销问题对RIS协助无线网络性能的负面影响，并对未来亟待解决的RIS开销问题及相关研究方向进行了总结。《基于视距概率模型的6G室内无线通信系统性能分析》针对室内环境，构建了视距信号概率模型，并基于信噪比分布研究了室内覆盖性能和能量效率，相关结论有助于指导室内无线通信系统部署，提升系统效率。为实现未来网络绿色低碳发展，《基于服务覆盖的无蜂窝边缘智能网络部署方法》一文将服务覆盖作为网络覆盖能力的新度量，通过深入挖

掘无蜂窝网络部署灵活的优势，将系统性能整体优化转换为每个覆盖单元的计算与通信服务能力与用户的计算和存储业务需求相匹配为目标。专家论坛栏目的《6G绿色网络发展趋势和关键技术》一文提出了构建泛在智简、绿能降碳的6G绿色网络的愿景，以及6G绿色网络在网络架构、关键技术、协议简化、绿色设备、绿能融合等方面的技术演进方向，打造绿色低碳的6G绿色网络。

本期的作者来自知名高校与科研机构。针对面向双碳的新一代无线通信网络技术，文章从技术挑战、系统架构和关键技术等方面介绍了最新的研究成果。希望本期的内容能为读者提供有益的启示和参考，在此对所有作者的大力支持和审稿专家的辛勤指导表示由衷的感谢！

# 基于多点信道海图的低能耗网络覆盖结构优化



## Optimization of Low-Energy Network Coverage Structure Based on Multipoint Channel Charting

叶远青/YE Yuanqing, 张四海/ZHANG Sihai,  
朱近康/ZHU Jinkang  
(中国科学技术大学, 中国 合肥 230026)  
(University of Science and Technology of China, Hefei 230026, China)

DOI: 10.12142/ZTETJ.202306002  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231212.1548.008.html>  
网络出版日期: 2023-12-13  
收稿日期: 2023-10-15

**摘要:** 资源小区覆盖结构是实现超高数据密度类业务场景的关键覆盖结构之一, 其具有弹性可扩展、高能效按需覆盖的特点。依托这种覆盖结构, 研究了一种基于多点信道海图的资源小区优化方法, 旨在发射功率约束下提高系统和速率。将多点信道海图感知到的相邻用户划分到同一资源小区, 通过在资源小区内分配正交资源块和在资源小区间复用正交资源块的方式, 减少系统干扰并提升系统和速率。仿真结果表明, 该方法能够有效提高系统和速率, 不需要用户位置信息和瞬时信道状态信息, 并且在相同系统和速率的情况下, 资源小区覆盖结构可以有效降低网络能耗。

**关键词:** 多点信道海图; 低能耗; 资源小区

**Abstract:** Resource cell coverage structure is one of the key coverage structures to achieve ultra-high data density scenarios, which has the characteristics of elastic scalability and high energy efficiency on-demand coverage. Relying on the above coverage structure, a multipoint channel charting-based resource cell generation method is proposed to improve sum system rate under transmit power constraint. Specifically, adjacent users on multipoint channel charting are grouped into the same resource cell to reduce system interference and improve sum system rate by allocating orthogonal resource blocks within the resource cell and multiplexing RBs between resource cells, respectively. Simulation results show that the proposed method provides significant gain in capacity without requiring user location information or instantaneous channel state information. It also indicates that the resource cell coverage structure can effectively reduce network energy consumption under the same sum system rate condition.

**Keywords:** multipoint channel charting; low-energy; resource cell

**引用格式:** 叶远青, 张四海, 朱近康. 基于多点信道海图的低能耗网络覆盖结构优化 [J]. 中兴通讯技术, 2023, 29(6): 3-10. DOI: 10.12142/ZTETJ.202306002

**Citation:** YE Y Q, ZHANG S H, ZHU J K. Optimization of low-energy network coverage structure based on multipoint channel charting [J]. ZTE technology journal, 2023, 29(6): 3-10. DOI: 10.12142/ZTETJ.202306002

人工智能的发展和智能设备的普及对6G移动网络提出了核心需求。6G网络将支持高速率、大容量、超密集类型的业务, 即超高数据密度业务(uHDD)<sup>[1]</sup>。随着数据流量的急剧增长和连接设备数量的不断增加, 无线网络的能源消耗也在快速增长。这将导致温室气体大量产生, 对环境保护和可持续发展造成重大威胁<sup>[2]</sup>。因此, 未来移动通信和

网络技术的发展势必要统筹兼顾能耗和资源, 打造绿色通信网络。

为了提高无线网络的能效, 在由分布单元(DU)、集中单元(CU)与核心网构成的5G三级网络结构的基础上, 创新演进更多的无线覆盖结构, 如5G云无线接入网(C-RAN)覆盖结构、雾无线接入网(Fog-RAN)覆盖结构、无蜂窝大规模输入输出(MIMO)覆盖结构等。但这些传统无线覆盖结构部署方式较为固定, 只能在局部范围内进行有限的调整, 难以与业务分布不均匀的uHDD场景匹配, 降低了网络的容量覆盖能力。

基金项目: 科技部重点研发项目(2022YFB2902302); 华为技术有限公司无线研究创新项目(FA2019051101)

资源小区覆盖结构<sup>[3]</sup>可消除容量覆盖漏洞，克服网络的复杂干扰，其生成方法（RCG）根据用户位置对用户进行聚类，同时根据用户位置信息和信道状态信息（CSI）调整覆盖结构以匹配业务分布，并通过贪婪启发式算法<sup>[4]</sup>朝着系统和速率最大化的方向迭代优化资源小区。然而，用户位置信息很难获取，例如：现有的用户定位方法（如三角测量）受限于视距传播条件<sup>[5]</sup>，根据开放系统互连分层原则，基站无法从全球导航卫星系统获取位置信息<sup>[6]</sup>等。此外，上述方法在很大程度上依赖于可用CSI的准确性和时效性。然而，由快速衰减引起的瞬时信道状态变化导致获取的CSI总是过时的。

为了解决用户位置难获取和CSI过时的问题，我们提出了一种基于多点信道海图（MPCC）的资源小区生成方法。传统通信系统的能效最大化通常由信噪比或速率约束来保证。本文中，我们在能量消耗约束下最大化系统和速率，从长期CSI获得的二阶信道统计量中提取有效特征，然后构建MPCC，以便更有效地保持用户之间的信道相关性。MPCC上相邻用户具有很强的空间相关性，如果用户共享同一个资源块（RB），将会产生严重的干扰。因此，采用基于MPCC的RCG算法，将MPCC上相邻的用户划分为同一簇，生成资源小区，在每个资源小区内分配正交RBs，并在资源小区间进行RBs的空间复用，可以减少系统干扰，提高系统和速率。结果表明，该算法可在用户位置不可用和RCG可用CSI不完善或过时的情况下提升系统和速率。

## 1 资源小区与系统模型

### 1.1 资源小区

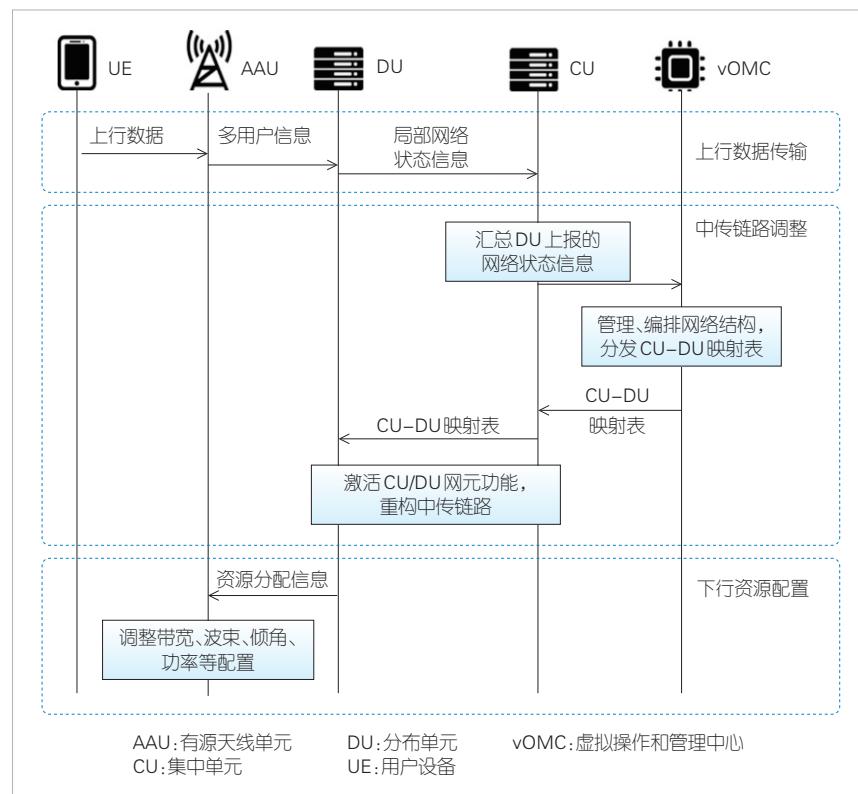
资源小区覆盖结构<sup>[3]</sup>是通过控制/数据平面分离和系统级资源管理实现超高数据密度服务的关键覆盖结构之一。在gNodeB架构中，CU负责处理非实时协议和服务，DU负责处理实时协议和物理层协议，有源天线单元（AAU）负责射频传输以及部分物理层功能的实现。其中，CU与DU之间通过中传链路（光纤）连接。控制层实现覆盖结构灵活适变的本质在于CU-DU之间中传链路的调整，即在形成资源小区基本空间单元后，为符合现有5G gNodeB CU/DU分离的网络结构需求，每个资源小区需选择一个基站激活CU模块功能，其余基站激活DU模块功能。

在实际部署中，可由一个中央处理单元

集中控制区域内的全部基站，由区域级计算中心执行计算任务。该处理单元部署于接入网侧虚拟操作和管理中心（vOMC）。资源小区覆盖结构生成的协议流程如图1所示。基站根据用户发送的上行导频序列与信令信息进行局部区域内的网络状态（如信道状态信息、用户业务信息等）特征提取与压缩。压缩后的信息通过回程链路传递至vOMC，然后vOMC综合接收到的信息并决定资源小区内的基站集合，在所有的信息处理完毕后，将结果用统一、简单的信令表达并分发至基站。基站读取信令后调整中传链路，以形成资源小区覆盖结构。资源小区与其他传统覆盖结构的本质区别在于：资源小区既能在宏观上实现覆盖结构随业务分布变化的灵活性和适应性，又能在微观上为覆盖范围内的用户准确分配资源，实现最终的干扰管理。

### 1.2 系统模型

我们考虑包括 $M$ 个基站、 $N$ 个单天线用户和 $S$ 个资源小区的二维平面上的超密集网络，其中基站配备 $A$ 根天线的均匀线性阵列（ULA）。基站集合为 $\mathcal{M} = \{1, 2, \dots, M\}$ ，用户集合为 $\mathcal{N} = \{1, 2, \dots, N\}$ ，资源小区集合为 $\mathcal{RC} = \{C_1, C_2, \dots, C_S\}$ ，任意资源小区内的基站集合为 $C_s^{\text{AP}} = \{1, 2, \dots, M_s\}$ ，小区内资源块集合为 $\mathcal{K}_s = \{1, 2, \dots, K\}$ ，小区内用户集合为 $C_s^{\text{UE}} =$



▲图1 资源小区覆盖结构生成的协议流程<sup>[3]</sup>

$\{1, 2, \dots, N_s\}$ 。

小区内基站与用户之间的映射关系可表示为：

$$x_{u,i,k}^{C_s} = \begin{cases} 1, & \text{用户 } u \text{ 占用基站 } i \text{ 的资源块 } k \\ 0, & \text{用户 } u \text{ 不占用基站 } i \text{ 的资源块 } k \end{cases} \quad (1)$$

由于每个基站只有  $K$  个独立正交资源块，因此在一个时隙内资源小区  $C_s$  最多能服务的用户数不超过  $M_s \times K$  个，其中  $M_s$  表示小区内的基站数，即：

$$\sum_{i=1}^{M_s} \sum_{u=1}^{N_s} \sum_{k=1}^K x_{u,i,k}^{C_s} \leq K \cdot M_s \quad (2)$$

资源块  $k$  在一个时隙内被复用的次数不会超过资源小区内的基站数  $M_s$ ，即：

$$\sum_{i=1}^{M_s} \sum_{u=1}^{N_s} x_{u,i,k}^{C_s} \leq M_s \quad (3)$$

此外，小区内的用户在一个时隙内至多被来自不同基站的正交  $M_s$  个资源块所服务，即：

$$\sum_{i=1}^{M_s} \sum_{k=1}^K x_{u,i,k}^{C_s} \leq M_s \quad (4)$$

根据小区内基站与用户的映射关系，用户  $u$  与所关联的基站  $i$  在资源块  $k$  的信干噪比表示为：

$$\text{SINR}_{u,i,k}^{C_s} = \frac{x_{u,i,k}^{C_s} P_{u,i,k} |\mathbf{h}_{u,i}|^2}{\sigma_0^2 + \sum_{j \in \mathcal{D}_i \setminus C_s} (1 - x_{u,j,k}^{C_s}) P_{u,j,k} |\mathbf{h}_{u,j}|^2}, \quad (5)$$

其中， $\mathcal{D}_i = \{j | d_{ij} \leq \varepsilon\}$  表示与用户  $u$  之间距离不大于  $\varepsilon$  的所有基站集合， $P_{u,i,k}$  表示基站  $i$  在资源块  $k$  上为用户  $u$  分配的功率， $\mathbf{h}_{u,i}$  表示基站  $i$  与用户  $u$  之间的信道矢量， $\sigma_0^2$  为用户端的加性高斯白噪声。基于多径空间相关信道模型， $\mathbf{h}_{u,i}$  可以表示为：

$$\mathbf{h}_{u,i} = \frac{1}{\sqrt{L}} \sum_{l=1}^L \beta_{u,l} \alpha_{u,i,l} \mathbf{e}(\theta_{u,i,l}), \quad (6)$$

其中， $L$  表示多径分量数， $\alpha_{u,i,l}$  是用户  $u$  与基站  $i$  之间第  $l$  条径的信道的小尺度衰落增益，其被建模为服从零均值方差为 1 的复高斯随机分布。 $\beta_{u,i}$  表示无线信道传播过程中的大尺度衰落，主要包括路径损耗与阴影效应，建模为：

$$\beta_{u,i} = \frac{z_{u,i}}{d_{u,i}^\gamma}, \quad (7)$$

其中， $d_{u,i}$  为用户  $u$  和基站  $i$  之间的距离， $\gamma$  表示经验衰落系数，一般取值为 3.8。另外  $z_k$  为用户  $u$  与基站  $i$  之间的阴影衰

落，满足  $10\log_{10}(z_{u,i})$  服从  $\mathcal{CN}(0, \zeta^2)$  分布， $\zeta$  为阴影衰落系数，一般取值为 8 dB。 $\mathbf{e}(\theta_{u,i,l})$  是基站端 ULA 的导向矢量，天线之间的距离为  $d = \lambda/2$ ，其中  $\lambda$  为电磁波波长。导向矢量可以表示为：

$$\mathbf{e}(\theta_{u,i,l}) = \left[ 1, e^{-j2\pi \frac{d}{\lambda} \sin(\theta_{u,i,l})}, \dots, e^{-j2\pi \frac{d}{\lambda} \sin((\theta_{u,i,l})(A-1))} \right]^T, \quad (8)$$

其中， $\theta_{u,i,l}$  为用户  $u$  与基站  $i$  间信道第  $l$  条径的到达角，对于任意  $u$  和  $i$  满足  $\theta_{u,i,l} \in [-\pi/2, \pi/2]$ 。用户  $u$  和基站  $i$  间  $L$  条径的用户到达角 (AoA) 都服从均匀分布  $U(\theta_{u,i} - \sqrt{3}\sigma, \theta_{u,i} + \sqrt{3}\sigma)$ ， $\theta_{u,i}$  表示用户  $u$  和基站  $i$  之间的入射角， $\sigma$  表示多径信道 AoA 的角度标准差。

用户端的下行速率可以表示为：

$$R_u^{Cs} = W \sum_{k=1}^K \log_2 \left( 1 + \sum_{i=1}^{M_s} \text{SINR}_{u,i,k}^{C_s} \right), \quad (9)$$

其中， $W$  为系统预先为每个资源块设定好的带宽， $R_u^{Cs}$  需要满足最低速率要求  $R_{\min}$ 。

对系统中的每个资源小区下的每个用户端的下行速率进行求和，得到整网下行速率：

$$R_{\text{total}} = \sum_{s=1}^S \sum_{u=1}^{N_s} R_u^{Cs}. \quad (10)$$

我们采用理想的网络功耗模型，即只考虑基站的发射功率<sup>[7]</sup>。假设将基站的总发射功率平均分配给用户，以约束整个网络的总功耗。基站  $i$  在资源块  $k$  上分配给用户  $u$  的发射功率为：

$$P_{u,i,k} = P_{\max} / \sum_{u=1}^{N_s} \sum_{k=1}^K x_{u,i,k}^{C_s}, \quad (11)$$

其中， $P_{\max}$  表示基站的最大发射功率。

基于上述假设条件与模型设计，可以得到整网下行速率优化模型。

$$\begin{aligned} \max_{\text{RC}, X} \quad & R_{\text{total}} = \sum_{s=1}^S \sum_{u=1}^{N_s} R_u^{Cs}, \\ \text{s.t.} \quad & \text{C1: } \sum_{i=1}^{M_s} \sum_{u=1}^{N_s} \sum_{k=1}^K x_{u,i,k}^{C_s} \leq K \cdot M_s, \forall C_s \in \text{RC}, \\ & \text{C2: } \sum_{i=1}^{M_s} \sum_{u=1}^{N_s} x_{u,i,k}^{C_s} \leq M_s, \forall C_s \in \text{RC}, \forall k \in K, \\ & \text{C3: } \sum_{i=1}^{M_s} \sum_{k=1}^K x_{u,i,k}^{C_s} \leq M_s, \forall C_s \in \text{RC}, \forall u \in N_s, \\ & \text{C4: } R_u^{Cs} \geq R_{\min}, \forall C_s \in \text{RC}, \forall u \in N_s, \end{aligned}$$

$$\begin{aligned} C5: & x_{u,i,k}^{C_s} = \{0, 1\}, \forall C_s \in RC, \forall u \in N_s, \forall i \in M_s, \forall k \in K, \\ C6: & \sum_{u=1}^{N_s} \sum_{k=1}^K P_{u,i,k} = P_{\max}, \forall C_s \in RC, \forall u \in N_s, \forall i \in M_s, \forall k \in K. \end{aligned} \quad (12)$$

## 2 基于多点信道海图的资源小区覆盖优化

针对未来无线通信系统需要满足移动用户对高速数据传输业务的需求，由于业务时空分布动态变化，缺乏适配动态业务灵活性的覆盖结构，将会导致资源的使用冲突和网络中复杂干扰，降低局部区域内的网络容量。受资源小区<sup>[3]</sup>的启发，本文中我们提出基于多点信道海图的资源小区覆盖结构优化方法。

由于用户间干扰与不同用户信道域之间的相关性密切相关，因此本文主要思想是通过提取无线信道特征构建多点信道海图并学习不同用户信道之间的相关性，从而将相关性强的用户划分为同一簇，将其相关联的基站划分到同一个资源小区，资源小区内基站进行联合传输并给用户分配正交资源块，以降低用户间干扰，资源小区间进行资源块复用，从而提升系统和速率。基于多点信道海图的资源小区覆盖结构生成框架如图2所示。

### 2.1 多点信道海图生成

图2虚线部分展示了多点信道海图的构造过程。首先，每个基站（BS）从所有用户信道状态信息  $H^m \in \mathbb{C}^{N \times P}$  中提取特征  $F^m$ ，计算局部相异度矩阵  $D^m \in \mathbb{C}^{N \times N}$ ，其中  $N$  和  $P$  分别是用户数和CSI的维数。不同相异度矩阵  $D^m$  表示基站  $m$  处CSI的局部视图，与用户间的无线电距离有关。然后，将  $D^m$  上报给vOMC并进行数据融合，从而构建全局相异度矩阵  $D \in \mathbb{C}^{N \times N}$ 。最后，通过主成分分析（PCA）<sup>[8]</sup>、Sammon's

Mapping (SM)<sup>[9]</sup>等流形学习方法将  $D$  变换成满足  $N_d \ll N$  的低维图  $Z \in \mathbb{C}^{N \times N_d}$ 。大多数信道海图的工作将  $N_d$  设置为 2。信道海图技术的核心思想可以表达为：在低维空间找到一组数据点  $\{z\}_{n=1}^N$ 。当  $d_p(\mathbf{p}_i, \mathbf{p}_j)$  较小时，存在：

$$d_z(z_i, z_j) \approx d_p(\mathbf{p}_i, \mathbf{p}_j), \quad (13)$$

其中  $\mathbf{p}_i$  和  $\mathbf{p}_j$  为原始空间中距离较近的两个数据点的三维坐标， $z_i$  和  $z_j$  分别为  $\mathbf{p}_i$  和  $\mathbf{p}_j$  降维后在海图空间的二维坐标。海图空间中的坐标没有绝对的物理意义，但是可以通过  $d_z(z_i, z_j)$  表示海图上数据点之间的差异程度。当  $d_p$  和  $d_z$  表示欧式距离计算且满足式 (13) 时，说明所构造的信道海图保留了数据点在原始空间中的局部相对位置关系。

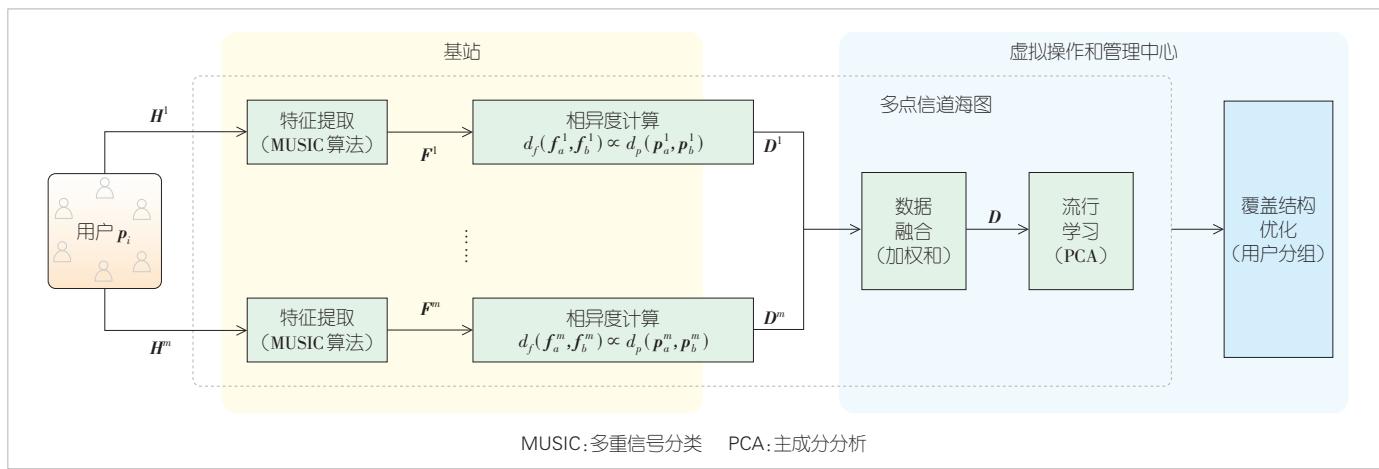
从信道状态信息中提取的特征是决定MPCC提供何种信息，以及其是否能有效解决通信问题的关键。由于空间几何上接近的位置会经历类似的多径分量，因此，可以从信道状态信息中挖掘多径分量的到达角  $\theta_{u,i,l}$  和信道增益  $\beta_{u,i}$  来构造MPCC。本文采用多重信号分类（MUSIC）算法进行特征提取<sup>[10]</sup>，其输入是上行链路CSI的协方差矩阵  $R = \mathbb{E}\{HH^H\}$ ，其携带与距离有关的大规模衰落特性的信息。在实际中，一般信道协方差矩阵计算方法为：

$$R = \frac{1}{T} \sum_{t=0}^{T-1} HH^H. \quad (14)$$

基于这些多径特征进行用户间的相异度度量可揭示信道图上用户间的近邻关系。在虚拟操作和管理中心融合各个基站上传的相异度矩阵，并进行降维，可实现其到信道海图的线性映射。

### 2.2 覆盖结构优化策略

系统模型假设用户与资源小区已匹配，即：任意用户  $u$



▲图2 基于多点信道海图的资源小区覆盖结构优化流程

只属于一个用户集合  $N_s$ ，且该用户集合  $N_s$  中的所有用户只被对应的资源小区  $C_s$  提供服务，因此可通过用户分组来优化资源小区覆盖结构。

用户在空间几何上的距离越近，用户之间的干扰相对较大。MPCC 上用户间的距离反映用户的最近邻关系，因此两个用户在 MPCC 上距离越近，其发生同频干扰的可能性就越大。基于 MPCC 的 K-means 聚类后，这些用户以更大概率被分配到同一簇（定义为资源小区）。首先，通过 K-means 算法形成初始用户簇，其中初始用户簇数量  $K_0$  定义为  $\min(\lceil N/K \rceil, M)$ 。为确保每个资源小区中有足够的正交资源块用于分配，我们将  $K_0$  设置为  $\lceil N/K \rceil$ 。随着用户数的增加， $\lceil N/K \rceil$  可能大于基站数量，因此将  $K_0$  设置为两者的最小值。在此基础上，根据用户最大参考信号接收功率（Max-RSRP）关联基站的原则，得到初始资源小区集。

为确保资源小区是完全独立的，我们需要若干验证。例如：1) 当不同资源小区之间存在相同基站时，该基站被分配给所关联用户最多的资源小区；2) 检查每个资源小区中的用户数量，以确保每小区中有足够的正交资源块用于分配。当资源小区中的用户数超过阈值  $N_{\text{thre}}$  时，就需要进行细粒度分区。具体来说，当资源小区  $C_s$  中的用户数量  $N_s$  大于  $M_s \times K$  时，需要将该资源小区  $C_s$  中的所有基站划分为独立的资源小区；当  $N_s$  大于细粒度分区阈值  $N_{\text{thre}}$  但小于  $M_s \times K$  时，资源小区  $C_s$  应根据用户数量  $N_s$  分裂为更多的资源小区。基于 MPCC 的资源小区生成算法的具体步骤如算法 1 所示。

#### 算法 1 基于多点信道海图的资源小区覆盖结构优化策略

```

输入：多点信道海图  $Z \in C^{N \times N_d}$ 
输出：资源小区集合  $\mathbf{RC} = \{C_1, C_2, \dots, C_s\}$ 
1: 用户根据 Max-RSRP 原则关联基站，得到初始用户 - 基站关联对
2: 计算聚类的数目  $S = K_0$ 
3: 基于多点信道海图通过 K-means 聚类算法初始化用户簇
4: if 关联同一个基站的不同用户被划分到不同的用户簇 then
5:   将基站划分到其关联用户最多的簇
6: end if
7: 输出初始资源小区集  $\mathbf{RC}_{\text{spre}} = \{C_1, C_2, \dots, C_s\}$ 
8: for 资源小区  $C_s$ ,  $s=1, 2, \dots, S$  do
9:   if  $N_s \geq M_s \times K$  then
10:    将  $C_s$  中的所有基站都视为一个独立的资源小

```

区，并分别输入到资源小区集合  $\mathbf{RC}$  中

```

11:   else if  $N_{\text{thre}} \leq N_s \&& N_s < M_s \times K$  then
12:     将  $C_s$  中的基站关联的所有用户 MPCC 向量作为输入，执行步骤 3—7，其中  $S = \lceil N_s/K \rceil$ ，将输出的资源小区输入到资源小区集合  $\mathbf{RC}$  中
13:   else
14:     将  $C_s$  输入到资源小区集合  $\mathbf{RC}$  中
15:   end if
16: end for
17: 输出资源小区集合  $\mathbf{RC}$ 

```

### 3 仿真结果分析

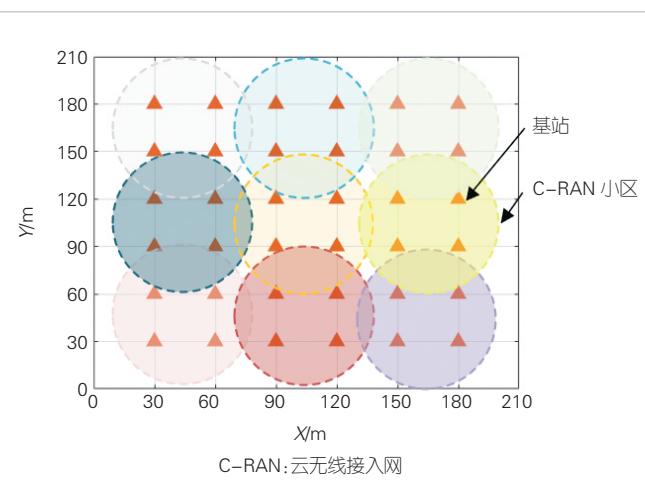
#### 3.1 实验设置

本部分内容针对所提的基于 MPCC 的资源小区优化策略，在能耗约束下评估系统和速率来验证性能。为了简单而又不失一般性，我们考虑一个  $210 \text{ m} \times 210 \text{ m}$  场景，其中基站均匀分布。参数汇总如表 1 所示。

我们以图 3 所示 C-RAN 静态覆盖结构为基准，采用基

▼表 1 仿真实验参数表

参数	值
基站数目	36
单个资源块带宽	0.36 MHz
资源块数目	100
噪声功率	-174 dBm/Hz
多径数目	3
基站天线数目	16
蒙特卡洛次数	10



▲图 3 传统 C-RAN 静态覆盖结构

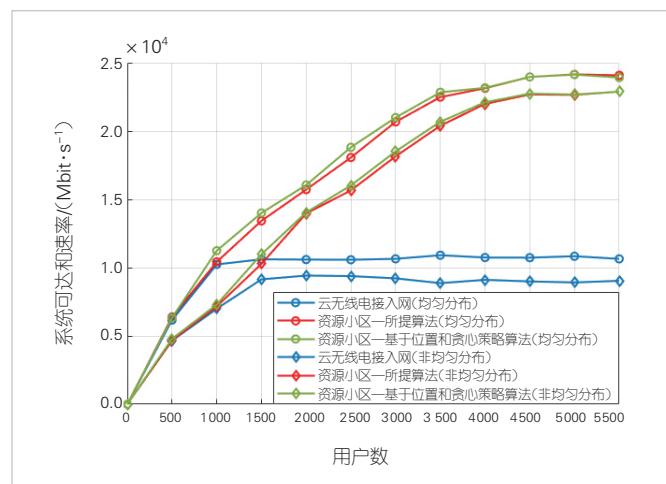
于位置和贪心策略的优化方法<sup>[3]</sup>作为对比，下文中将其称为LG算法。LG方法使用基于用户地理位置的K-means算法形成用户初始簇，基于用户地理位置通过K-means算法进行用户分簇形成初始用户簇，在此基础上进行贪心迭代形成资源小区，使系统和速率达到最大。

### 3.2 实验结果及分析

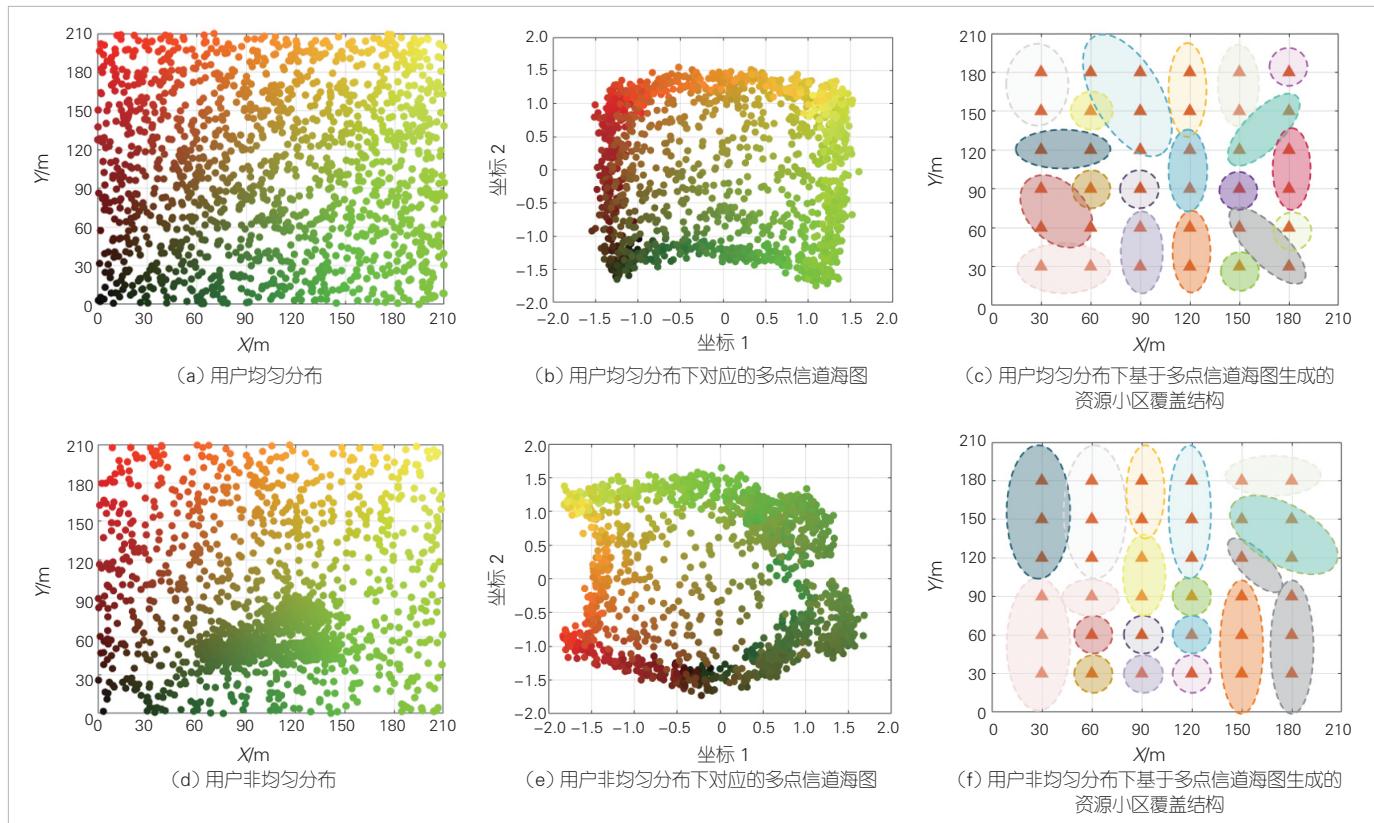
图4为资源小区生成示意图，其中用户数为1500。图4(a)展示了用户均匀分布情况，其中用户根据均匀点过程独立生成。不同颜色的点代表不同空间位置的用户，用户间标记的颜色越相似，其在通信系统中的实际物理距离就越近。图4(b)多点信道海图显示，用户在海图上的位置分布形状与原始位置相似，证明MPCC可有效保持信道域中用户之间的最近邻关系。图4(c)中展示了基于多点信道海图的资源小区优化算法得到的资源小区覆盖结构，其中不同的颜色分别代表不同的资源小区。图4(d)和图4(e)分别为用户非均匀分布情况和对应的多点信道海图。其中，600个用户是由泊松点过程独立生成，其余用户均服从均匀分布。从直观上看，虽然MPCC中用户的空间位置分布与原始位置相比发生了变化，但海图转换后用户的颜色标记并没有出现

明显的混叠现象。为了更直观地反映资源小区的覆盖效果，图4(f)展示了所提算法优化得到的资源小区。可以看出，在用户密集分布的区域，资源小区尽可能分裂出更多的资源小区，而在用户稀疏分布时，尽可能多地合并接入点，从而发挥BS协作的优势。

图5展示了系统下行速率与用户数的关系。随着用户数量的增加，可实现的系统下行速率在初始阶段呈上升趋



▲图5 系统可达和速率随用户数的变化趋势( $P_{\max} = 10 \text{ W}$ )



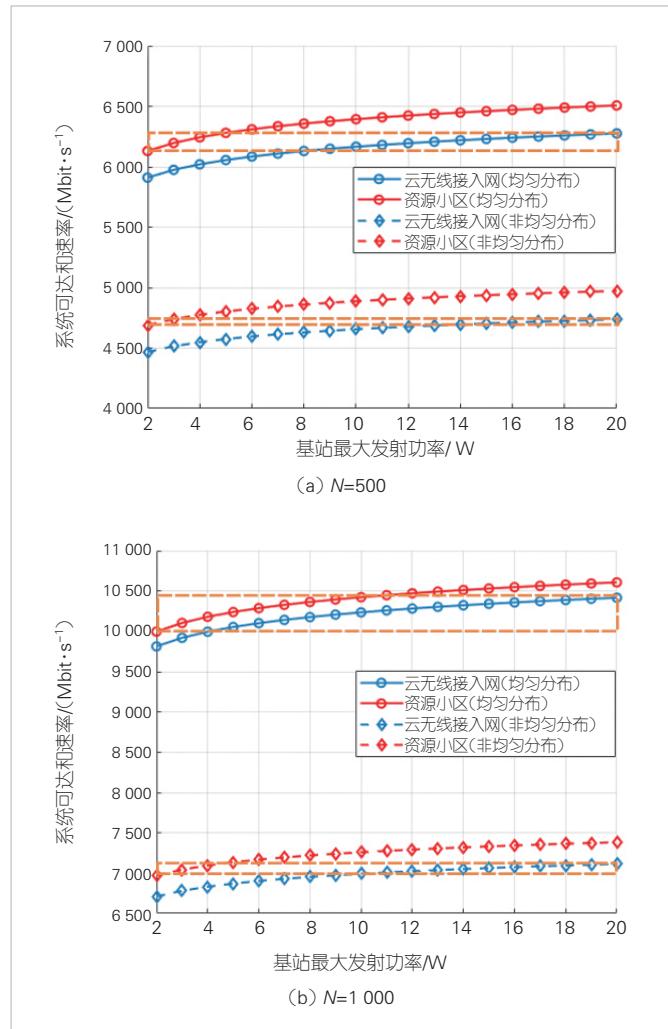
▲图4 资源小区生成示意图( $N=1500$ )

势，随后逐渐趋于稳定。值得注意的是，随着用户数增多，用户速率需求也在变化，即分为两种业务场景：用户连接数少但速率需求高，用户连接数多但速率需求低。9个小区的C-RAN静态覆盖结构中，由于正交资源块数量的限制，全网最大接入用户数为 $9 \times 100 = 900$ 。因此，在用户均匀分布的情况下，接入用户数超过1 000个时，C-RAN静态覆盖结构的网络承载能力将达到上限，而用户非均匀分布时，在超过用户数1 500的情况下C-RAN静态覆盖结构不再有增益。对于资源小区这种弹性覆盖结构，小区的数量随用户数量的增加而灵活变化。当用户数量不断增加时，原始小区开始分裂为更多的资源小区，以确保允许更多的用户访问，直到超过4 500个用户时接近承载能力的上限。绿色实线代表基于瞬时CSI的LG方法的性能，红色实线代表我们所提的利用平均CSI基于MPCC的算法。可以看出，本文算法性能略差或相当于基于位置的贪心策略的优化方法，但都优于C-RAN静态覆盖结构（蓝色实线）。在用户均匀分布的情况下，本文方法与LG方法性能相比降低最大为7.8%，此时用户数为1 000，其相比于C-RAN性能提升1.84%；当用户数为5 500时，本文算法和LG算法性能相当，相比于C-RAN系统和速率分布提升126.27%。在用户非均匀分布情况下，用户数为1 500时，本文方法与LG方法性能相差最大为6.35%，此时，其相比于C-RAN性能提升12.45%；当用户数为5 500时，本文算法和LG算法性能相当，相比于C-RAN系统和速率提升153.23%。

通过提高基站总发射功率可以提高系统和速率，图6给出了 $N=500$ 和 $N=1\ 000$ 时系统可达和速率与基站最大发射功率对应关系。对于两种覆盖结构可达的系统和速率区间（橙色虚线框），资源小区的某个系统和速率 $R_{RC}$ 对应的基站最大发射功率为 $P_{RC}$ ，找到 $R_{RC} \pm 5$ 范围内的 $R_{C-RAN}$ ，其对应的基站最大发射功率为 $P_{C-RAN}$ ，然后再按照公式 $\eta = \frac{P_{C-RAN} - P_{RC}}{P_{C-RAN}} \times 100\%$ 进行计算，可以得到资源小区相比于C-RAN达到相同系统和速率的能耗降低百分比 $\eta$ 。 $N=500$ 时，在用户均匀分布情况下，资源小区相比于C-RAN达到相同系统和速率的能耗约为75.66%；在用户非均匀分布情况下，能耗降低约为85.46%。 $N=1\ 000$ 时，资源小区的能耗降低百分比在用户均匀分布和分均匀分布两种情况下分别为69.05%和77.90%。

## 4 结束语

针对日益增长的用户速率需求，我们提出基于多点信道海图的资源小区覆盖结构优化策略，通过构建信道海图来感



▲图6 系统可达和速率与基站最大发射功率对应关系

知用户空间分布，将信道特征相关性强的用户划分到同一资源小区。资源小区内基站对用户进行联合传输并分配正交资源块，以减少干扰，提升用户速率。资源小区间通过正交资源块复用以提升系统和速率。仿真结果表明，在相同基站总发射功率约束下，所提方法优化得到的资源小区弹性覆盖结构对比静态C-RAN覆盖结构有效提升了系统和速率。在能耗方面，在系统达到相同系统和速率的情况下，资源小区弹性覆盖结构有效降低了所需的基站发射功率。本文所提算法能够解决现有基于贪婪的资源小区覆盖结构优化算法中用户位置很难获取、瞬时信道状态信息不完美的问题，对网络覆盖结构优化具有重要参考价值。

## 致谢

感谢中国科学技术大学陈玥珊、张佳杰同学对本论文做出的贡献！

## 参考文献

- [1] YOU X H, WANG C X, HUANG J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts [J]. *Science China information sciences*, 2021, 64(1): 110301. DOI: 10.1007/s11432-020-2955-6
- [2] CHEN Y, ZHANG S Q, XU S G, et al. Fundamental trade-offs on green wireless networks [J]. *IEEE communications magazine*, 2011, 49(6): 30–37. DOI: 10.1109/MCOM.2011.5783982
- [3] 张夏雨, 李建东, 刘俊宇, 等. 资源小区: 面向下一代超密集网络的无线覆盖结构 [J]. 物联网学报, 2023, 7(1): 1–17. DOI: 10.11959/j.issn.2096-3750.2023.00336
- [4] COLORNI A, DORIGO M, MAFFIOLI F, et al. Heuristics from nature for hard combinatorial optimization problems [J]. *International transactions in operational research*, 1996, 3(1): 1–21. DOI: 10.1111/j.1475-3995.1996.tb00032.x
- [5] GARCIA N, WYMEERSCH H, LARSSON E G, et al. Direct localization for massive MIMO [J]. *IEEE transactions on signal processing*, 2017, 65(10): 2475–2487. DOI: 10.1109/TSP.2017.2666779
- [6] DAHLMAN E, PARKVALL S, SKÖLD J. *Background of lte//4G: LTE/LTE-Advanced for Mobile Broadband*. Amsterdam: Elsevier, 2014: 1–15.. DOI: 10.1016/b978-0-12-419985-9.00001-5
- [7] KARL P. On lines and planes of closest fit to systems of points in space [J]. *The london, Edinburgh, and Dublin philosophical magazine and journal of science*, 1901, 2(11): 559–572. DOI: 10.1080/14786440109462720
- [8] SAMMON J W. A nonlinear mapping for data structure analysis [J]. *IEEE transactions on computers*, 1969, C-18(5): 401–409. DOI: 10.1109/TC.1969.222678
- [9] SCHMIDT R. Multiple emitter location and signal parameter estimation [J]. *IEEE transactions on antennas and propagation*, 1986, 34(3): 276–280. DOI: 10.1109/TAP.1986.1143830
- [10] HAN S Q, YANG C Y, MOLISCH A F. Traffic-aware base station doze in cooperative multicell systems [C]//*Proceedings of 2013 IEEE China Summit and International Conference on Signal and Information Processing*. IEEE, 2013: 394–398. DOI: 10.1109/ChinaSIP.2013.6625368

## 作者简介



叶远青, 中国科学技术大学在读硕士研究生; 主要研究领域为无线通信网络和资源管理。



张四海, 中国科学技术大学副教授; 主要研究领域为无线大数据、机器学习和IC工艺优化等; 先后主持和参加项目10余项, 已发表论文80余篇。



朱近康, 中国科学技术大学教授; 主要研究领域为无线移动通信与网络、通信信号处理及未来无线技术研究; 先后主持和参加项目20余项, 出版4本专著, 已发表论文150余篇。



# 面向5G-A的无线网络节能关键技术

## Key Technologies for 5G-A Wireless Network Energy Saving

郭诚/GUO Cheng, 陈梦竹/CHEN Mengzhu

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

DOI: 10.12142/ZTETJ.202306003

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231211.1116.004.html>

网络出版日期: 2023-12-12

收稿日期: 2023-10-15

**摘要:** 建立网络能耗模型, 对基站在数据发送状态、接收状态、休眠模式及模式转换的相对功耗进行建模, 可以评估各种网络节能技术的节能增益。面向5G-A, 从时域、空域、频域和功率域等方面, 研究包括公共信号轻量化设计、极简小区架构、小区不连续发送/接收模式、基站唤醒机制、节能测量上报增强等网络节能的关键技术, 使能更动态、高效、精准的无线传输, 并通过终端和基站间信息交互的方式提升节能效果。随后, 仿真验证了这些技术的正确性和有效性。

**关键词:** 网络节能; 能耗建模; 极简小区架构; 测量上报增强

**Abstract:** The network energy consumption model is established to analyze the relative power consumption of base stations for transmission/reception, sleep modes, and mode transmission, and evaluate the energy-saving gains of various network energy-saving techniques. In terms of time domain, space domain, frequency domain, and power domain, key techniques for network energy saving, including light-weight design of common signals, simplified cell architecture, cell discontinuous transmission/reception mode, base station wake-up mechanism, enhanced channel state information measurement/reporting for network energy-saving are researched, and enabling more dynamic, efficient, and accurate wireless transmission. In addition, the energy-saving gain is improved through information interaction between the device terminal and the base station, and then the effectiveness of the techniques is verified and analyzed through system-level simulation.

**Keywords:** network energy saving; energy consumption modeling; simplified cell architecture; enhanced measurement report

**引用格式:** 郭诚, 陈梦竹. 面向5G-A的无线网络节能关键技术 [J]. 中兴通讯技术, 2023, 29(6): 11-15. DOI: 10.12142/ZTETJ.202306003

**Citation:** GUO C, CHEN M Z. Key technologies for 5G-A wireless network energy saving [J]. ZTE technology journal, 2023, 29(6): 11-15.  
DOI: 10.12142/ZTETJ.202306003

### 1 5G-A无线网络节能演进

全球气候变化正在对全人类生存发展带来日益严峻的影响, 走向碳中和已成为全球的共识。作为数字化转型重要底座的移动通信网络, 聚焦数据中心、通信基站、通信机房3类重点设施。工业和信息化部发布的《信息通信行业绿色低碳发展行动计划(2022—2025年)》<sup>[1]</sup>要求推动基站主设备及配套设施节能技术的应用推广, 使5G基站能效提升20%以上。通过信息通信行业数字化转型来赋能全社会降碳, 促进居民低碳环保生活, 从而更好地推动城乡绿色智慧发展。

电信运营商及设备商自2G时代起就不断启动节能减排计划, 研究网络节能手段, 应对气候及能耗的挑战。在过去

的时间里, 随着产业链的不断发展与成熟, 芯片技术的不断提高, 通过不断优化电路设计、硬件实现算法, 网络节能已经取得了很大的进步。一些基础节能功能包含符号关断、通道关断、载波关断和设备深度休眠等被提出, 以适应不同的通信网络负荷<sup>[2]</sup>。

随着5G-A时代到来, 无线通信将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, 同时全息通信、智能交互、感官互联、数字孪生、通信感知等新业务需求不断涌现, 这给网络流量带来了百倍甚至千倍的激增。要想在保持比特增长的同时尽量降低功耗, 需要基于高集成芯片、高效率器件、新升级工艺、强算力算法, 打造新一代绿色设备, 实现极致的关断深度和响应时间<sup>[3]</sup>。

1) 极致关断深度。无线软件节能技术的核心理念为通过关闭部分无线能力,降低网络冗余容量,使得网络能耗匹配分布在在网络中的话务负荷。未来,通过硬件模块化设计和软件可弹缩架构设计、精细化设备休眠机制,持续提升器件关断深度,从而实现硬件功耗随负载逐渐逼近线性变化,最大程度降低中低负载下设备能耗。

2) 极致响应时间。节能状态下的硬件响应时间是影响网络指标和用户体验的关键因素。硬件响应时间将实现分钟级到毫秒级的跨越,使得从仅闲时节能扩展到全天都可节能。

在低能耗硬件基础之上,面向5G-A,我们将从时域、空域、频域和功率域等方面展开网络节能新一轮的研究工作,通过定义基站节能状态和节能模式,用更准确的能耗模型,评估更加灵活的节能策略和关断方式,最终赋能百业,共建绿色生态。

## 2 网络能耗建模

### 2.1 基站节能模型

为了评估节能方案的节能增益,我们设深休眠状态下的基站功耗为单位1,定义了不同状态下的基站功耗模型<sup>[4]</sup>,具体如表1。

在没有数据发送和接收情况下,基站可以关闭部分器件,进入休眠状态。根据节能效果,我们可将基站休眠状态分为深休眠、浅休眠和微休眠。鉴于不同休眠状态下关闭器件的不同,基站唤醒时间和状态切换的能耗也有所差异,基站激活态和休眠态转换如图1所示,其中微休眠可以认为转换时间T和能耗E忽略不计。表2分别给出了深休眠和浅休眠的转换时间T和能耗E<sup>[4]</sup>。

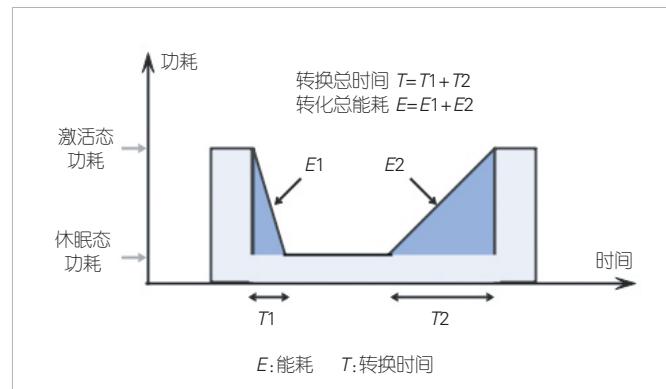
### 2.2 基站能耗模型

基站功耗可以分为静态功耗和动态功耗,其中静态功耗是指与业务负载和输出传输功率无关的能耗占用,而动态功耗随着负荷增加而增加。

▼表1 参考配置下基站功耗模型

节能状态	FR1 TDD	FR1 FDD	FR2 TDD
深休眠P1		1	
浅休眠P2		2.1	
微休眠P3	5.5	5	3
下行激活状态P4	32	26	17.6
上行激活状态P5	6.5	5.8	4.2

FDD:频分双工 FR:频率范围 TDD:时分双工



▲图1 基站激活态和休眠态转换

▼表2 休眠态的转换总时间和总能耗

节能状态	转换总时间	转换总相对能耗
深休眠	10 s	17 000
浅休眠	640 ms	1 088

$$P_{\text{dynamic}}^{\text{DL}} = P_{\text{static}}^{\text{DL}} + P_{\text{dynamic}}^{\text{DL}}, \quad (1)$$

$$P_{\text{dynamic}}^{\text{UL}} = P_{\text{static}}^{\text{UL}} + P_{\text{dynamic}}^{\text{UL}}, \quad (2)$$

$$\text{其中, } P_{\text{static}}^{\text{DL}} = P_{\text{static}}^{\text{UL}} = P_3.$$

为了进一步深入分析不同负载对基站功耗的影响,我们定义了基站动态功耗在随不同因素变化时的计算公式<sup>[4]</sup>。在下行激活状态下,基站动态功耗部分计算公式为:

$$P_{\text{dynamic}}^{\text{DL}} = s_a \times \left( P_{\text{dyn, ante}} + \frac{s_f \times s_p}{\eta(s_f, s_p)} \times P_{\text{dyn, joint}} \right). \quad (3)$$

仅和通道数相关的部分功耗计算公式为:

$$P_{\text{dyn, ante}} = \alpha \times (P_4 - P_{\text{static}}^{\text{DL}}). \quad (4)$$

和通道数、带宽、功率谱密度等多因素相关的部分功耗计算公式为:

$$P_{\text{dyn, joint}} / \eta(s_f = 1, s_p = 1) = (1 - \alpha) \times (P_4 - P_{\text{static}}^{\text{DL}}), \quad (5)$$

其中,  $s_a$ 、 $s_f$ 、 $s_p$  分别是激活的通道数相对基线配置中通道数的比值, 激活的带宽相对基线配置中最大带宽的比值, 每通道的实际发射功率谱密度相对基线配置中功率谱密度的比值,  $\alpha$  为仅和通道数相关的部分能耗占比, 参考业界数据取值为 0.4<sup>[5]</sup>,  $\eta$  为功放 (PA) 效率。

类似地,在上行激活状态下,基站动态功耗部分计算公式为:

$$P_{\text{dynamic}}^{\text{UL}} = s_a \times P_{\text{dyn, ante}}^{\text{UL}}, \quad (6)$$

$$\text{其中, } P_{\text{dyn, ante}}^{\text{UL}} = P_5 - P_3.$$

### 3 网络节能关键技术

#### 3.1 基站不连续发送/接收模式

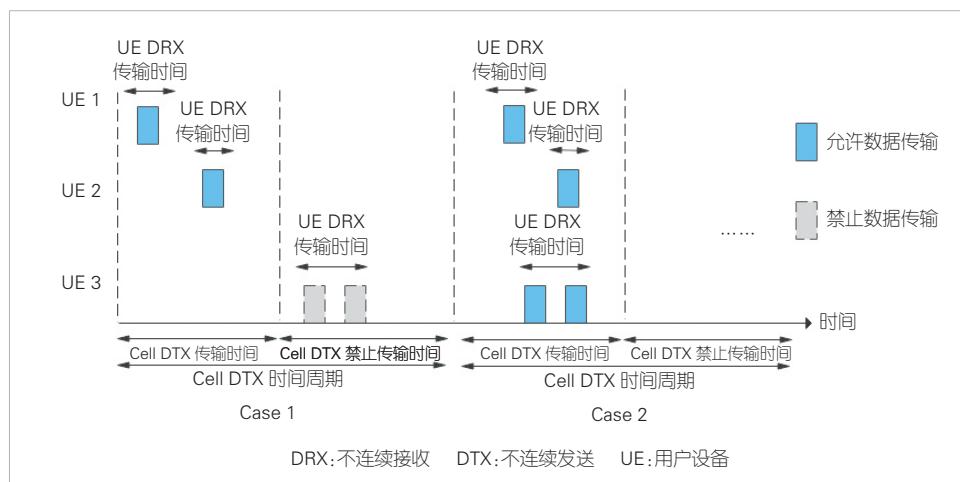
为终端配置不连续接收（DRX）模式以减少不必要的物理下行控制信道（PDCCH）监听功耗，是实现终端节能的有效途径之一<sup>[6]</sup>。与之类似，引入小区不连续发送（DTX）/DRX功能，可以减少基站不必要的下行信号发送/上行信号接收，实现基站节能目的。其中，小区DTX模式用于控制小区的下行传输，小区DRX模式用于控制小区的上行传输。

小区DTX/DRX模式是周期性行为，由终端专用无线资源控制（RRC）信令配置，配置参数包括周期、起始时隙、位置偏移量和持续时间等，支持以小区为粒度配置不同的小区DTX/DRX参数，从而满足不同小区的节能需求。同时，为了适配不同的业务需求，还需要减少信令开销，采用组播信令指示动态激活/去激活小区DTX/DRX配置。

在小区DTX/DRX生效期间，基站周期性地不执行数据业务的发送/接收。为了减小对网络中终端的影响，小区DTX/DRX维持同步信号块（SSB）、接入信道（RACH）、寻呼消息（Paging）和系统消息（SIBs）的正常接收和发送。小区DTX可与终端DRX协同应用，在两者相重叠的传输时间段内，才允许下行数据进行传输。如图2所示，Case1中用户设备（UE）1~UE2均可正常传输数据，UE3因为终端DRX与小区DTX可传输时间段不重叠因此禁止传输数据，应用时需优先按照Case2方式配置，UE1~UE3均可正常传输数据。

#### 3.2 公共信令轻量化传输

基站在持续性不发送信号时，可以关闭器件，进入休眠



状态，以实现节能。但现有网络往往需要小区持续提供服务，那么即便在仅少量数据传输甚至无数据传输时，由于SSB、RACH、Paging和SIBs持续性地接收和发送，使得基站无法有效进入休眠模式。为解决此问题，我们提出了公共信令轻量化传输机制。

对于载波聚合（CA）场景，广播及公共信令可由主小区（Pcell）承载，终端在SSB-less辅小区上的激活流程以及同步定时等依赖于SSB传输的主小区。减少辅小区（Scell）上的SSB传输可以增加基站睡眠机会，从而达到减少基站功耗的目的。

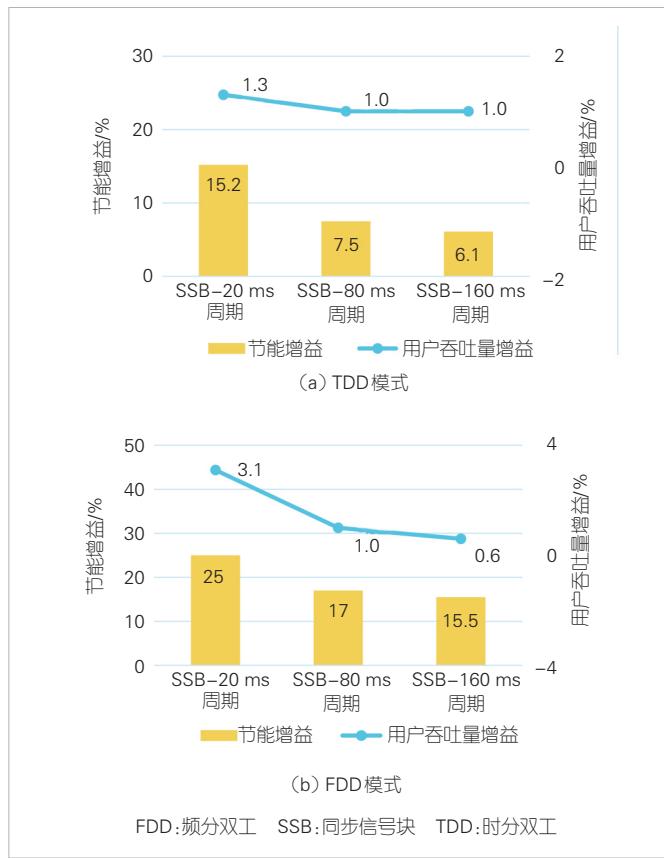
在不同的业务负载条件下，SSB-less辅小区的节能效果不同。图3给出了SSB-less辅小区的节能效果：在低负载场景，时分双工（TDD）模式可以获得6.1%~15.2%的节能增益，频分双工（FDD）模式可以获得15.5%~25%的节能增益。此外，由于不需要发送SSB，基站可以利用更多资源调度数据，仿真结果显示可增加0.6%~3.1%的用户感知吞吐量（UPT）增益。

同样地，这一节能方法也可扩展至非CA场景。为保证终端在节能小区的正常接入，需要引入锚点小区。处于非连接态的终端可以从锚点小区获取所需的系统信息、同步信息和寻呼消息。当终端接入小区时，可以在SIB-less/SSB-less节能小区（即非锚点小区）发起初始接入，并进行连接态业务传输。当节能小区的SIBs在锚点小区上传输时，会导致锚点小区上的功耗少量增加。但综合考虑节能小区和锚点小区的总功耗，我们认为基站仍能够获得可观的节能增益。

基于基站功耗模型，我们对SIB-less/SSB-less节能小区方案进行仿真评估。我们将不同的SSB/SIB传输周期作为对比基线，在不同负载条件下，SSB-less/SIB-less节能小区可以获得9.4%~24.5%的节能增益。由于节能小区的系统信息在锚点小区上进行传输，锚点小区的功耗增加2.3%~7.5%。

#### 3.3 动态通道关断及测量上报增强

在5G现网中，大规模多输入多输出（MIMO）由于其空分复用及发送接收分集的优势被广泛使用。大规模MIMO在带来高容量的同时，其大量的收发通道及相关的硬件处理单元也会带来基站功耗的增加。一种有效的网络节能方案是根据流量负载或UE数量动态调整基站激活通道的数量，



▲图3 SSB-less节能和吞吐量增益仿真结果

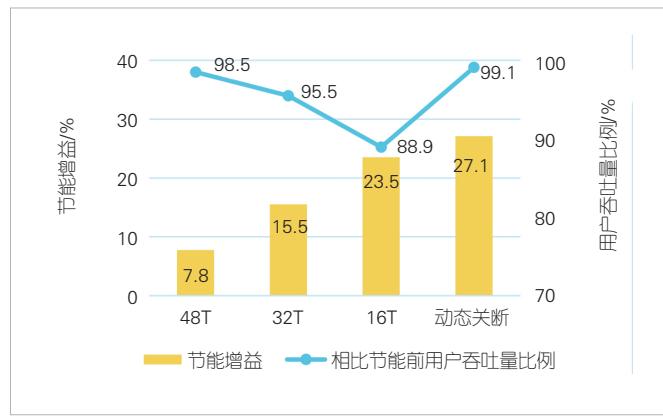
但通道关断会导致基站发射功率和赋形增益降低。为了实现更高的节电增益并保持较低的性能损失，基站需要获取关断前后网络UE的信道状态信息(CSI)，从而评估通道关断的性能影响，辅助决策通道关断的生效判决和生效时机。

基站可以通过UE测量上报的CSI报告获得基站与UE间的信道状态<sup>[7]</sup>，将不同的通道关断状态与CSI测量报告一一对应，从而获取到不同通道关断状态下的信道状态。但过多的CSI报告会增加空口开销，也会导致UE测量能力受限，影响其他用途的CSI报告配置。因此，对于动态通道关断方案中的一个重要研究方向就是CSI测量上报增强。

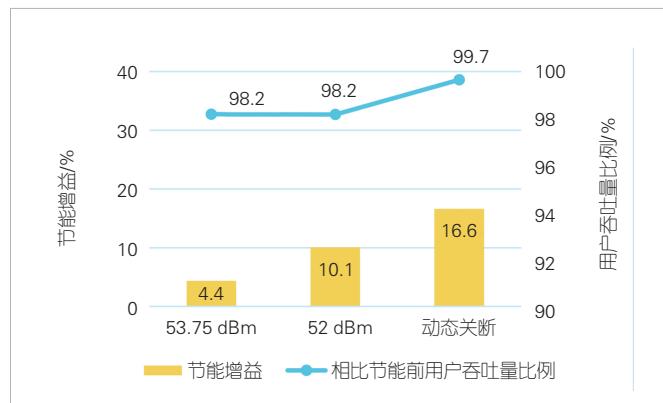
为了在一组CSI报告中包含多组CSI测量，需要引入Multi-CSI测量上报方案，并与多种天线关断状态的参考信号的CSI(CSI-RS)配置进行关联，具体可分为如下两种场景：

场景1：当通道关断导致CSI-RS port数量变化时，不同通道关断状态可共用同一CSI-RS配置。该公共CSI-RS配置对应非天线关断状态，其他CSI-RS配置为该公共CSI-RS配置的子集，对应于有天线关断的状态。

场景2：当通道关断没有导致CSI-RS port数量变化时，



▲图4 通道关断的节能增益和吞吐量仿真结果



▲图5 动态功率调整的节能增益和吞吐量仿真结果

不同通道关断状态的CSI-RS需要独立配置，即每个CSI-RS配置对应一种天线关断状态。

基于64通道基站，我们对动态通道关断方案的节能增益和用户吞吐量(UPT)影响进行了仿真评估。当对比基线方案为静态通道关断时，即静态关断至48通道、32通道、16通道时，节能增益分别为7.77%、15.53%、23.53%，UPT损失为1.5%~11.06%；而通过CSI测量上报增强方案辅助的动态通道关断能提供27.09%基站节能增益，并且UPT损失不到1%。由评估结果也可以看出，基于Multi-CSI测量上报增强的通道关断方案可以在提供较大节能增益的同时带来较少的UPT损失。

### 3.4 动态功率优化及测量上报增强

功率动态调整是功率域节能中的关键技术。基站在与终端交互的过程中，通常使用固定的发射功率进行下行数据发射。如果基站采用较小的发射功率，则可能导致小区边缘UE无法准确的获取传输的信息。基站采用较大的传输功率会带来功耗额外增加，因此，为了降低网络能耗的同时保证数据的正确传输，引入动态的自适应功率调整技术至关

重要。

当对业务信道进行动态功率调整时，可能会导致CSI报告/反馈不准确。为了减少对网络性能的影响，基站需要在功率调整时获得更准确、更实时的CSI信息，因此我们提出为终端配置多个功率偏移值的CSI-RS配置，与不同功率档位的CSI-RS配置进行关联，获取不同功率偏移档位所对应的CSI信息，辅助网络进行动态发射功率调整；并引入Multi-CSI测量上报方案，实现一组CSI报告中包含多组CSI测量。

基于最大发射功率为55 dBm基站，我们对动态功率控制方案的节能增益和UPT进行了仿真评估。当对比基线方案为静态功率回退时，即静态回退功率至53.75 dBm和52 dBm时，基站节能增益分别为4.38%、10.07%，UPT损失不到2%；而通过Multi-CSI辅助的动态功率控制方案能提供16.63%基站节能，并且UPT损失不到1%。由评估结果可以看出，基于Multi-CSI测量上报增强的动态功率调整方案可以在提供较大节能增益的同时带来较少的UPT损失。

### 3.5 基站节能唤醒机制

基站节能技术可以通过减少信号传输、增加基站睡眠时间等方案有效地减少基站功耗，获取基站节能增益。然而，当没有终端反馈参与时，仅依靠基站获取的信息进行基站节能，可能会对系统性能（如吞吐量、延迟等）和用户体验产生影响，最终导致基站节能方案不能被应用。为了在实现基站节能的基础上尽可能保证系统性能和用户体验，可以考虑引入终端辅助的基站节能方案，即基站唤醒信号。

基站唤醒信号是UE向基站发送的上行信号，用以辅助基站在节能状态和非节能状态之间的切换，例如：指示节能小区唤醒，触发下行公共信号传输、指示信号传输周期更新等。基站唤醒信号可以及时地将终端的需求反馈给基站，以确保基站快速响应/服务，并将对用户体验的影响降至最低。

## 4 结束语

基站多域节能技术的研究，旨在保证网络性能的基础上，减少网络能耗，提高信息通信行业绿色低碳发展质量，促进信息通信行业绿色转型。

1) 建立网络能耗模型。网络能耗模型通过建模基站不同的数据发送、接收状态，不同深度的睡眠模式及模式转换期间的相对功耗，来评估各种网络节能技术。

2) 制定评估方法。在对候选节能技术的评估中，除了考虑网络节能增益，还要兼顾网络关键绩效指标（KPI）、终端能耗，以及终端复杂度等。评估方法和能耗模型是选择能够平衡节能效果和网络性能影响的节能技术的方法。

3) 研究并识别出网络节能的关键技术点。从包括时域、空域、频域和功率域等各个方面使能更动态、高效、精准的无线传输，并通过终端和基站间信息交互的方式提升节能效果。每种节能技术均可匹配网络特征灵活设置触发条件，多种节能技术之间即可单独应用也可叠加部署。面向5G-A，通过对无线网络节能技术进行优化和升级以更好地满足绿色网络的发展和运营需求。

### 参考文献

- [1] 工业和信息化部, 国家发展改革委, 财政部, 等. 信息通信行业绿色低碳发展行动计划(2022—2025年) [R]. 2022
- [2] 中兴通讯. 4G/5G网络节能降耗技术白皮书 [R]. 2020
- [3] 中兴通讯. B5G技术白皮书 [R]. 2020
- [4] 3GPP. Technical specification group radio access network; study on network energy savings for NR: TR 38.864 [S]. 2022
- [5] 中国移动研究院. 5G基站节能技术白皮书 [R]. 2020
- [6] 3GPP. Technical specification group radio access network; medium access control (MAC) protocol specification: TR 38.321 [S]. 2019
- [7] 3GPP. Technical specification group radio access network; physical layer procedures for data: TR 38.864 [S]. 2019

### 作者简介



郭诚，中兴通讯股份有限公司无线网络节能首席规划专家；长期从事无线通信系统的研发工作，主要负责无线网络节能和多频协同技术规划；所负责的中兴通讯绿色节能方案多次获得全球性行业奖项；拥有发明专利10余项。



陈梦竹，中兴通讯股份有限公司3GPP RAN1标准专家；长期从事无线通信系统的物理层标准工作，主要研究方向包括信道编码、干扰管理、终端节能、基站节能以及元宇宙等；发表论文5篇，拥有发明专利50余项。

# 面向节能减排的 跨制式融合感知通信系统



## Energy Conservation and Emission Reduction Oriented Perception and Communication Fused Wi-Fi and Cellular Integrated Communication System

周均翼/ZHOU Junyi, 周琳/ZHOU Lin,

张舜卿/ZHANG Shunqing

(上海大学, 中国 上海 200444)

(Shanghai University, Shanghai 200444, China)

DOI: 10.12142/ZTETJ.202306004

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231214.1420.002.html>

网络出版日期: 2023-12-14

收稿日期: 2023-10-15

**摘要:** 提出了一套面向节能减排的跨制式融合感知通信系统, 该系统包括通信融合、感知融合以及边缘协同管控平台3部分。经过相关实验的验证, 该系统能够对各类节能算法进行快速的脚本化原型测试, 有望对未来面向节能减排的跨制式网络设计提供可靠验证与指导。

**关键词:** 绿色网络; 跨制式网络; 原型验证平台; 协同管控

**Abstract:** A perception and communication fused Wi-Fi and cellular integrated communication system is proposed, including fused communication, fused perception, and edge collaborative control platform. With the proposed fusion perception communication system, rapid script-based prototyping tests can be conducted on various energy-saving algorithms, providing reliable validation and effective guidance for future energy-conserving cross-disciplinary network designs.

**Keywords:** green network; Wi-Fi and cellular integrated communication system; prototype verification platform; collaborative management and control

**引用格式:** 周均翼, 周琳, 张舜卿. 面向节能减排的跨制式融合感知通信系统 [J]. 中兴通讯技术, 2023, 29(6): 16–22. DOI: 10.12142/ZTETJ.202306004

**Citation:** ZHOU J Y, ZHOU L, ZHANG S Q. Energy conservation and emission reduction oriented perception and communication fused Wi-Fi and cellular integrated communication system [J]. ZTE technology journal, 2023, 29(6): 16–22. DOI: 10.12142/ZTETJ.202306004

近10年, 中国的碳排放总量年均增长2.3%, 超过世界平均水平<sup>[1]</sup>。截至2019年, 信息与通信技术产业已经构成了全球5%的碳足迹。而到2020年, 该产业75%的碳排放均涉及无线通信行业, 全网碳排放量在持续增长<sup>[2]</sup>。以上海的实际布网情况为例, 固网以及4G/5G的实际接入数大概是网络实际接入数的2倍。多制式的独立组网调度将导致巨大的资源浪费和重复覆盖, 从而使全网碳排放比理想情况增加20%左右<sup>[3]</sup>。为了实现“双碳”目标, 可以将固网和移动网络进行统一管控, 实现跨制式资源的灵活调度与机会式休眠, 从根本上解决资源浪费问题。然而, 目前跨制式网络的

协同管控存在无法通过对协议栈的简单适配来达到跨制式资源的统一管控、无法实现对用户业务的全场景感知等问题。现有感知技术如上行到达时间差定位法(UTDOA)<sup>[4]</sup>、到达角度定位法(AOA)<sup>[5]</sup>等都极度依赖接入点的主动参与, 在站点休眠时无法进行。由于感知信息由相应的接入节点掌握, 因此无法进行跨制式的感知信息共享。现有的5G蜂窝网络、4G蜂窝网络以及Wi-Fi网络的灵活调度存在以下挑战:

1) 资源管理: 不同网络之间的资源分配和管理是一个复杂的问题。蜂窝网络和Wi-Fi网络具有不同的特性, 例如: 传输速率、覆盖范围和信号干扰等。调度算法需要考虑这些因素, 并在多个网络之间有效地分配资源, 以确保网络性能和用户体验。

2) 时延和吞吐量: 4G、5G和Wi-Fi网络的时延和吞吐量要求不同。5G网络通常具有较低的时延和较高的吞吐量,

**基金项目:** 国家重点研发计划(2022YFB2902304、2022YFB2902005、2022YFB2902022); 国家自然科学基金项目(62071284、61904101); 广东省重点领域研发计划(2020B0101130012); 上海市科学技术委员会创新计划(21ZR1422400)

适用于实时应用和大规模数据传输。4G 和 Wi-Fi 网络可能在某些方面性能较低，需权衡时延和吞吐量的需求进行调度。

3) 扩展性和互操作性：灵活调度算法需要适应不断增长的网络规模和日益复杂的网络架构。对于多个移动设备和基站之间的调度决策，需要具备良好的扩展性和互操作性，以便实现跨不同网络的无缝切换和协同工作。

针对上述挑战，本文中我们提出了一个面向节能减排的感知融合智能通信系统。该系统首先实现对蜂窝网络以及 Wi-Fi 网络的软件灵活调度，构建了基于互联网协议（IP）层分流的双连接系统。然后，借助毫米波雷达模块以及无线指纹拓扑，构建了主动感知与被动感知协同的全场景解决方案。最后，利用时序数据库和并发通信接口，构建了协同控制平台。这一系统能够实现跨制式的协同传输、通过被动和主动感知的融合，实现全场景的跨制式感知，减少通信资源浪费的问题。

## 1 相关技术介绍

### 1.1 通信网络

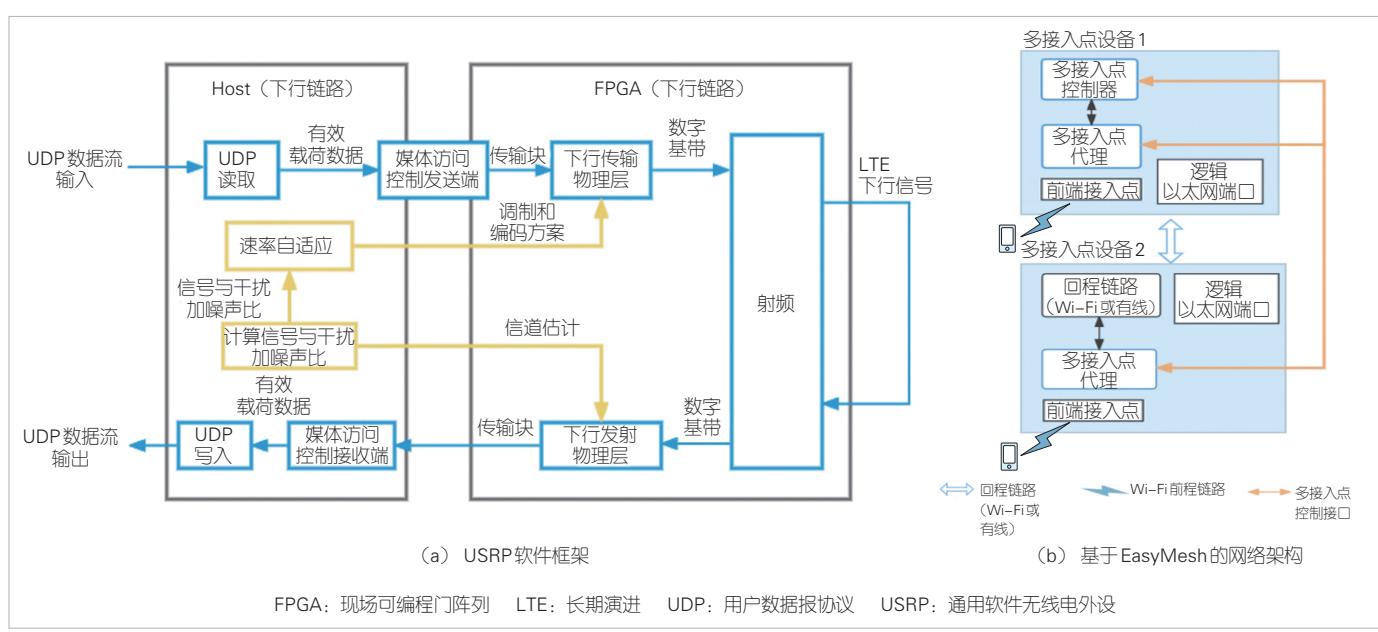
通信网络分为固定通信网络、移动通信网络等。固定通信网络的主要布网方式为光纤+Wi-Fi 网络，为室内场景进行大带宽的稳定网络覆盖；而移动通信网络不受光纤布网的限制，为室内室外等多种场景下的用户提供多种服务。然而，固网和移动网不同的部署特点会带来网络重复覆盖和资源浪费，从而导致全网碳排放增加。

蜂窝网络是一种移动通信硬件架构，主要由移动站、网络子系统、基站子系统组成。为实现对蜂窝网络的调控，本文基于通用软件无线电外设（USRP）定义了蜂窝网络。USRP则是一种开源的、低成本的无线电硬件平台，软件框架<sup>[6]</sup>如图 1 (a) 所示。HOST 部分主要涉及收发的配置，在上位机中实现。而物理层主要在现场可编程门阵列（FPGA）中实现。

对于固定通信网络，本文通过 OpenWrt 系统实现软件定义的自组织 Wi-Fi 网络。自组织 Wi-Fi 网络是一种无中心、多跳的网络，能够应用在没有基站覆盖的偏远地区。OpenWrt 是一个用于嵌入式设备的 Linux 系统。本文提出的通信系统在 OpenWrt 系统的基础上安装了 Wi-Fi 组网协议（EasyMesh）<sup>[7]</sup>，如图 1 (b) 所示。EasyMesh 网络包含 2 个逻辑实体：1 个多接入点控制器（MAPC）和 1 个或多个多接入点代理（MAPA）。MAPC 从 MAPA 接收前端接入点、客户端和回程链路的测量和能力数据，并在 MAPA 上执行接入点控制相关的命令和操作。MAPA 执行从 MAPC 收到的命令，并与多接入点设备上的前端接入点和后向回传终端（Back-haul STA）的 Wi-Fi 子系统对接，以获得测量和能力数据，应用配置变化和执行接入点控制功能。

### 1.2 跨制式双连接传输系统

跨制式指不同的网络制式间的切换与交互。双连接是第 3 代合作伙伴计划（3GPP）Release-12<sup>[8]</sup>版本引入的重要技术。通过双连接技术，长期演进（LTE）宏站和小站可以利用现有的非理想回传 X2 接口来实现载波聚合，从而为用户



▲图1 USRP软件框架与EasyMesh网络架构<sup>[6-7]</sup>

提供更高的速率，并可以利用宏/微组网提高频谱效率，负载平衡<sup>[9]</sup>。支持双连接的终端可以同时连接两个基站，增加单用户的吞吐量。

跨制式双连接传输机制的系统由1个边缘服务器进行协同管控；数据传输部分的接入点由1个USRP、1个自组织Wi-Fi固网路由器组成，终端为跨制式终端。通信流程如下：用户向边缘服务器请求相应的数据业务，边缘服务器做出应答，并通过跨制式机制建立传输。具体实现方式为：将传输层的数据包切割并排序，分发给与用户连接的接入点；接入点按照各自的协议栈进行传输；用户收到来自各个制式的数据包后，将数据包根据原来的顺序进行重新排序，并恢复成原来的数据流；同时用户利用上行链路进行相应协议栈缓存积压状态上报。

## 2 跨制式融合感知通信系统架构设计

本文所提出的面向节能减排的跨制式融合感知通信系统由通信融合、感知融合网络、边缘管控平台3部分组成。通信融合和感知融合网络部分可以实现对用户的分层感知以及对跨制式节点的节能参数的软件化编排。边缘管控平台实现统筹调度，支持各类场景下的节能设计。该通信系统架构设计及实物图如图2所示。

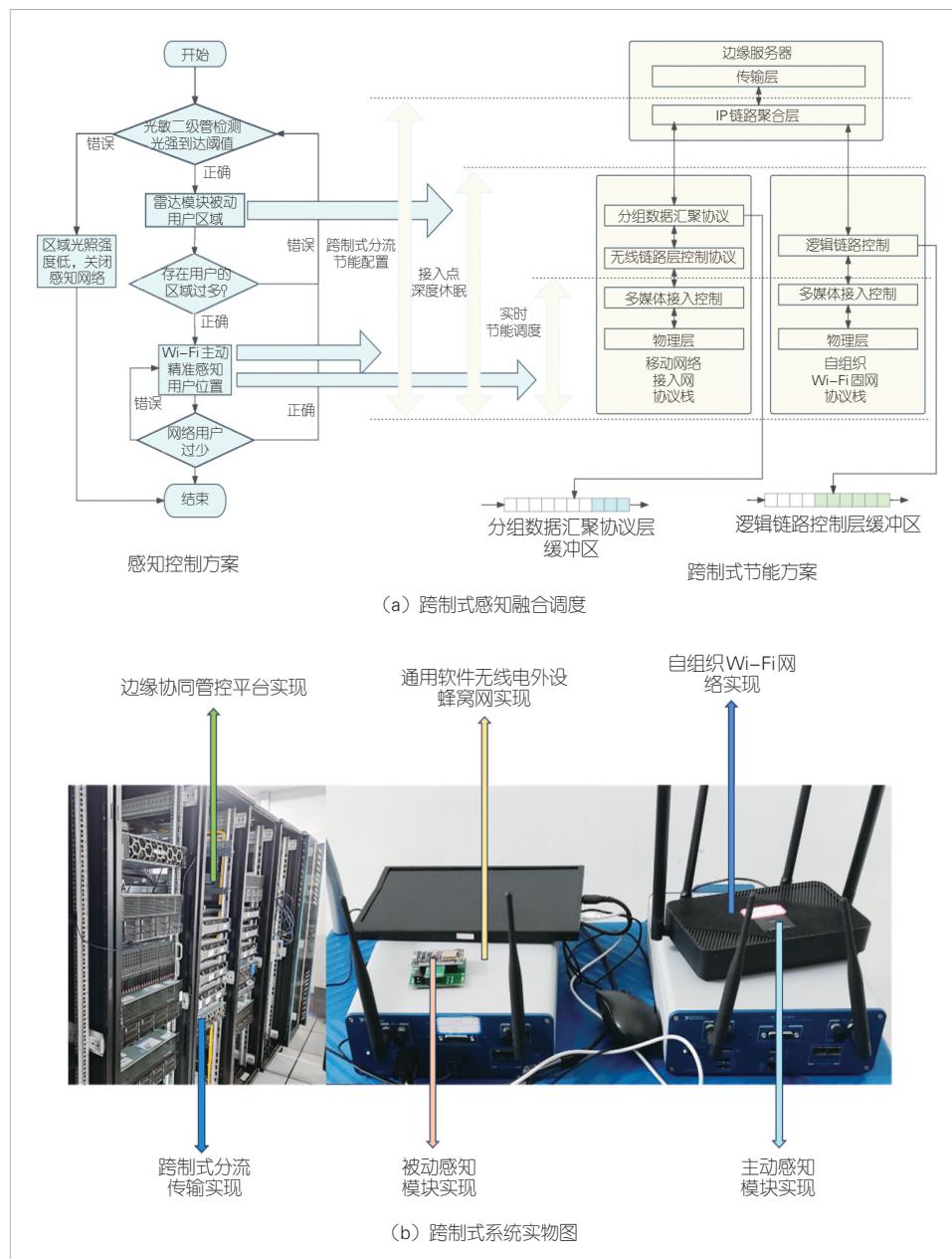
根据上述介绍，用户利用上行链路进行相应协议栈缓存积压状态上报，根据积压上报情况，对数据进行分流，以达到节能减排的效果。然而，跨制式系统的缓存积压不光来自用户需求与接入点通信能力的不匹配，还来自数据包乱序等，因此我们提出了相应的系统模型来建模上述问题的场景，并提出相应的算法。

### 2.1 通信融合

为实现对跨制式资源的统一管控，该通信系统首先通过USRP实

现对4G/5G系统的软件调度，再通过OpenWrt开源系统实现软件定义的自组织Wi-Fi网络，最后通过跨制式IP层分流，实现跨制式的协同传输。

通过对USRP软件的下行HOST部分、下行FPGA部分进行修改，该通信系统支持节能关键参数的软件调度。为了认识基站的能耗构成以便后续的节能应用，我们引入了功耗模型<sup>[10]</sup>，并结合实际系统关键参数，进行软件开发，以实现无线能耗统计模块的设计。基于USRP定义的蜂窝网络能够提供丰富的性能统计和参数调用，如调制编码策略（MCS）、资源块（RB）分配策略等。同时，部分参数为可控参数，



▲图2 通信系统架构设计及实物图

利用Labview图像化编程语言并使用传输控制协议(TCP)的通信方式,可以近实时地从边缘智能服务器获取系统调度参数并将其作用在USRP系统中,从而实现灵活的边缘智能管控。

此外,该通信系统在OpenWrt系统的基础上安装了Wi-Fi组网协议(EasyMesh)<sup>[7]</sup>,该组网协议实现了网关接入/休眠拓扑更新、路由智能寻优等功能。

为了使能节能减排的应用,我们为固网通信系统提供了多层休眠支持,包括基于智能物联网开关的接入点休眠和天线关断休眠。由于自组织固网的路由器接入点不能提供基带单元等的深度关断,因此考虑在低负载情况下,直接对自组织固网的接入点进行整站关断,以实现网络节能减排。该通信系统的智能物联网开关基于ESP8266芯片<sup>[11]</sup>开发板硬件,通过Arduino平台进行二次开发。这样可以通过自组织网络主路由器网关连接边缘服务器,从而获取开关控制指令并进行上电、断电的操作。然后在初次上电时,连接ESP8266并将其作为路由器,然后发射Wi-Fi信号,在进入其系统的网络后台后,再设置与其通信的主路由器。经过以上配置,智能物联网开关实现通过路由器和边缘服务器上运行的TCP Server程序通信。该系统每5 s向边缘智能服务器发起一次TCP通信请求,然后接收开闭信号,以减少智能开关的通信能耗。

在OpenWrt系统中,通过对相关文件进行配置,可以控制天线的开关。

为了优化网络资源和网络负载管理,我们构建了基于IP层分流的双连接系统。该系统通过对不同制式的智能感知,实时调整不同制式的流量分配比例,使用户在一定网络功耗下获得尽可能高的服务质量。同时,不同制式的功率也在系统的监控和调节范围内,因此可以使能跨制式节能分流调度。首先,边缘服务器通过USRP和Wi-Fi传输的缓存状态差 $\Delta Q$ ,获取USRP和Wi-Fi当前工作功率 $\Delta P$ ;然后根据 $\Delta Q$ 和 $\Delta P$ 进行USRP和Wi-Fi的分流比例决策和工作功率决策;再根据决策结果调整二者的功率并在IP链路聚合层将流量分配给USRP和Wi-Fi。

针对具有等停机制的TCP传输协议,分流算法采用的是Fuzzy-PID算法<sup>[12]</sup>。经验证,在上层协议具有等停机制时,最小化缓存差等价于最大化吞吐量。因此,该算法通过建立基于模糊逻辑的PID动态调整机制,来最大化跨制式传输的缓存差。此外,该算法还对用于分流决策的PID参数进行动态调整,能够在未知非线性、处理延迟和干扰的情况下达到预期的性能。

## 2.2 感知融合

为了实现全场景的智能感知系统,解决传统感知系统依赖于对应的接入点参与,无法应对站点休眠以及跨制式协同场景的问题,我们构建了主动感知与被动感知协同的全场景解决方案。被动感知系统在站点处于休眠等场景下,即时感知可能业务的接入;主动感知则是在中高负载时,感知用户的精准位置,从而支持更加精细的节能调度。

被动感知网络采用毫米波雷达模块,相较于市面上常见的车载雷达模块或者人体感应雷达模块,具有低功耗、易部署、高可信度的优势,感知范围为0~20 m。经实验验证,该被动感知系统能够在多种场景下实现高可信的粗粒度人员感知。

主动感知系统采用了基于图卷积网络(GCN)的室内感知定位算法<sup>[13]</sup>。该主动感知方案分为两步:首先感知定位系统,然后感知全场景数据,并将数据经过自组织Wi-Fi网络上传至服务器,服务器运行该GCN算法,最终得出人员的精确位置。

上述两个感知系统协同运作,被动感知应对通信系统休眠等低负载业务场景,而主动感知主要处理业务密集的高负载场景,为跨制式通信系统提供全场景的业务数据支撑。

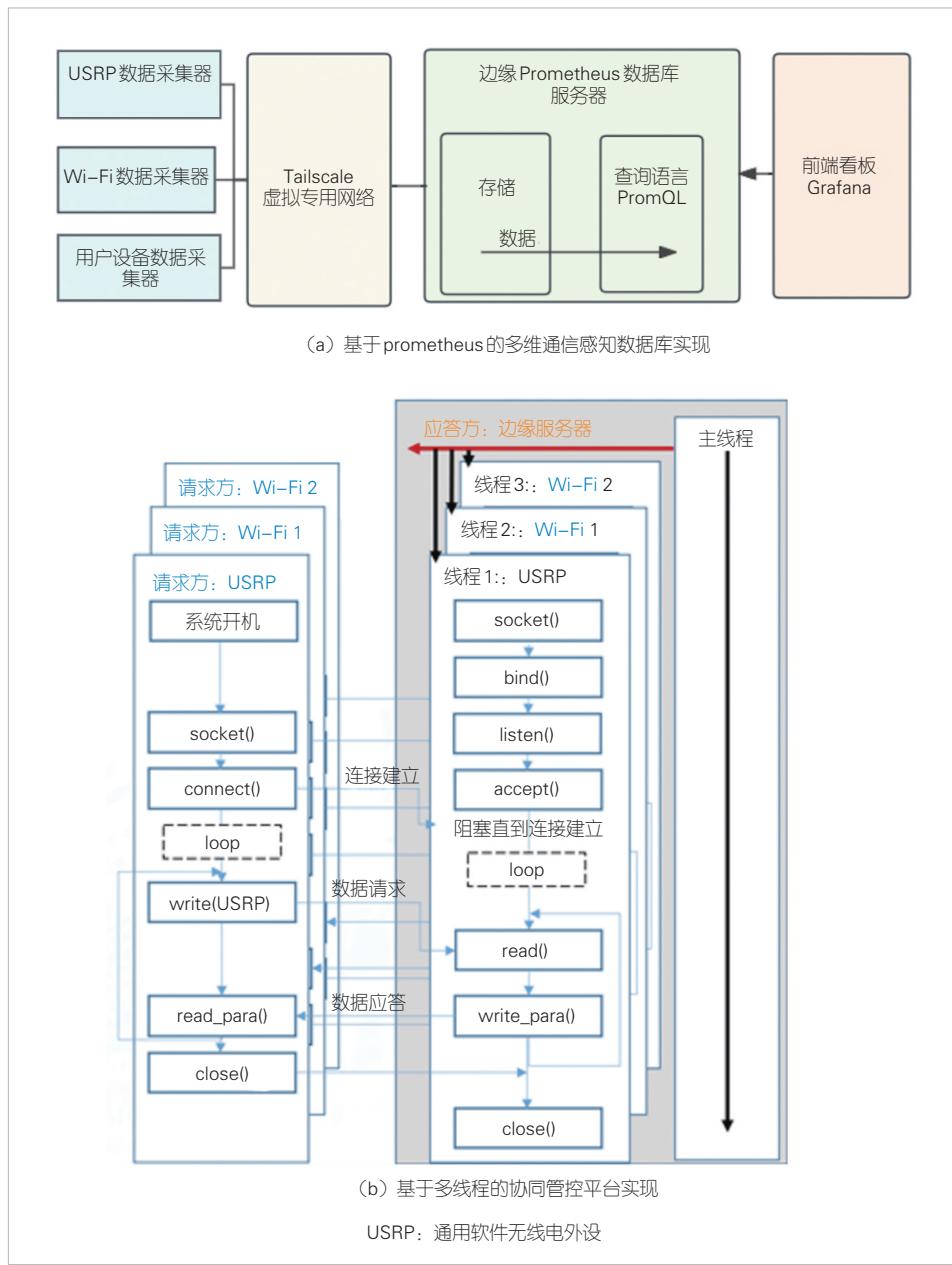
## 2.3 边缘管控平台实现

为了支持各类场景下的节能设计,我们首先基于Prometheus建立了多维通信感知数据库,并在此基础上设计了基于边缘服务器的协同管控平台。如图3(a)所示,该多维通信数据库主要包括以下几个部分:数据采集器Exporter、用于安全上报数据的通道Tailscale虚拟专用网络(VPN)、边缘Prometheus数据库服务器,以及前端看板Grafana。

数据采集器Exporter通过文件读取、页面抓取等方式获得USRP、OpenWrt等各个操作平台上的数据并进行存储。Tailscale VPN建立了一套虚拟的专用网络将所有的接入点以及终端的数据进行采集和上报。Grafana则对USRP、Wi-Fi、终端上采集到的性能与能耗数据等进行实时显示,方便网络运维人员进行实时监控以及问题排查。

协同控制平台通过可靠的通信方式实现多个跨制式接入点的并发控制,支持全栈式跨制式网络参数编排,以及各类节能算法的脚本化快速验证。

该系统通过使用基于websocket的多线程并发通信流程,实现并发的接入点网络参数获取,如图3(b)所示。边缘服务器的网络参数控制器主线程主要用来在阻塞循环中监听接入点的连接请求,在建立连接后,将该连接交给一个并发子线程进行数据请求的接收和参数的应答。依靠上述机制,



▲图3 边缘平台实现

边缘服务器可以在多个接入点请求参数时并发地根据需求应答网络参数，实现多接入点的协同管控。

此外，该系统提供了两个不同接口来设计协同控制平台，分别是节能参数调度接口和参数接口。该系统针对整个网络的各个协议栈提供了节能参数调度接口，实现了跨制式全栈式的网络参数编排；该系统提供的IP层的跨制式分流比例的参数接口，在网络部署时可以通过引入自动化分流策略脚本，根据有边缘多维通信感知数据库中信息动态调整分流比例，实现跨制式业务在各接入点间的按需分配。系统通过上述的跨制式节点间并发参数传递机制，将USRP节点的

射频关断以及自组织Wi-Fi路由器的物联网智能开关关断指令发送给响应的TCP通信链路，在读取被动感知参数以后，可以部署自动化脚本并通过简单的逻辑实现接入点深度关断，从而实现接入点的深度休眠。向接入点提供MAC层、物理层的节能调度参数，如天线关断、MCS、射频发送功率等，结合跨制式分流机制，实现接入点的实时节能调度。这一精细化的无线参数调度能够在满足用户体验的同时，达到节能的效果。

### 3 系统测试与分析

#### 3.1 接入点深度休眠

为验证本文所提的接入点深度休眠带来的节能效果，我们搭建了表1所示的实验场景，将所提方案与普通节能调度策略进行了对比。

为了保证对比实验的客观性，我们在1 h内让所有人员离开测试场景12 min（20%的时间），然后再对比所提方案与普通节能调度策略的节能效果。我们发现在普通节能策略下，Wi-Fi接入点仍存在约为5.3 W的静态功耗；而所提的接入点深度休眠策略通过物联网开关将接入点整站关断，能耗降为0，直到有人进入场所，数据恢复发送。

我们统计1 h内系统的总能耗：所提方案消耗了17.13 J

▼表1 能耗数据

参数	接入点深度休眠	跨制式节能调度	
网络配置	Wi-Fi网络	Wi-Fi网络	蜂窝网络
接入点个数	1	2	1
天线个数	2	2	1
静态能耗/ mW	5 300	5 300	11 000
发射功率/ dBm	0~30	0~30	-10~30
带宽/MHz	20	20	18

的能量，普通节能调度方案消耗了 19.91 J 的能量。比起普通节能调度方案，本文所提方案降低能耗 13.96%。

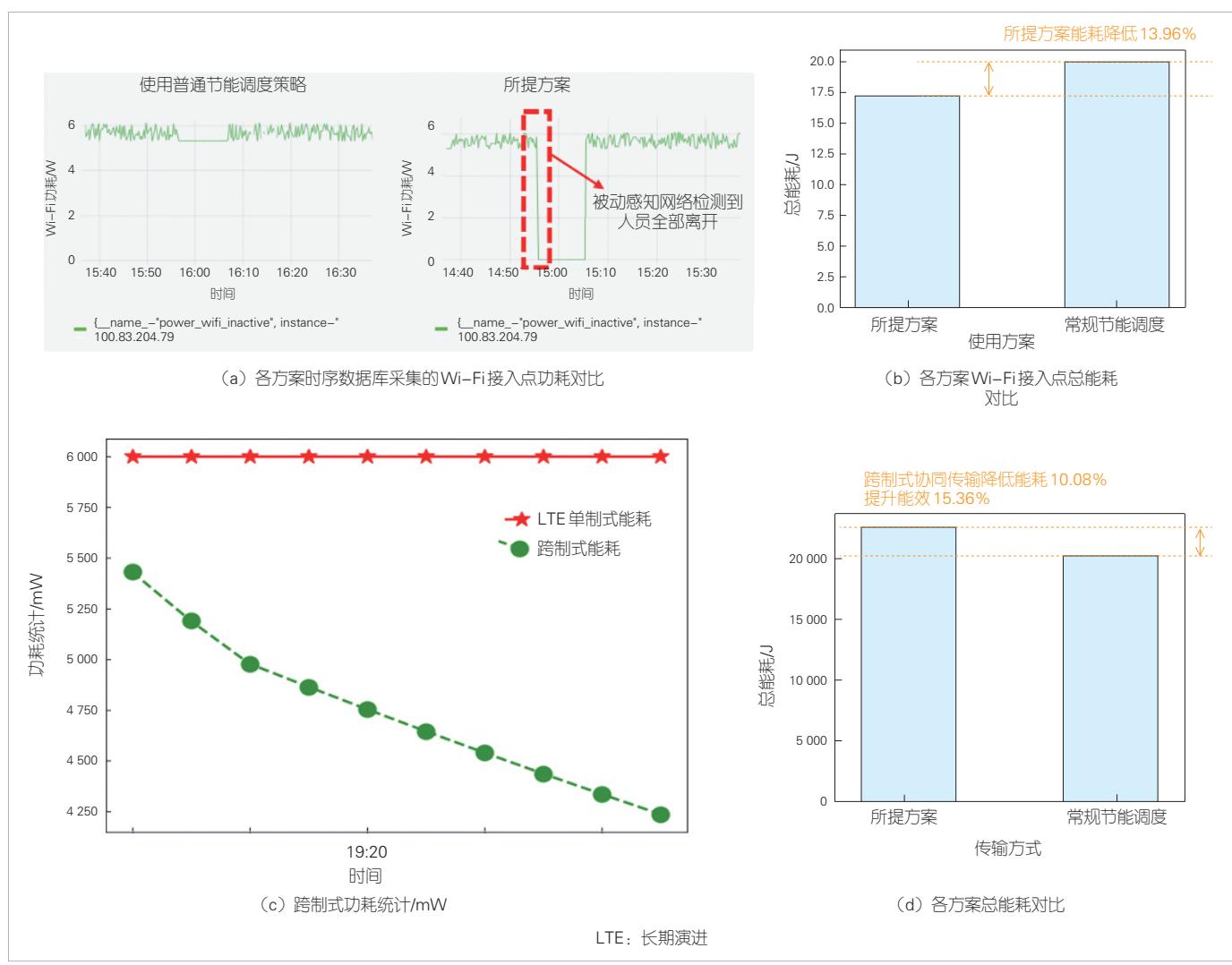
### 3.2 跨制式节能调度节能效果

为了验证所提的跨制式节能调度方案带来的节能效果，我们搭建表 1 所示的实验场景，将所提方案与单制式传输进行了对比实验。

为了保证对比实验的客观性，我们设置相同的用户吞吐量要求：两种传输方式的用户吞吐量为 26 Mbit/s。我们发现，在跨制式传输场景下，用户在某一制式信道发生变化后，跨制式分流机制迅速响应，并向另一制式分配更多流量，保证业务传输要求得到满足。此时能耗也产生了相应的波动。在自组织 Wi-Fi 网络自动切换到更好的接入点时，系统选择了更加节能的分流比，让信道条件和能效更高的制式

传输更多信息。此时，系统的能耗明显下降。

如图 4 所示，图 4 (a) 对比了所提方案与普通节能调度策略的节能效果。可以看到在普通节能策略下，用户离开后，服务 Wi-Fi 接入点仍存在约为 5.3 W 的静态能耗。所提的接入点深度休眠策略通过物联网开关将接入点整站关断，能耗降为 0。直到有人进入场所，数据恢复发送。图 4 (b) 通过统计 1 h 内系统的总能耗，发现我们所提的接入点深度休眠方案消耗了 17.13 J 的能量，普通节能调度方案消耗了 19.91 J 的能量。比起普通节能调度方案，所提方案降低能耗 13.96%。图 4 (c) 比较了在一特定时间段，跨制式系统与单制式系统的能耗，跨制式能耗相较于单制式能耗有明显下降。图 4 (d) 统计了 10 min 内系统的总能耗，我们发现所提方案消耗了 20 271.29 J 的能量，单制式传输方案消耗了 22 544.74 J 的能量。比起常规的节能调度方案，所提方案降



▲图 4 各方案能耗对比

低能耗 10.08%，能效提高了 15.36%。总的来说，本文所提出的接入点休眠与跨制式系统的总体能耗相较于传统的单制式系统有明显下降。

#### 4 结语

随着全网能耗的日益增长，大量部署的现有跨制式节点却没有得到充分利用，这势必会带来资源的浪费。针对这一矛盾，本文首先提出了面向节能减排的跨制式融合感知通信系统。然后，借助软件定义无线电 USRP 平台以及开源操作系统 OpenWrt，实现了对蜂窝网络以及 Wi-Fi 网络的软件灵活调度，同时借助毫米波雷达模块以及无线指纹拓扑，实现被动感知和主动感知方式的共存。最后，利用时序数据库和并发通信接口，实现资源的集中管控。通过系统实测数据，我们发现所提的休眠策略与跨制式调度策略分别能够带来约 14% 和 10% 的总能耗降低。所提跨制式系统能够对各类节能算法进行快速的脚本化原型测试，有望对未来面向节能减排的跨制式网络设计提供可靠验证和有效指导。

#### 参考文献

- [1] 张亚飞. 信息与通信技术在碳达峰、碳中和战略中的应用 [J]. 电子世界, 2021 (22): 146–147
- [2] GreenIT. fr Frédéric Bordage. The environmental footprint of the digital world [EB/OL]. [2023-10-12]. <https://thecalm.ca/wp-content/uploads/2020/02/The-Environmental-Footprint-of-the-Digital-World-Study.pdf>
- [3] 上海市通信管理局. 上海通信行业发展概况(2021年) [EB/OL]. (2022-02-17) [2023-10-12]. [https://shca.miit.gov.cn/zwgk/xyfz/xyjt/art/2023/art\\_828e4150f785419b9e78b970ae45e2dc.html](https://shca.miit.gov.cn/zwgk/xyfz/xyjt/art/2023/art_828e4150f785419b9e78b970ae45e2dc.html)
- [4] MCDERMOTT K, VAGHEFI R M, BUEHRER R M. Cooperative UTDOA positioning in LTE cellular systems [C]//Proceedings of 2015 IEEE Globecom Workshops (GC Wkshps). IEEE, 2015: 1–6. DOI: 10.1109/glocomw.2015.7414171
- [5] XHAFA A, DEL PERAL-ROSAZO J A, LÓPEZ-SALCEDO J A, et al. Evaluation of 5G positioning performance based on UTDoA, AoA and base-station selective exclusion [J]. Sensors, 2021, 22(1): 101. DOI: 10.3390/s22010101
- [6] EMERSON. LabVIEW communications system design suite release notes [EB/OL]. (2023-06-14) [2023-10-12]. <https://www.ni.com/zh-cn/support/documentation/release-notes/product.labview-communications-system-design-suite.html>
- [7] Wi-Fi Alliance. Wi-Fi EasyMesh [EB/OL]. (2023-06-14) [2023-10-12]. <https://www.wi-fi.org/zh-hans/discover-wi-fi/wi-fi-easymesh>
- [8] 3GPP. Evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN) V13.4.0: TS 36.300 [S]. 2016
- [9] ROSA C, PEDERSEN K, WANG H, et al. Dual connectivity for LTE small cell evolution: functionality and performance aspects [J]. IEEE communications magazine, 2016, 54(6): 137–143. DOI: 10.1109/mcom.2016.7498101
- [10] XU Z K, YANG C Y, LI G Y, et al. Energy-efficient configuration of spatial and frequency resources in MIMO-OFDMA systems [J]. IEEE transactions on communications, 2013, 61(2): 564–575. DOI: 10.1109/tcomm.2012.100512.110760
- [11] ESPRESSIF. ESP8266EX 技术规格书 [EB/OL]. (2023-06-14) [2023-10-12]. [https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex\\_datasheet\\_cn.pdf](https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_cn.pdf)
- [12] YU J, ZHANG S Q, SUN J Y, et al. High-performance transmission mechanism design of multi-stream carrier aggregation for 5G non-standalone network [EB/OL]. (2021-04-05) [2023-10-12]. <http://www.cic-chin.communications.cn/EN/10.23919/JCC.fa.2021-0405.202308>
- [13] SUN Y Z, XIE Q G, PAN G J, et al. A novel GCN based indoor localization system with multiple access points [C]//Proceedings of 2021 International Wireless Communications and Mobile Computing (IWCMC). IEEE, 2021: 9–14. DOI: 10.1109/iwcmc51323.2021.9498616

#### 作者简介



周均翼，上海大学通信与信息工程学院在读硕士研究生；主要研究方向为多模态融合感知定位等。



周琳，上海大学通信与信息工程学院在读本科生；主要研究方向为宽带移动通信技术。



张舜卿，上海大学通信与信息工程学院教授；主要研究方向包括节能 5G/5G+ 通信网络、混合计算平台以及联合射频和基带设计；入选“青年千人计划”，并于 2017 年获得 IEEE 通信学会通信进展论文奖；发表论文 60 余篇，获得授权专利 50 余项。

# 基于多时隙业务联合整形的低能耗资源调度方法



## A Multi-Slot Traffic Shaping Based Resource Allocation Method for Low Energy Consumption

李建东/LI Jiandong, 牛淳隆/NIU Chunlong,  
赵晨曦/ZHAO Chenxi, 刘俊宇/LIU Junyu  
(西安电子科技大学, 中国 西安 710071)  
(Xidian University, Xi'an 710071, China)

DOI: 10.12142/ZTETJ.202306005  
网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231211.1535.008.html>  
网络出版日期: 2023-12-12  
收稿日期: 2023-10-15

**摘要:** 面向未来6G移动通信系统超高数据密度的业务需求场景,为保障用户服务质量(QoS)并降低系统能耗,首先分析了移动通信系统的能耗构成,发现了系统能耗的非线性特征。然后在此基础上,设计了多时隙业务联合整形的低能耗资源调度方法。该方法通过感知用户业务流量和时延要求等需求侧的数据特征,利用深度强化学习算法在给定的多个时隙内动态调整基站资源分配策略。该资源分配策略降低了用户业务请求的非平稳性,从而减少了基站的非线性传输特性产生的额外能耗。最后通过软件仿真对比不同方法,验证了基于多时隙业务联合整形的理论和算法的正确性及有效性。

**关键词:** 系统能耗的非线性特征; 多时隙业务联合整形; 低能耗资源调度; 深度强化学习

**Abstract:** Facing the business requirements scenario of ultra-high data density of future 6G mobile communication systems, to ensure user quality of service (QoS) and reduce system energy consumption, the energy consumption composition of mobile communication systems is analyzed and the nonlinear characteristics of system energy consumption is studied. On this basis, a low-energy resource allocation method for multi-slot traffic joint shaping is designed. By sensing the data characteristics of the demand side such as user service traffic and delay requirements, this method uses the deep reinforcement learning algorithm to dynamically adjust the base station resource allocation strategy within a given multiple time slot. The resource allocation strategy reduces the non-stationarity of user service requests, thereby reducing the additional energy consumption caused by the nonlinear transmission characteristics of the base station. Finally, by comparing different methods through software simulation, the correctness and effectiveness of the theory and algorithm based on multi-time slot business joint shaping are verified.

**Keywords:** nonlinear characteristics of system energy consumption; multi-slot traffic joint shaping; low energy resource allocation; deep reinforcement learning

**引用格式:** 李建东, 牛淳隆, 赵晨曦, 等. 基于多时隙业务联合整形的低能耗资源调度方法 [J]. 中兴通讯技术, 2023, 29(6): 23–28. DOI: 10.12142/ZTETJ.202306005

**Citation:** LI J D, NIU L C, ZHAO C X, et al. A multi-slot traffic shaping based resource allocation method for low energy consumption [J]. ZTE technology journal, 2023, 29(6): 23–28. DOI: 10.12142/ZTETJ.202306005

随着5G移动通信的全面商用,未来的6G移动通信系统的网络性能<sup>[1]</sup>将在原有的业务场景下进一步提升。6G移动通信将组成一个巨大的分布式神经网络,实现通信、感知、计算、存储一体化<sup>[2]</sup>,深度融合物理世界、生物世界和数字世界,真正实现“万物互联”。同时,随着人工智能(AI)技术的进一步发展,6G通信网络将会更加智能化<sup>[3]</sup>。基于数据驱动的算法赋能网络动态的自我学习能力和调节能

力,充分挖掘用户的潜在需求并制定个性化的服务。此外,超密集异构网络(UDN)<sup>[4]</sup>在未来6G移动通信系统的发展中依然占据主导地位,6G基站的密度进一步提高,用户的业务数据请求维度进一步增大,数据服务的个性化要求进一步提升。上述的发展需求对未来的移动通信网络提出了更高的要求,如更加密集的基站数目、更多的天线阵列和更多维度的资源管控问题<sup>[5]</sup>。

截至2022年底,全国移动通信基站总数达1083万个。基站越多,对应的频谱资源管控问题越复杂,同时如果资源

基金项目:国家重点研发计划项目(2022YFB2902302)

管控方案不合理，则会产生巨大的能量损耗。根据估算，三大运营商一年需要支付的电费已超500亿元。针对未来超密集基站和超高用户数据等现象导致的能耗过高问题，未来6G无线移动通信的宏观驱动力之一便是实现环境可持续发展。2020年9月，中国宣布力争二氧化碳排放于2030年前达到峰值，2060年努力实现碳中和<sup>[6]</sup>。“双碳”战略对6G移动通信系统的网络性能提出了更高的要求。为保障用户服务质量，实现能量的极致利用，平衡能源的供需关系，降低异质网络架构的系统碳排放，我们需要研究智能化的组网传输方法。

## 1 基站能耗分析

基站是移动通信系统最重要的组成部分之一，它实现了大规模的无线覆盖，保障了移动用户的业务需求被及时满足。一个完整的基站包含的设备数量较多，根据耗电量调查，传输设备、基站主设备、动力配套设备、环境温度控制设备、监控设备、照明设备等的耗电量都相对较高，其中基站主设备、环境温度控制设备的电能消耗量大。功率放大器是射频通信系统中极为重要的器件，射频放大器将输入到输入端的信号进行放大并输出。作为通信电路中的主要设备，基站的功率放大器所消耗的功率无疑是低能耗组网的重点研究内容<sup>[7]</sup>。当前5G基站的主要组成部分有：集中单元(CU)、分布单元(DU)和有源天线单元(AAU)，其中各个组成部分和对应的功耗来源占比分析如图1所示。

基站的AAU是主要的功耗占比模块，而功率放大器是AAU的主要能量消耗模块。此外，当基站工作在中高负载区域时，功率放大器的能量损耗是占比最大的，接近60%。因此，着重研究功率放大器在中高负载区域工作时的性能特点并设计相应的算法，对降低基站功耗有着重要的意义。

### 1.1 传输能耗的非线性特征

建模通信网络场景时，计算每个用户和基站之间的信道容量可以采用香农公式<sup>[8]</sup>：

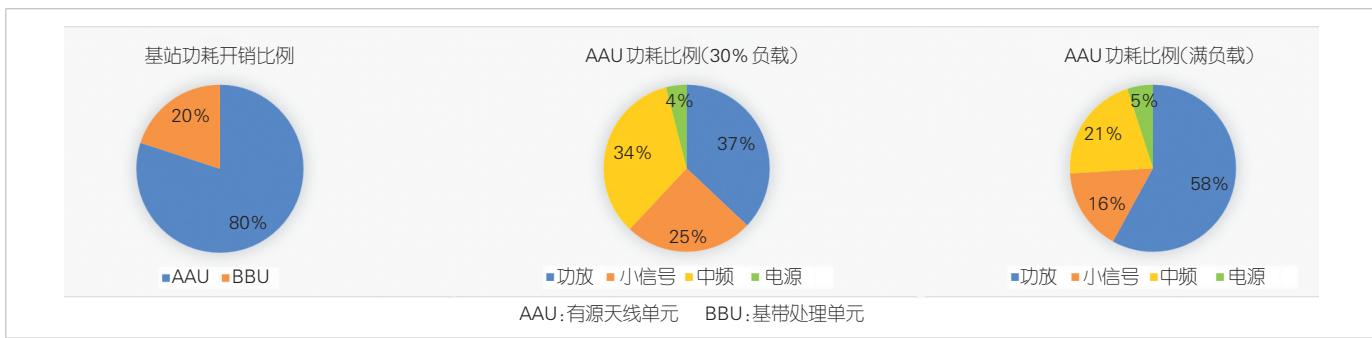
$$C = W \cdot \log_2 \left( 1 + \frac{S}{I + N} \right). \quad (1)$$

由  $S = G \cdot P_{\text{out}}$  可得  $C(P_{\text{out}}) = W \cdot \log_2(1 + G \cdot P_{\text{out}}/(I + N))$ ，易证明香农公式是凹函数。根据凹函数性质可得： $n \cdot C(P_{\text{out}}) \geq C(n \cdot P_{\text{out}})$ ，其中当  $P_{\text{out}} = 0$  或者  $n = 1$  时，等号成立。所以在实际传输时，我们假设信道特征不变，在用户的时延要求内，将当前给定的总功率拆分为多个小功率进行多时隙发送，那么系统总容量将会增大。也就是说，传输相同的数据量，相比于单时隙大功率传输，多时隙小功率传输将给系统带来更低的能耗。不仅如此，每个时隙分配的功率资源越少（总能耗不变的情况下），系统传输的容量越大，可节约的能量越多。

我们用软件仿真两个时隙的传输过程，具体如图2所示。横轴表示基站的发射功率，纵轴表示单位传输数据量，假设  $t_1, t_2, t_3$  时隙长度均为 1 s，若在  $t_1, t_2$  时隙分别消耗了 1 J 能量，总共消耗 2 J 能量，那么传输数据量为 2 个单位；若仅在  $t_3$  时刻单次消耗 2 J 能量，此时传输的数据量减少到 1.585 个单位。所以在同等能量消耗的条件下，两个时隙的小功率传输方案可以使得可服务数据量提升 26%，故可以通过感知和预测多个时隙的信道信息和业务到达情况进行业务整形，进而实现资源整形，将一定的功率分成多份后分配在各个时隙上进行数据传输，从而大大增加系统容量，提高网络能量效率。

### 1.2 功放能耗的非线性特征

在通信资源分配问题中，基站给每个用户分配不同的资源块，根据用户的需求，将携带用户数据的小信号通过基站的功率放大器进行一定比例的放大处理，并最终通过天线端发送给用户。目前，移动通信系统资源调度的大量研究工作主要考虑的是基站的输出功率如何分配，并没有发现实际传输时一段时间内功放能耗的非线性特征。理想的功率放大器的输入功率与输出功率是一种考虑截断效应的映射关系，在



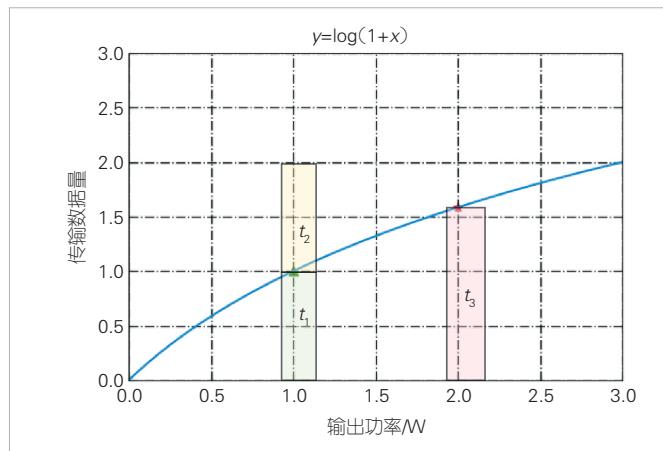
▲图1 基站能耗占比

饱和输出功率内，输出功率与输入功率是线性的关系。而实际的功率放大器的总输入功率与输出功率的映射是非线性的关系，功放在较大负载时会出现非线性的特点。

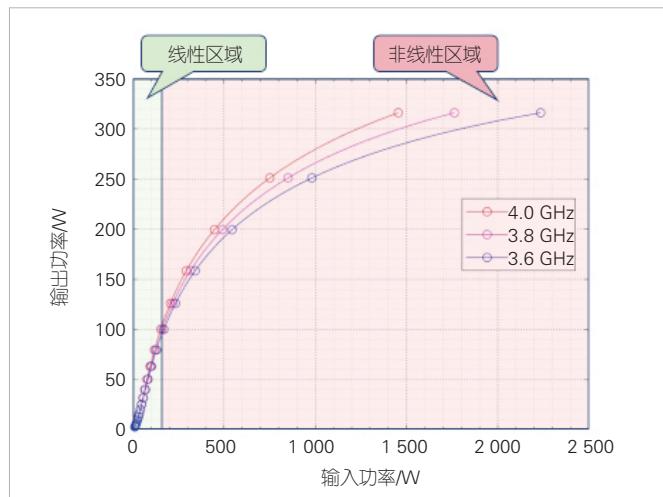
根据调研所绘制的测试数据如图3所示，基站工作在低负载区域时，功率放大器的输入功率和输出功率呈现近似线性的关系。在数据密度较大的业务场景中，因为基站工作在中高负载区域时，非线性特点非常突出，输出功率的增加趋势变缓，所以在单位时间内功放的输出能耗不会随输入能耗一直线性增长，在高数据密度业务场景下功放的输出能耗增长缓慢并最终趋于饱和。传统的资源调度方法只考虑优化输出功率，因此，结合系统传输能耗的非线性特征，我们提出了多时隙业务联合整形的思想，通过对业务请求的联合整形，让基站工作在低负载区域，从而降低系统能耗。

### 1.3 业务抖动与功放能耗的关系

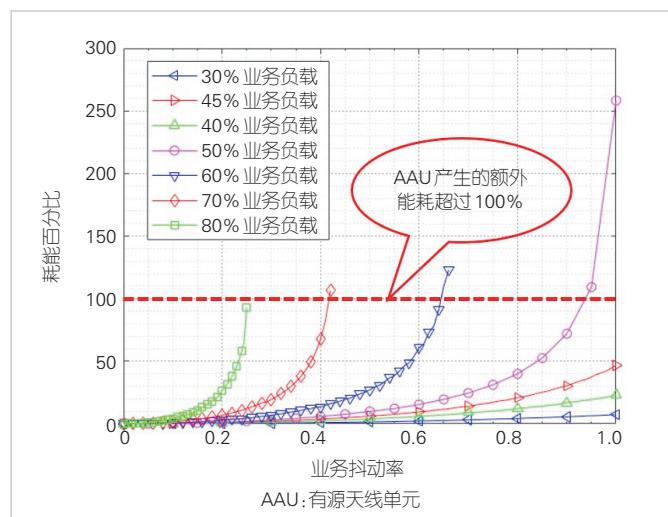
根据调研及实测数据，我们绘制了如图4所示的曲线。



▲图2 软件仿真时隙传输过程中对数函数的性质



▲图3 功放输入输出功率关系



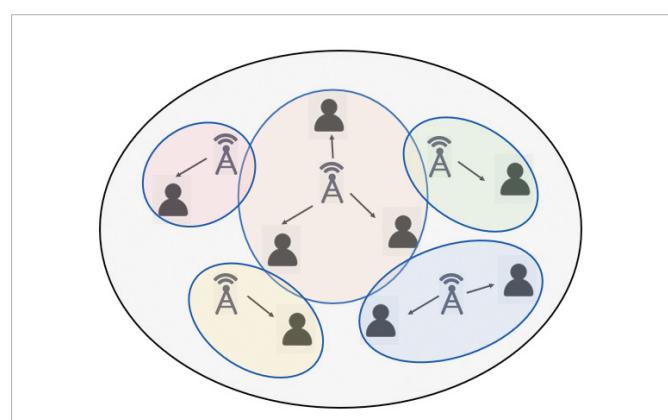
▲图4 业务抖动对功放能耗的影响

我们发现，如果基站的输出功率在一段时间内波动较大，则会产生额外的能量损耗，尤其当基站工作在中高业务负载区域时，由于业务抖动所引起的额外能耗会超过100%。额外能耗的产生是由于用户的业务请求在一段时间内突然增大导致，这使得基站工作在中高负载的非线性区域。因此，后续我们设计的重点是如何降低用户的业务请求的突发性，从而获得平稳的业务处理流，最终实现基站远离非线性工作区域、减少无线网络的额外能耗的目标。

## 2 通信网络建模

### 2.1 低能耗资源调度模型

通信网络简化示意如图5所示。我们将通信网络建模为图模型<sup>[9]</sup>  $G(V, E)$ ，每个基站、用户对应节点  $V$ ，基站和用户之间的连接关系为边  $E$ 。引入时隙的概念，对于每一个时隙  $t$ ，时隙长度设为  $Td$ ，单位为  $s$ ，观察时间窗口长度为  $T$ ，单位为  $s$ 。基站的节点特征变量包括：发送功率  $P_{on}(t)$ 、输入



▲图5 通信网络示意图

功率  $P_i(t)$ 、输出功率方差  $Pv_n(t)$ ；用户的节点特征变量包括：传输速率  $C_m(t)$ 、当前时隙用户剩余需要传输的数据  $Dr_m(t)$ 、当前时隙用户的总数据业务完成比例  $CR_m(t)$ ；基站与用户之间的信道增益和接入矩阵分别为  $G_{NM}(t)$  和  $X_{NM}(t)$ 。 $\mathcal{N}$  表示基站集合， $\mathcal{N} = \{1, 2, \dots, N\}$ ， $N$  表示基站的总数； $\mathcal{M}$  表示用户集合， $\mathcal{M} = \{1, 2, \dots, M\}$ ， $M$  表示基站的总数。

基站与用户进行通信时，会进行相应的匹配，并发射无线电信号。这些信号经过一定的路径损耗到达用户<sup>[10]</sup>。由香农公式，我们可以得到每个用户  $m$  在时隙  $t$  的服务速率，具体如公式 (2)：

$$C_m(t) = B \cdot \log_2(1 + \text{SINR}_{nm}(t)) \quad (2)$$

信噪比 (SINR) 是指信号干扰噪声比，即有用信号比其他信号的干扰加噪声。那么，SINR 在正交多址接入 (OMA) 系统下的计算公式为：

$$\text{SINR}_{nm}(t) = \frac{P_o_n(t) \cdot G_{nm}(t)}{\sum_{i \in \{N/n\}} P_o_i(t) \cdot G_{im}(t) + \sigma^2} \quad (3)$$

根据上述基础公式，我们建立了低能耗资源调度优化模型：

$$\begin{aligned} & \max_{\{X, P\}} \sum_t \frac{\sum_m CR_m(t)}{\sum_n (P_i_n(t) + Pv_n(t))} \\ \text{s.t. } & \left\{ \begin{array}{ll} x_{nm}(t) \in \{0, 1\} & \forall n \in N, \forall m \in M, \forall t \in T \\ \sum_m x_{nm}(t) \leq LK_n & \forall n \in N, \forall t \in T \\ \sum_m P_{nm}(t) \leq PO_n & \forall n \in N, \forall t \in T \\ 0 \leq CR_m(t) \leq 1 & \forall m \in M, \forall t \in T \end{array} \right. \end{aligned} \quad (4)$$

为实现优化目标，我们将业务完成率  $CR_m(t)$  进行最大化，基站输入能量  $P_i_n(t)$  和输出功率波动  $Pv_n(t)$  进行最小化。约束条件为：每一时隙每个用户最多连接一个基站；每个基站最多可接入用户数不超过其最大子载波数；每个基站分配给所有其接入用户的功率不超过该基站的最大功率阈值；每个用户的服务质量范围规定在 0~1 之间。

## 2.2 模型求解算法设计

随着人工智能 (AI) 技术的发展，深度学习<sup>[11]</sup>和强化学习<sup>[12]</sup>被逐渐应用于资源调度领域。根据上述建立的系统模型，我们设计了多时隙业务联合整形的深度强化学习 (MSRS-HDDPG) 算法，并在核心架构中采用了深度确定性策略梯度 (DDPG) 算法<sup>[13]</sup>。

具体的强化学习三元组设计如下：

1) 状态空间  $s_t$ ：具体包括用户业务量  $D_m(t)$ 、用户未完成传输的数据量  $Dr_m(t)$ 、每个用户的完成率  $Cr_m(t)$ 、信道信息  $G_{nm}(t)$ 、基站输入功率  $P_i_n(t)$ 、基站输出功率方差  $Pv_n(t)$ 。

2) 动作空间  $a_t$ ：具体包括用户和基站的连接关系  $X_{nm}(t)$ 、基站的发射功率  $Po_n(t)$ 。

$$3) \text{ 奖励函数 } r_t: r_t = \frac{\sum_m CR_m(t)}{\sum_n (P_i_n(t) + Pv_n(t))}.$$

具体的网络架构设计如下：

1) 动作网络输入状态信息  $s_t$ ，输出需要优化的动作  $a_t$ 。

2) 评价网络输入  $(s_t, a_t)$ ，输出当前状态下选择该动作的价值  $Q_t$ 。

3) 目标动作网络输入  $s_{t+1}$ ，输出  $a_{t+1}$ ，即下一时刻的动作估计；目标评价网络输入  $(s_{t+1}, a_{t+1})$ ，输出下一时刻的动作价值  $Q_{t+1}$ 。

多时隙业务联合整形的模块具体设计如下：

1) 假设时隙为  $t$ ，用户  $m$  的数据业务请求量为  $D_m(t)$ ；

2) 根据历史数据估计每个基站  $n$  的负载情况；

3) 若基站工作在非线性区域，那么将功率的最大值约束设为线性区域的最大发射功率，并计算用户在时隙  $t$  的未完成业务量  $Dr_m(t)$ ， $D'_m(t+1) = Dr_m(t) + D_m(t+1)$ ；

4) 若基站工作在线性区域，则根据强化学习算法直接完成当前时隙的全部业务请求；

5) 重复上述步骤 3)、4) 至时隙  $T$ 。

算法的具体执行流程如算法 1。

### 算法 1 多时隙业务联合整形的深度强化学习算法

初始化评价网络、动作网络和对应目标网络的参数

初始化缓存空间，存放探索的状态转移向量作为神经网络训练样本

**for** episode = 1, 2, ..., EPISODES **do**

    初始化探索状态  $s_0$

    感知和预测一段时间内信道信息

**for**  $t = 1, 2, \dots, T$  **do**

        根据多时隙的请求和信道估计进行业务联合整形

        随机输出探索动作  $a_t$  并将输出功率映射到输入功率

        根据整形后的业务需求设计当前时隙的奖励函数  $r_t$

        执行后观测当前时隙的奖励和下一时刻的状态  $s_{t+1}$

        将状态转移  $(s_t, a_t, r_t, s_{t+1})$  存入缓存空间

```

从缓存空间中随机取样 M 批样本状态转移( $s_i, a_i, r_i, s_{i+1}$ )
根据最小化目标价值和当前网络价值均方误差更新评价网络
    根据最大化价值网络策略梯度更新动作策略网络
    根据滑动更新算法更新目标网络参数
end for
end for

```

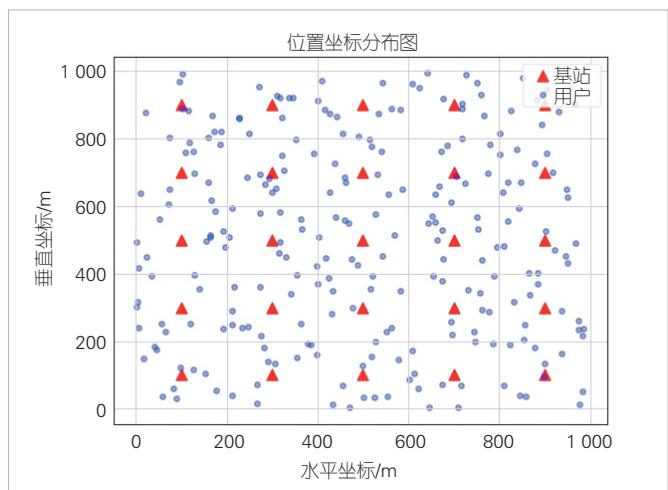
### 3 仿真结果与分析

本文中，我们设计在面积为  $1 \text{ km}^2$  的区域内随机生成 250 个用户的坐标，同时固定生成了 25 个基站，用户和基站的分布如图 6。

具体仿真参数如表 1 所示。

此外，在仿真时，针对基站功放输入功率和输出功率的映射关系，本文中我们选取了具有凹函数性质的映射函数。可以证明，系统总的优化目标是输入功率的凹函数，一段时间内的功放能耗存在非线性特性。

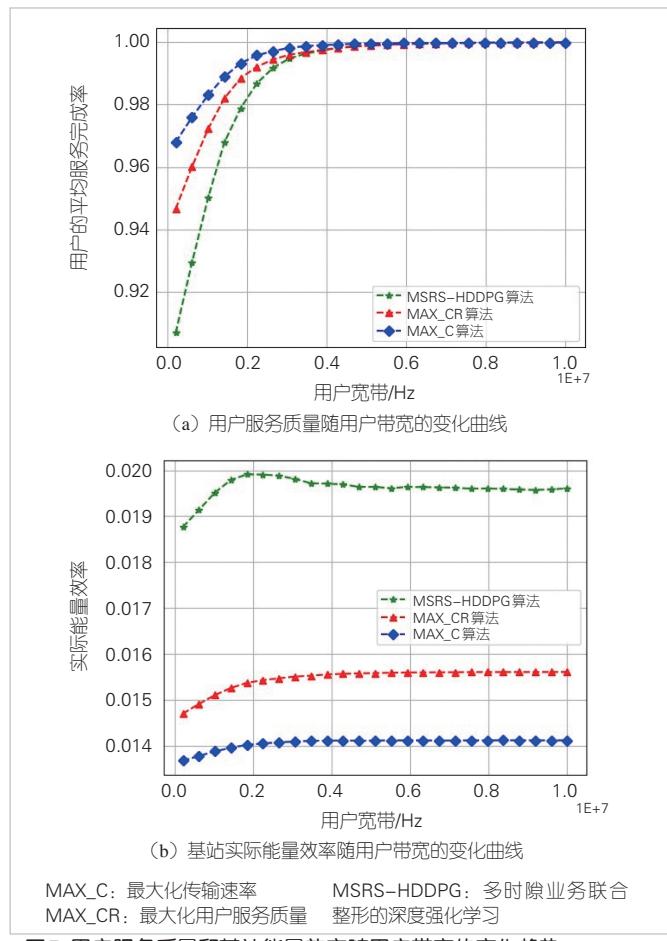
用户服务质量质量和基站能量效率随用户带宽的变化趋势如图 7 所示，通过仿真不同带宽下的各种算法性能可知，我们提出的算法的实际能量效率远高于另外两种算法（分别高出 25.6% 和 40%）。实际能量效率的定义为：基站的服务完成率和基站消耗的总输入能量的比值，这充分反映了用户的服务质量和基站能耗的平衡关系。



▲图 6 基站与用户的随机分布图

▼表 1 仿真参数

参数符号	数值	参数符号	数值
基站数目 $N$	25	传输频段 $F$	5 GHz
基站最大发射功率 $PO_n$	20 W	最大多普勒频移 $f_d$	20 Hz
基站最大连接数目 $LK_n$	10	时隙长度 $Td$	0.01 s
用户数目 $M$	250	时间窗口长度 $T$	0.1 s

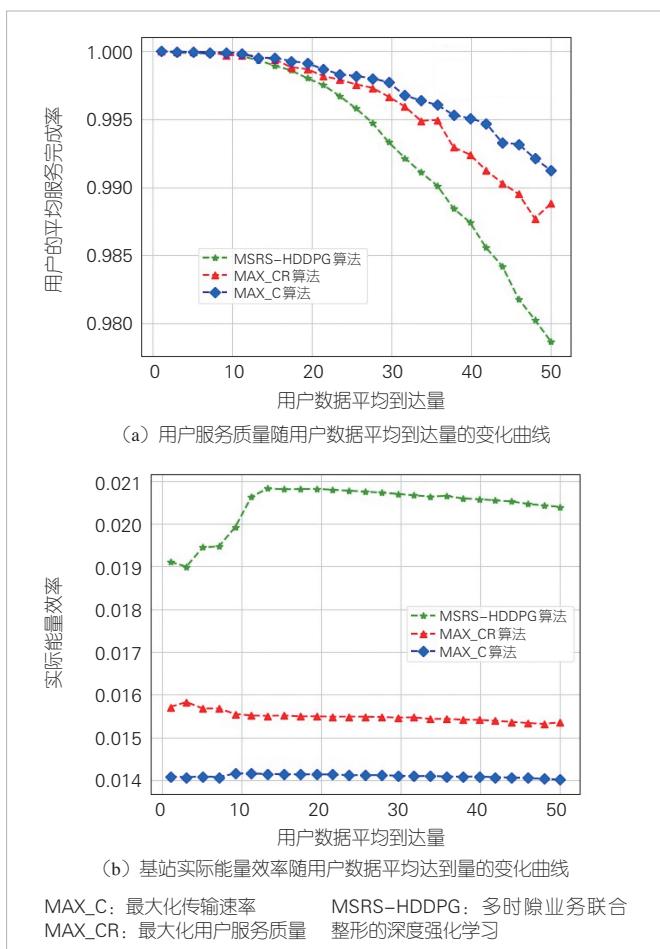


▲图 7 用户服务质量质量和基站能量效率随用户带宽的变化趋势

用户服务质量质量和基站能量效率随业务量的变化趋势如图 8 所示，通过仿真不同数据访问量下各种算法的性能可得：在较大数据量的访问情况下，本文所设计的算法的业务完成率仅损失约 1%，但时实际能量效率分别高出 32.9% 和 45.7%。

### 4 结束语

有限的通信资源和日益增长的个性化用户数据业务导致未来 6G 移动系统的通信资源管控问题越来越突出。针对这种供需匹配问题，同时为了响应国家提出的“双碳”战略，本文中我们提出了多时隙业务联合整形的深度强化学习算法。该算法充分考虑了系统的传输能耗和功放能耗的非线性特征，从时间尺度上进行了长期优化策略的训练，从而对优化目标做出了更加准确的调整。该算法旨在基本相同的服务质量条件下，降低系统的总能耗。对比最大化用户服务质量 MAX\_CR 和最大化传输速率 MAX\_C 两种算法的性能，我们提出的算法在服务质量上约损失 6%，但节约了近 33% 的基站总输入能量。本文定义的实际能量效率指标分别比对比算法高出 28.57% 和 38.46%；增大带宽和用户的数据量时，本



▲图8 用户服务质量及基站能量效率随业务量的变化趋势

文所提算法的实际能量效率  $\eta_{AEE}$  均有所提高，实现了低能耗资源管控的目标。

未来，在应对6G超大数据密度业务场景，如虚拟现实(VR)等多维数据业务场景时，本文所设计的算法的优势会更加明显：可以在保障一定服务质量下充分节约基站的功放能耗，从而极大地改善整个网络的性能。当网络规模逐渐扩大时，算法的收敛速度问题亟待解决。因为时间成本过高会导致资源调度效果较差，而且不同的场景规模重新训练的成本较高，所以未来需要集中研究超大动态规模网络的算法设计问题，实现不同场景规模下的算法快速移植和快速收敛，最终实现网络对当前场景特征的快速挖掘和学习。

## 参考文献

- [1] 华为.《6G:无线通信新征程》白皮书 [R]. 2022
- [2] 王友祥,裴郁杉,黄蓉,等.6G通感算一体化网络架构和关键技术研究 [J].移动通信,2023,47(9):1-10. DOI: 10.3969/j.issn.1006-1010.20230904-0002
- [3] 刘光毅,邓娟,郑青碧,等.6G智慧内生:技术挑战、架构和关键特征 [J].移动通信,2021,45(4): 68-78
- [4] ZHANG H J, JIANG C X, BENNIS M, et al. Heterogeneous ultra-dense networks: part 1 [J]. IEEE communications magazine, 2017, 55(12): 68-69. DOI: 10.1109/mcom.2017.8198804
- [5] 李剑飞.异构蜂窝网络的资源管控技术 [D]. 西安:西安电子科技大学, 2023
- [6] 科技部,国家发展改革委,工业和信息化部等.科技支撑碳达峰碳中和实施方案(2022—2030年) [EB/OL].(2022-06-24)[2023-10-13]. [https://www.gov.cn/zhengce/zhengceku/2022-08/18/content\\_5705865.htm](https://www.gov.cn/zhengce/zhengceku/2022-08/18/content_5705865.htm)
- [7] 彭钊,张旭.5G基站能耗分析与节能探讨 [J].通信与信息技术,2021(2): 49-50
- [8] SHANNON C E. A mathematical theory of communication [J]. Bell system technical journal, 1948, 27(3): 379-423. DOI: 10.1002/j.1538-7305.1948.tb01338.x
- [9] 陈恩艺.基于图论的超密集组网技术研究 [D].北京:北京理工大学, 2016
- [10] 任帅.无线通信衰落信道建模及其在信道仿真器中的应用 [D].合肥:合肥工业大学
- [11] 王莎莎.基于深度学习的通信网络流量实时监测方法 [J].长江信息通信, 2023(8): 161-163
- [12] 赵若帆.基于深度强化学习的资源调度研究 [D].北京:北京交通大学, 2022
- [13] LILLICRAP T P, HUNT J J, PRITZEL A, et al. Continuous control with deep reinforcement learning [EB/OL]. [2023-10-12]. <https://arxiv.org/abs/1509.02971.pdf>

## 作者简介



李建东，西安电子科技大学教授、“长江学者”特聘教授、IEEE Fellow、中国电子学会会士、中国通信学会会士、西安电子科技大学空天地一体化综合业务网全国重点实验室学术委员会副主任、国家新一代宽带无线移动通信网重大专项总体组专家、中国首批“新世纪百千万人才工程”国家级人选、第3届教育部“青年教师奖”获得者、2007年国家杰出青年科学基金获得者、科技部国家“863”计划15周年先进个人等；主要研究领域为宽带无线通信、智能无线网络、大规模自组织网等；主持“863”计划重大课题、国家自然科学基金重大项目等，获得国家技术发明二等奖2项、省部级科技进步奖6项；已发表论文80余篇，出版教材和专著5部。



牛淳隆，西安电子科技大学在读硕士研究生；主要研究领域为低能耗移动通信网络资源调度、强化学习等。



赵晨曦，西安电子科技大学副教授；主要研究领域为空天地一体化网络组网及资源调度、低能耗移动通信网络资源调度、多智能体强化学习等；主持国家自然科学基金1项，参与国家级科研项目4项；已发表论文7篇，申请国家发明专利3项。



刘俊宇，西安电子科技大学教授；主要研究方向为无线覆盖技术、异构密集无线网络容量理论及组网技术、天地一体化网络智能组网技术等；主持国家重点研发计划课题、国家自然基金面上项目、陕西省重点研发计划等国家级及省部级项目10余项；已发表论文60余篇，授权国家发明专利30余项。

# 低开销智能反射面辅助无线通信研究综述



## Survey on Low-Overhead Reconfigurable Intelligent Surface Assisted Wireless Communication

陈天贝/CHEN Tianbei, 李娜/LI Na, 陶小峰/TAO Xiaofeng

(北京邮电大学, 中国 北京 100876)

(Beijing University of Posts and Telecommunications, Beijing 100876, China)

DOI: 10.12142/ZTETJ.202306006

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231212.1443.006.html>

网络出版日期: 2023-12-13

收稿日期: 2023-10-15

**摘要:** 智能反射面 (RIS) 为无线网络注入新型绿色资源, 但也带来了新的组网开销, 成为制约其实际应用的关键。在考虑开销的基础上, RIS 是否依然能够发挥其赋能无线通信的潜在价值、如何降低 RIS 开销, 以及如何减小 RIS 开销的影响, 是亟待探索的问题。对低开销 RIS 辅助无线通信进行了系统综述, 梳理了 RIS 开销模型, 包括 RIS 通信开销模型和 RIS 电路开销模型, 分析了开销问题对 RIS 协助无线网络性能的负面影响, 并整理了 5 种常见的低开销方案。结合 6G 网络发展趋势, 对未来亟待解决的 RIS 开销问题及相关研究方向进行了总结。本研究有助于激发 RIS 的无线赋能潜力, 从而推进 RIS 技术在更多场景中的应用。

**关键词:** 智能反射面; 低开销设计; 信道估计; 信道反馈; 相位配置; 功耗开销

**Abstract:** Reconfigurable Intelligent Surfaces (RISs) are new green resources for wireless networks, but still bring in extra networking overhead, which may seriously limit their performance in practical networks. Based on these considerations, it is important to study whether and how RIS can still enhance wireless communication performance, and how to mitigate the impact of RIS overhead. A systematic review of low-overhead RIS-assisted wireless communication is presented. Two models of the RIS overheads are introduced, including communication overhead and circuit power consumption of the RIS. The corresponding impact on the communication performance is analyzed, then five low-overhead solutions are introduced. Finally, in line with the development trends of 6G networks, the unresolved issues are summarized and the corresponding research directions that need urgent attention in the future are envisioned. Our work will contribute to inspiring the potential of RIS-assisted wireless communication and promote RIS's application in more scenarios.

**Keywords:** reconfigurable intelligent surfaces; low-overhead design; channel estimation; channel feedback; phase configuration; power consumption

**引用格式:** 陈天贝, 李娜, 陶小峰. 低开销智能反射面辅助无线通信研究综述 [J]. 中兴通讯技术, 2023, 29(6): 29–38. DOI: 10.12142/ZTETJ.202306006

**Citation:** CHEN T B, LI N, TAO X F. Survey on Low-overhead reconfigurable intelligent surface assisted wireless communication [J]. ZTE technology journal, 2023, 29(6): 29–38 DOI: 10.12142/ZTETJ.202306006

全球范围内已广泛开启 6G 移动通信技术研究的新篇章。国际电信联盟 (ITU) 发布《IMT 面向 2030 及未来发展的框架与总体目标建议书》<sup>[1]</sup>, 描述了 6G 的场景、需求、能力、关键技术推动因素等。其中, “可持续性”作为新增指标, 要求未来通信网络采用高效、低功耗技术, 为可持续发展目标作贡献<sup>[2-3]</sup>。同时, 中国在 2020 年提出“双碳”战略目标, 近年来稳步推进能源产业的结构调整、其他产业的低碳转型等。积极构建绿色低碳的移动通信网络是重要内容之一。

5G 移动通信网络主要采用优化硬件架构设计、加快半导体工艺, 以及提高硬件集成度的硬件节能技术与包括符号关断、通道关断、小区关断等节能技术<sup>[4]</sup>, 一定程度上支撑了移动通信产业的绿色低碳需求, 但距离实现最终的“双碳”目标仍有较大差距。根据华为提供的数据显示, 典型的 5G 基站能耗为 11.5 kW<sup>[5]</sup>, 其中, 超过 75% 的能耗来自有源天线单元中的功放、数字基带、收发信板等关键器件<sup>[6]</sup>。为满足日益增长的通信速率需求, 6G 网络将引入更多有源节点, 采用更高通信频段等。这些措施会带来更高的能量消

耗、硬件开支和维护成本<sup>[7]</sup>。

6G网络在能耗问题上将面临更高的需求和更大的挑战<sup>[8]</sup>，亟须突破新的高效组网架构及组网技术。智能反射面（RIS）是一种新型物理维度的无线传输技术，具备独特的低成本、低能耗、可编程、易部署等优点，被业界视为一种绿色的、可持续的、生态友好的潜在技术方案，通过对空间电磁波的幅度、相位、极化等特征的按需调控，可以构建智能可编程的无线环境<sup>[9]</sup>。目前，RIS协助的无线通信技术受到全球学界和产业界的广泛关注。一方面，现有研究对传统有源天线的波束赋形和RIS节点的反射系数进行联合设计，可优化提升无线网络的能量或频谱效率<sup>[10]</sup>，增强可靠性或安全性<sup>[11]</sup>等。另一方面，产业界已开展RIS原型的研发与测试，例如：日本DoCoMo公司于2018年完成RIS外场测试，指出在28 GHz频段可将通信速率提升500 Mbit/s<sup>[12]</sup>；美国麻省理工学院于2019年搭建2.4 GHz非授权频段的测试平台RFocus，表明RIS可将信号强度提升约10倍<sup>[13]</sup>；清华大学戴凌龙教授团队于2020年成功研制256单元、2 bit离散相位的RIS原型，在28.5 GHz频段可实现19.1 dBi的天线增益<sup>[14]</sup>；中国移动联合东南大学于2021年完成室外测试验证，证明RIS可使边缘用户吞吐量提升10倍以上<sup>[15]</sup>；同年，中兴通讯联合中国电信完成5G高频外场的RIS测试验证，指出在26 GHz频段可使处于覆盖盲区或弱区的用户接收信号强度提升12.5 dB<sup>[16]</sup>。

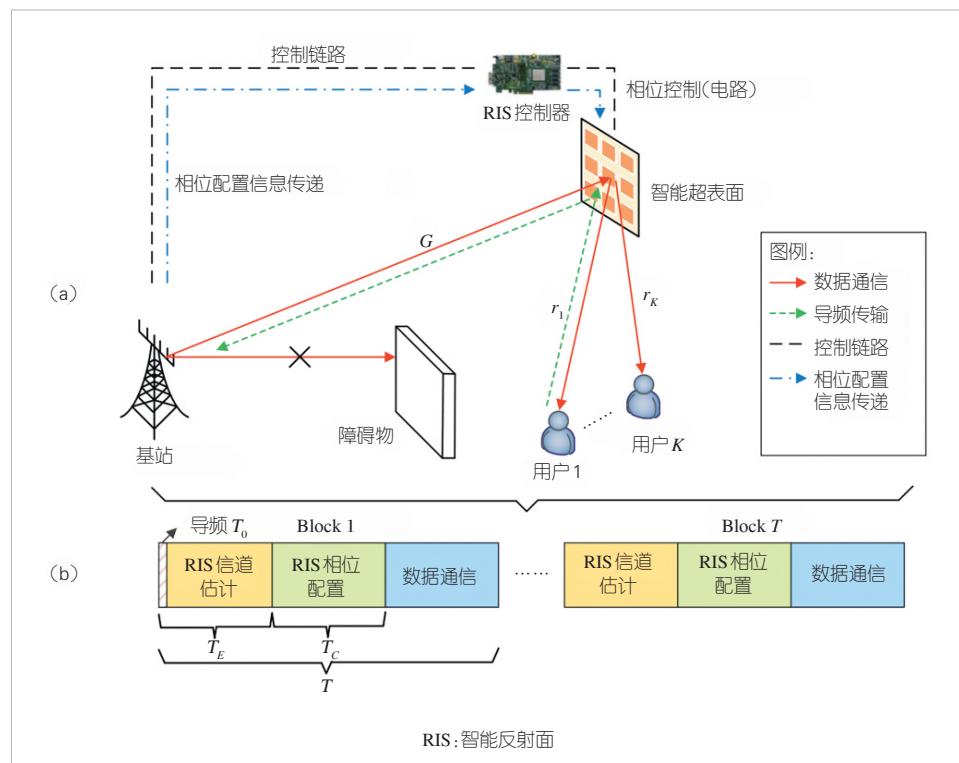
上述工作从理论和实践两个角度证明了RIS优化提升网络性能和用户体验的重要价值。然而，东南大学崔铁军院士指出，RIS的大规模阵列将导致庞大的信道规模，造成巨大的训练开销，同时由于受到硬件和功耗的限制，训练过程中反射单元的切换也将带来不可避免的时间开销<sup>[17]</sup>；欧洲电信标准协会（ETSI）RIS工业规范组的创始成员和学术副主席、IEEE Fellow MARCO D. R.教授指出，绝大多数关于RIS协助无线通信的研究仅聚焦于数据传输阶段，忽视了信道状态信息获取、RIS相位信息反馈及配置阶段产生的系统开销，而这些开销会降低甚至抵消RIS带来的性能增益<sup>[18]</sup>；清华大学戴凌龙教授的研究表明，RIS信道状态信息获取的导频开销与基站天线数及RIS单

元数成正比，这在实际大规模部署或移动场景下将是不可承受的，且会导致实际性能与理论性能出现较大偏差<sup>[19]</sup>。因此，在考虑上述系统开销的前提下，RIS是否依然能够为无线网络带来性能增益，以及如何容忍或降低RIS开销的负面影响，是值得探索的关键问题。

## 1 RIS协助无线网络中的开销模型

相关研究表明<sup>[20]</sup>，当RIS节点靠近基站或终端时，可更好地发挥其赋能无线通信的潜在价值。典型的RIS协助无线网络模型如图1（a）所示。由于多数RIS设备不具备完善的基带信号处理能力，需要由基站或终端来完成信道状态信息估计、RIS调控参数设计以及资源分配等复杂计算，之后通过有线<sup>[21]</sup>或无线<sup>[22]</sup>的控制链路将RIS调控参数传递到RIS控制器，由其完成对RIS单元相位、幅度等参数的配置，最后，便可基于RIS建立的辅助通信链路进行无线数据传输。典型的传输阶段示意图如图1（b）所示。

当RIS单元或调控参数信息的数量较大时，相应的信道信息估计、控制信息交互所产生的传输开销不容忽视，将造成显著的时间、能量等开销，进而压缩数据通信资源，制约通信系统性能。此外，RIS控制电路的运行也会消耗额外的能量，一定程度上增加移动通信网络的能耗负担。下面我们将分别介绍RIS协助无线网络中的开销模型。



▲图1 RIS协助无线网络模型及传输阶段示意图<sup>[23]</sup>

### 1.1 RIS 通信开销模型

正式数据通信阶段之前需要经历 RIS 信道信息的提取以及 RIS 调控参数的传递过程，然而多数 RIS 由全无源的反射单元构成，不具备基带处理能力，需要基站与终端的配合来完成对基站-RIS-终端级联信道信息的提取<sup>[24]</sup>，包括信道信息估计及必要的信道信息反馈。文献[18, 25]以单 RIS 协助的点对点无线通信为例，针对两种不同的导频发送模式，即逐一发送导频模式和正交导频传输模式，建立了由信道信息估计造成的时间与能耗开销的数学模型。两种传送模式下，相应的时间与功耗开销均与 RIS 反射单元数  $N$  呈线性关系。面对实际中部署的成百上千的反射单元，信道信息估计的开销无法忽略。

终端侧获得估计的级联信道信息后，需要将估计获得的信道信息反馈给基站，以进行资源优化设计<sup>[26]</sup>。对于时分双工（TDD）系统而言，上下行信道具有互易性，下行信道信息可直接由上行信道估计获得，而频分双工（FDD）系统则需要借助额外的反馈链路，以获取下行信道信息<sup>[27]</sup>。尽管已有很多经典的信道反馈方法，例如随机向量量化（RVQ）反馈方案<sup>[28]</sup>、基于码本的反馈方案<sup>[29]</sup>、基于深度学习的反馈方案<sup>[30]</sup>等。然而，这些针对传统通信系统的设计并不能直接应用于 RIS 协助的通信系统，例如：传统 RVQ 码本尺寸呈  $2^B$  级指数增长<sup>[31]</sup>，其中， $B$  表示每个接收天线单元对应信道的反馈比特数。为充分实现复用增益所需的反馈比特数将随着天线数的增加而线性增加<sup>[28]</sup>。面对具有大量反射单元的 RIS 协助无线通信，反馈链路开销将更加不可忽视。研究证明<sup>[32-33]</sup>，与传统无线网络相比，RIS 协助无线网络面临更为严重的信道开销挑战。特别地，当 RIS 用于高速移动场景时，信道环境的快速变化需要频繁执行信道信息估计，进而造成更加显著的时间开销<sup>[34]</sup>。

紧接着，基于前述获得的信道信息，基站或终端将进一步计算最佳的 RIS 单元调控策略，并将相应参数传递给 RIS 控制器，之后由其完成对 RIS 单元的相位、幅度等参数的配置。然而，受到控制链路带宽的约束，RIS 调控参数的传递过程会产生必要的时间开销<sup>[35]</sup>。同样以单 RIS 协助的点对点无线通信为例，文献[18, 25]建立了无线控制链路下 RIS 调控参数传递导致的时间与功耗开销的数学模型，其中，时间开销受到 RIS 单元数、控制链路带宽以及控制链路发送功率的影响。类似地，文献[36]建立了有线控制链路下 RIS 调控参数传递带来的时间开销的数学模型，该模型受到控制信息总比特数与控制链路数据速率的制约。另外，文献[37]提出采用 RIS 分组策略下配置的时间开销模型仅与单个 RIS 单元的调控时间及 RIS 分组数有关。

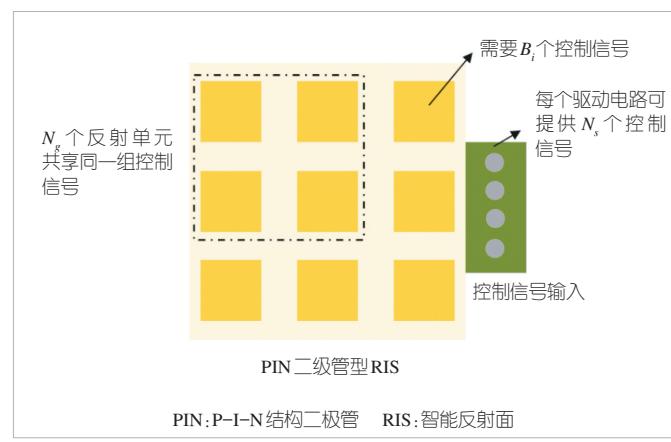
### 1.2 RIS 电路开销模型

RIS 单元的调控需要通过改变外部施加的偏置电压来实现，产生的硬件功耗将会影响整个网络的能量效率（EE）或频谱效率（SE）。同时由于 RIS 本身由数量庞大的反射单元构成，这种影响无法忽视。文献[35, 38-39]建立了单个 RIS 的理论功耗开销模型，相应的功耗开销与 RIS 单元数、每个 RIS 单元的离散相位量化阶数  $b$  密切相关。具体地，文献[35]列举出 3~6 bit 离散相位量化阶数下每个 RIS 单元的理论功耗值为 1.5 mW、4.5 mW、6.0 mW、7.8 mW。不难看出，当离散相位量化阶数或 RIS 单元数较大时，RIS 面板整体的功耗开销将十分显著。针对多 RIS 协助无线通信场景，RIS 功耗将随着面板数量的增加而线性增加。除理论模型外，东南大学崔铁军院士、唐万恺副研究员、金石教授等实测验证了 RIS 的功耗开销模型<sup>[40-41]</sup>，包括静态功耗与动态功耗。其中，静态功耗又进一步包括控制面板的功耗和驱动电路的功耗，前者视为常数，取值为 4.8 W，后者受到控制信号总数、驱动电路的输出控制信号数的影响，如图 2 所示。典型实测驱动电路功耗值为 1 720 mW。动态功耗与 RIS 单元的极化方式、离散相位量化阶数以及编码状态有关，35 GHz 频段下每个 1 bit 相位控制的 RIS 单元的典型实测功耗值为 12.6 mW。

### 2 开销对 RIS 协助无线网络性能的影响

考虑前述 RIS 开销的前提下，RIS 是否依然具备赋能无线通信的潜在价值，是需要探讨的根本问题之一。已有研究表明，RIS 开销制约下的无线网络性能将严重偏离其可达的理论性能上界，甚至出现零增益或负增益<sup>[18]</sup>现象。

针对经典的 RIS 协助点对点多天线无线通信场景，信道估计以及相位信息反馈带来的时间开销对系统的 EE 或 SE 的影响得以研究<sup>[18]</sup>，发现开销的存在将严重降低系统的 EE 或 SE，甚至出现 RIS 本身的性能增益远远无法弥补开销带来的



▲图 2 驱动电路功耗示意图<sup>[40]</sup>

性能损失的情况。例如，当RIS单元数为50时，系统性能比理论上界下降约20%；而当单元数增加至150时，系统性能将下降约300%。类似地，文献[42]对导频/控制开销与可达速率之间的权衡关系展开研究，发现为缓解导频开销对网络性能的负面影响，当单元数较大时需要选择较低的离散相位量化阶数以减少控制信号的数量。此外，当开销达到一定程度时，单元数的增加反而使得可达速率逐渐下滑，比几乎不考虑开销的情形下降近72.4%。

针对RIS协助多用户无线通信场景，文献[37]对相位配置时间开销与系统吞吐量之间的权衡关系展开研究，发现对全部RIS单元的反射系数进行完全重配可以实现较高的瞬时数据速率，然而这一举措势必会缩短数据传输的时间。例如，当配置的时间占总时间的10%时，系统吞吐量比不考虑配置开销的情形降低78.6%左右。

针对多RIS协助无线通信场景，控制链路资源对系统遍历速率的影响得以研究<sup>[43]</sup>，发现多RIS协助用户通信的方式虽然能显著提升信号功率，但会带来极高的同步信令开销。随着RIS数量的增加，不同RIS之间的严格同步会带来更大的信令开销，进而严重挤占控制链路资源。例如，相比于控制链路容量较大的情形 ( $C_B^{\max} = 150 \text{ Mbit/s}$ )，系统的遍历速率将降低17.0%~47.9%。

另外，文献[35, 38]针对RIS自身功耗对系统EE性能的影响展开研究，发现EE随RIS单元相位量化阶数的增加而降低。例如，当量化阶数从4 bit增加至6 bit时，EE降低约37.3%。与此同时，该研究还发现EE与RIS单元数之间存在权衡关系。当单元数相对较少时，RIS引入带来的空间自由度将提升EE；然而当单元数相对较多时，RIS自身难以承受的功耗将抵消其对EE的性能增益，造成EE的下滑。

### 3 RIS协助无线网络的低开销方案

如前所述，信道信息的获取、RIS调控参数的传递和配置均会导致额外开销，严重时将制约通信系统的整体效率。因此，如何降低或容忍这些开销是需要探讨的另一个根本问题。针对此问题，一方面，业界专家广泛研究低开销的信道估计与反馈方法，充分挖掘RIS信道的空间相关性、稀疏性等特性，保障RIS信道估计与反馈阶段较低的开销；另一方面，一些研究综合考虑RIS的价值和成本，提出例如贪婪的RIS面板开/关控制算法等新算法，可有效提升系统的能量效率。

#### 3.1 低开销的信道估计方案

RIS的信道估计可以采用传统思路或方法。例如：为克

服RIS本身无法收/发导频信号的缺点，部署有源RIS节点并将输出连至接收射频链路，以直接实现基带信道估计<sup>[44]</sup>；逐个开启/关闭RIS单元，以实现信道估计的开启/关闭型信道估计策略<sup>[45]</sup>，以及基于DFT的最小二乘（LS）估计方法<sup>[23]</sup>和最小均方误差（MMSE）估计方法<sup>[46]</sup>。然而，这些方法中有源RIS节点的引入以及RIS单元逐个开启/关闭的操作将为系统带来更高的功耗、硬件设计复杂度以及信道估计开销，其根本原因是未考虑RIS信道的特点，例如RIS信道的空间相关性、稀疏性、双时间尺度特性等。利用这些特点，可设计低开销的信道估计方案。

利用RIS信道的空间相关性，文献[47]提出RIS分组策略将RIS划分为若干组，每组RIS单元共享相同的反射系数。对此信道估计阶段只需开启/关闭对应分组的RIS单元，使得导频数从 $N + 1$ 降至 $N_{\text{group}} + 1$ 。当 $N_{\text{group}}$ 分组数较多时，可以显著减少开销。

利用RIS信道的稀疏性，文献[48–49]将信道估计问题转化为级联稀疏信道的恢复问题，提出正交匹配追踪（OMP）、高斯混合近似消息传递（GAMP）算法，并引入残差网络，以降低对应的开销。另外，文献[50]提出基于快速交替的LS算法（FALS），只需观测信道矩阵的特定行/列即可恢复完整信道，相比于OMP算法可有效降低约45%的训练开销。

实际网络部署后，基站与RIS的位置往往就固定不变，基站-RIS之间的信道维数较高但变化缓慢，具有准静态特性；反之，用户处于移动状态，用户-基站/RIS之间的信道维数较低但变化较快，具有时变特性。这一性质被称为RIS信道的双时间尺度特性<sup>[51]</sup>。基于这一性质，文献[32, 52]提出低开销的分段信道估计方案，针对基站-RIS信道提出双链路导频传输方案。基站发送下行导频并接收经RIS反射的上行导频，以此基于坐标下降算法恢复信道。由于基站-RIS信道具有准静态特性，长期观察内它的估计频率相对较低，对此可以将开销降低至 $2(N + 1)/\alpha + K\lceil N/M \rceil + K$ ，其中， $\alpha$ 为描述导频开销的因子， $M$ 表示基站天线数， $K$ 表示用户数。

利用不同用户RIS信道之间的相关性，文献[52–54]提出三阶段的信道估计框架，以有效降低不同用户级联信道之间存在的大量冗余。具体做法是：首先第一/二阶段实现典型用户的级联信道估计，接着第三阶段利用与典型用户信道的强相关性实现其他用户的级联信道估计，可以将原本 $KMN + KM$ 个信道系数降低至 $K + N + \max(K - 1, \lceil (K - 1)N/M \rceil)$ 。类似地，文献[55]则提出双结构正交匹配追踪算法（DS-OMP）来减少导频开销。

除此之外，基于深度学习的降低导频开销方法也备受瞩

目。文献[56]提出时域信道子采样和双深度神经网络的方法，神经网络中采用神经普微分方程（ODE）描述RIS动态信道以提高循环神经网络对时间序列的重建能力，进而降低动态时变信道的训练开销。文献[57]提出模型驱动的深度展开神经网络框架以降低训练开销，使训练开销相比于传统的LS可降低约25%。另外，文献[58-59]提出基于码本的信道估计与波束赋形协同方案。该方案利用位置信息离线生成具备环境感知能力的反射系数组码本，根据需要选择对应的最佳码字，可以将信道估计开销降低至 $QK$ ，即与RIS单元数无关，仅与码本的大小 $Q$ 和用户天线数 $K$ 有关，进而有助于显著降低开销。

### 3.2 基于统计或混合CSI的低开销设计研究

对RIS协助无线网络性能的分析大多是基于瞬时信道状态信息（CSI）。然而，瞬时CSI的准确获取具有较大挑战，庞大的RIS单元数将造成极高的信令交换开销。与此同时，RIS不同状态之间的切换通常比较缓慢，约为几兆赫兹<sup>[60]</sup>。这种硬件的不完美将导致导频传输的非连续性，增加估计的时间开销。为实现信道估计开销和网络性能之间的权衡，相关研究基于统计或混合CSI来设计RIS的相位偏移。

由于统计CSI变化相对较缓，系统有充足的时间来获取相关信息，无须对其频繁地更新，因此信令开销可以大大减少<sup>[61]</sup>。文献[62-63]采用长期观察获得的统计CSI联合设计，发送波束赋形以及RIS的相位，以最大化遍历容量或速率。另外，当网络中存在窃听者时，由于它往往不会与基站主动交互，因此窃听信道的完美CSI无法获得。对此，文献[64]基于统计CSI联合设计以最大化遍历安全速率。

然而，前述方案中瞬时CSI的缺失可能会降低系统实时性能，对此采用瞬时和统计CSI的混合解决方案可以进一步改善开销与系统性能之间的权衡问题。文献[39, 65]联合设计用户侧的波束赋形和RIS的相位，以最大化RIS协助上行多用户多输入多输出（MIMO）系统的全局能量效率<sup>[39]</sup>以及加权EE和SE的总和<sup>[65]</sup>。另外，由于网络中窃听者位置的不确定性，因此难以实时获取与窃听者有关的瞬时CSI。对此，文献[66]基于混合CSI联合设计基站与RIS侧的波束赋形，最大化遍历安全速率。

### 3.3 低开销的CSI反馈技术研究

传统信道信息反馈方法包括RVQ反馈方案、基于码本的反馈方案以及基于深度学习的反馈方案等。然而，这些方法应用于RIS时，会导致量化码本的大小随着天线数呈指数

级增长，带来严重的反馈开销问题。因此，需要合理利用不同用户RIS信道的相关性、改进的学习算法、RIS信道定制的能力，来有效降低信道反馈信息的数量。

由于不同用户在共享完全一致的基站-RIS稀疏信道的同时，又有着各自独特的用户-RIS信道，因此最终级联信道矩阵有着数量有限的非零列向量，并且不同用户的非零列索引相同。文献[33]利用该性质挖掘级联信道的单结构稀疏性，将下行CSI反馈信息划分为“用户特定”的信道信息（对应着非零列向量）以及“用户无关”的CSI（对应着非零列索引）。前者经不同用户单独反馈给基站，而后者仅需经某特定用户反馈给基站进而降低下行CSI的开销。相比于传统的基于统计的信道反馈方案，后者可以降低约80%的下行CSI反馈开销。级联信道的三重稀疏结构得以进一步挖掘<sup>[67]</sup>。不同用户共同的CSI信息包括路径角度值、偏移值以及幅度比例，这些信息仅需经部分用户反馈给基站，而其余信息可以经压缩后再反馈。相比于单结构稀疏反馈方案<sup>[33]</sup>，“三重稀疏结构”的CSI方案反馈开销可进一步降低约56.8%。另外，文献[31]提出基于压缩感知的信道反馈方案，采用基于聚类的码本将下行CSI压缩为数量较少的信道向量以减小反馈开销。

然而，实际中并非所有的信道信息都具备完全稀疏性，采用压缩感知技术不能很好地提取出信道状态信息。对此，文献[27]提出名为Quan-Transformer的网络架构以实现CSI压缩与重构。该方案在有效压缩CSI的同时还能大幅提升CSI的恢复精度，相比于传统基于深度学习的信道反馈框架CsiNet能够提升约3.32%的NMSE性能。类似地，文献[68]提出基于数据驱动的深度学习信道获取网络（CAN）。下行导频传输、UE侧的CSI反馈以及基站侧的CSI重建过程被建模为基于变换的端到端神经网络，从而以较低的导频与反馈信令开销获取下行CSI。

除此之外，文献[69]聚焦于路径增益的反馈设计，提出级联码本的方案来量化子信道中的视距路径和非视距路径增益。而面对多RIS协助场景，多RIS的引入将带来更高的信道维数、更复杂的传播环境，从而造成难以忍受的反馈开销。文献[70]基于路径选择以及相移器的设计，将丰富的散射环境定义为稀疏信道，对此仅需反馈主导路径的信道参数。相比于原本的 $3N_{\text{RIS}} + 3 \sum_{k=1}^K (L_{T,k} + L_{k,R})$ ，反馈开销可以显著降低至 $2N_{\text{custom}}$ 。其中， $N_{\text{RIS}}$ 表示RIS的数目， $L_{T,k}$ 和 $L_{k,R}$ 分别表示非视距路径和传播路径数， $N_{\text{custom}}$ 表示信道定制后的路径数。

### 3.4 低开销的相位配置信息反馈技术研究

RIS相位配置信息传输是RIS特有的传输开销。当单元数量大、相位精度高时，该开销将变得难以承受。为减小相位配置信息的反馈开销，文献[71]提出低秩张量建模的方法，例如PARAllel FACTors (PARAFAC) 和 Tucker，将优化得到的RIS相位配置表示为维数更小的预定义因子的Kronecker乘积形式。调整对应预定义因子的数量可以控制反馈的相位配置数。相比于文献[18]，该方案可以在保证可观的SE前提下极大地降低相位配置信息的反馈开销。基于传统的自编码，文献[72]移除批量归一化(BN)层并引入去噪模块，提出相移压缩和去噪网络(PSCDN)方案，将RIS相位信息随机映射到特征空间内维数更小的码字中，从而能够实现有限带宽下的高压缩比反馈。

针对控制链路带宽受限的情形，文献[36]提出新颖的基于自适应码本的有限反馈协议，同时给出两种自适应码本的解决方案，分别是随机邻近性(RA)和基于深度神经网络策略的RIS控制(DPIC)。这两种解决方案仅需要端到端的级联信道，无须对用户位置和入射角进行估计，可显著减小控制开销。

虽然前述方案一定程度上降低了相位配置信息的反馈开销，然而这些方案中RIS的相位配置优化依然依赖于CSI。面对快速变化的信道，频繁的RIS配置带来的巨大开销往往难以承受。对此，文献[73]将RIS相位配置与信道估计分离，当且仅当处于RIS波束覆盖范围内的用户服务质量(QoS)无法满足时才需要重新配置RIS的相位。在这种情况下，RIS相位重新配置的频率不再取决于信道相干时间，而取决于用户的移动速度、RIS波束的覆盖范围以及用户的QoS需求，进而有助于进一步减少开销。

### 3.5 考虑电路开销的RIS通信技术研究

兼顾电路开销的RIS通信技术研究是近年来的研究热点之一。文献[25]在兼顾信道估计和相位配置能耗的基础上，对RIS参数进行优化设计，提出一种针对点对点通信系统的高能效RIS通信方案，可提升系统能效约27.8%。文献[35]则针对多用户通信系统，提出基于Dinkelbach算法的功率分配方案，可将系统能效提升约300%。考虑非完备CSI条件下的上行MIMO通信系统，文献[39]基于闭合表达式推导最优功率分配策略，可实现能效与谱效之间的权衡。此外，文献[22, 38, 74]等还分别针对设备对设备(D2D)系统、多播系统、共生无线电系统提出最优能效方案。

针对多RIS协助无线通信网络，文献[75]联合优化基站侧的有源波束赋形、RIS侧的无源波束赋形以及RIS的相位

偏移，以最小化网络整体的功率消耗，这其中包括RIS节点自身的电路功耗。研究指出，通过动态关闭某些RIS面板，可以实现以非常小的性能损失大幅降低功率消耗。文献[76]则提出一种低复杂度的贪婪搜索算法，通过控制每块RIS面板的开/关状态，以最大化整个系统的能量效率。

当无线信道环境和RIS可用能量模型不确定时，文献[77]基于深度强化学习(DRL)优化发送功率分配以及每个RIS单元的工作状态，以提升系统整体能效。类似地，文献[78]采用DRL的方法求解RIS辅助的车辆自组织网络的能效最大化问题，可实现约87.5%的能效提升。文献[79]则提出一种基于长短时记忆(LSTM)联合资源分配算法，可基于实测数据预测最优RIS配置参数，相比传统的贝叶斯网络方法可提升约20.1%的系统能效。

新型的同时透射与反射型RIS(STAR-RIS)可以进一步增强覆盖能力。针对STAR-RIS辅助的多用户MIMO非正交多址(NOMA)系统，文献[80]基于Dinkelbach方法、对偶分解理论以及连续凸逼近(SCA)技术等，联合优化STAR-RIS的传输与反射波束赋形向量以实现最优的能效性能。文献[81]针对边缘移动计算(MEC)场景，通过联合优化STAR-RIS反射系数、功率分配以及卸载数据的大小，降低整个系统的总能量消耗。文献[82]进一步提出一种基于深度确定性策略梯度(DDPG)算法，通过联合优化基站侧的有源波束赋形和STAR-RIS侧的反射系数来最大化系统能效。

## 4 新研究主题

学术界广泛关注RIS开销问题，已针对典型场景下的基础模型提出多种有效的解决方案。然而，随着6G研究的推进，许多新的应用场景、技术需求陆续出现。如何在更加复杂多样的网络环境中有效发挥RIS的潜在价值，仍需要持续探索。本文从新场景、新型软硬件设计、新型组网架构、新技术交叉赋能等方面展开论述。

### 1) 全空间RIS协助的无线通信

相比于传统的纯反射型半空间RIS，STAR-RIS可实现360°的全空间覆盖<sup>[83]</sup>。然而，STAR-RIS的每个单元需同时满足传输和反射需求，且相应传输和反射相位紧密耦合、相互制约。这对硬件复杂度、电路运行能耗、信道估计开销等均提出新的挑战。针对传统RIS模型开展的性能分析、低开销方案设计等结论，将不能很好地应用于STAR-RIS模型，因此需要开展新的技术方案研究。

### 2) 基于统计/混合CSI的联合波束赋形

基于统计/混合CSI优化设计RIS辅助通信系统的有源波

束赋形和无源波束赋，可降低对信道估计及信息反馈的需求。该方法已被证明可实现通信性能增益与系统开销的折中权衡。现有研究已针对典型的基础通信场景提出了高效的联合波束赋形方案，而针对多点协作 RIS、多跳 RIS、高速移动等复杂场景，如何基于有限的 CSI 实现有效的联合波束设计仍有待探索。

### 3) 低功耗的 RIS 硬件平台

实际中具有多种不同的 RIS 硬件方案，在工作频率、码本切换时间、工作电压、功耗和部署成本等方面存在差异，因此可分别应用于不同的场景。例如，实测数据指出<sup>[84]</sup>，基于 RF-MEMS 和 MOSFET 设计的 RIS 具有较低能耗，但硬件成本较高；基于 PIN 二极管的 RIS 具有更低的硬件成本和更短的码本切换时间，但能耗较大且仅支持 110 GHz 以下的频段。随着电子信息材料科学与工程技术的进步，很多新型 RIS 方案将被提出。其中，低成本、低功耗、广适用范围的 RIS 硬件设计是潜在方向。

### 4) 低开销的 RIS 组网模型

在实际网络中部署 RIS 时需综合考虑用户分布、传播环境、成本控制等诸多因素。一方面，当将相同数量的 RIS 单元分别放置在用户与基站附近时，可实现比集中式部署更好的用户通信速率<sup>[85]</sup>；另一方面，障碍物较多、通信频率较高时，需要多个 RIS 接力建立通信链路<sup>[86]</sup>。此时，多个 RIS 站点之间以及 RIS 与基站之间的相互协调，会显著增加同步挑战和信令开销。因此，如何权衡 RIS 协作组网增益与开销是极具挑战的研究课题。

### 5) AI 赋能的 RIS 通信方法

人工智能（AI）技术已被证明可以有效求解复杂应用中的非线性资源分配问题。因此，AI 赋能的 RIS 通信网络是备受关注的研究方向。AI 模型训练通常由处理能力较强的基站执行，相应的训练数据收集、传递等过程会造成额外通信开销。这可能要求 RIS 节点具备一定的信息感知能力，增加其硬件成本。同时，模型训练过程本身需要耗费庞大的计算时间与能量等<sup>[87]</sup>。联邦学习、迁移学习<sup>[88]</sup>等技术框架有望为低开销 RIS 组网提供新思路。

### 6) 低开销的 RIS 近场通信

为补偿严重的路径衰落带来的系统性能损失，预计未来网络将部署更多的 RIS 反射单元。然而，由于瑞利距离与天线口径的平方成正比，随着 RIS 面板孔径的不断增加，相应的近场通信范围将不断扩大，导致越来越多的用户、散射体处于近场环境中<sup>[89]</sup>。对此，近场环境下的低开销方案设计面临新的挑战。例如，近场环境下，天线阵列导向矢量与散射体的坐标密切相关，对此原先的低开销码本方案将不再适

用<sup>[90]</sup>。更糟糕的是，相比于远场码本，近场码本将包含更多的候选码字，使得近场的波束赋形设计更为复杂，造成更高的导频开销<sup>[91]</sup>。传统远场环境下的低开销设计方案不能很好地解决这些问题，亟须开展新的技术方案研究。

### 7) 低开销的 RIS 协助通信感知系统

相比于传统的通信感知一体化（ISAC）网络，RIS 的引入可以在为处于非视距的感知目标创建虚拟直射链路的同时抑制干扰信号，有助于进一步增强感知性能<sup>[92]</sup>。然而，ISAC 系统设计中需要综合考虑感知和通信的需求，除传统的通信参数例如信道信息的估计之外，还需要实现对感知目标的角度、速度、距离等参数的估计<sup>[93]</sup>，因此将占据更多的系统资源，造成额外的资源开销。除此之外，为满足未来高精度的感知需求，系统中将引入更多的 RIS 以提供更高的设计自由度<sup>[94]</sup>，然而，这些 RIS 之间的协同控制机制以及部署方案设计也将带来额外的资源开销。

### 8) 低能耗的 RIS 协助边缘计算网络

与传统的 MEC 网络相比，RIS 技术的引入，一方面，有助于改善无线链路的质量，提高传输速率，另一方面，有助于显著扩大无线卸载的覆盖范围<sup>[95]</sup>。然而，现有研究大多聚焦于对 RIS 反射系数的优化来提升卸载效率<sup>[96]</sup>，忽视了网络中 RIS 节点的部署成本以及反射单元相位配置过程中产生的能耗开销。相关的研究工作处于起步阶段，有待进一步的探索。

## 5 结束语

RIS 是面向 6G 的创新性高效无线技术，其根本优势在于打破了传统无线信道的不可控局限，通过构建智能可编程的无线环境，有效提升无线网络的覆盖范围、通信效率、安全隐私能力等。然而，RIS 节点的引入也会带来额外开销。在考虑信道估计、无线控制等开销的前提下，如何尽量发挥 RIS 的无线赋能潜力是根本问题之一。

## 参考文献

- [1] BISOYI S, PASUPULETI M, KUCHI K, et al. IMT 2030 candidate technologies: a white paper by IIT hyderabad & wisig networks. [EB/OL]. (2023-07-26)[2023-10-15]. <https://arxiv.org/abs/2307.14978v1>
- [2] LIU R Q, LI R Y N, DI RENZO M, et al. A vision and an evolutionary framework for 6G: scenarios, capabilities and enablers [EB/OL]. (2023-05-23)[2023-10-15]. <https://arxiv.org/abs/2305.13887.pdf>
- [3] SELVA E, GATI A, HAMON M H, et al. Towards a 6G embedding sustainability [EB/OL]. (2023-07-13) [2023-10-15]. <https://arxiv.org/abs/2307.06636.pdf>
- [4] PIOVESAN N, LÓPEZ-PÉREZ D, DE DOMENICO A, et al. Machine learning and analytical power consumption models for 5G base stations [J]. IEEE communications magazine, 2022, 60(10): 56–62. DOI: 10.1109/MCOM.001.2200023

- [5] FAN C, DIAO Z, LIU W. Research on 5G power consumption and communication power supply scheme [J]. Communication world, 2019, 26: 28–30
- [6] ZOU Y, WANG Q G, CHI Y, et al. Electric load profile of 5G base station in distribution systems based on data flow analysis [J]. IEEE transactions on smart grid, 2022, 13(3): 2452–2466. DOI: 10.1109/TSG.2022.3150074
- [7] WANG C X, YOU X H, GAO X Q, et al. On the road to 6G: visions, requirements, key technologies, and testbeds [J]. IEEE communications surveys & tutorials, 2023, 25(2): 905–974. DOI: 10.1109/COMST.2023.3249835
- [8] 孙善球, 林学进, 王昊, 等.“双碳战略”给移动通信天线产业带来的机遇和挑战 [J]. 移动通信, 2022, 46(3): 99–102
- [9] LIU R Q, WU Q Q, DI RENZO M, et al. A path to smart radio environments: an industrial viewpoint on reconfigurable intelligent surfaces [J]. IEEE wireless communications, 2022, 29(1): 202–208. DOI: 10.1109/MWC.111.2100258
- [10] WU Q, ZHANG R. Beamforming optimization for wireless network aided by intelligent reflecting surface with discrete phase shifts [J]. IEEE transactions on communications, 2020, 68(3): 1838–1851. DOI: 10.1109/TCOMM.2019.2958916
- [11] AI Y, DEFIGUIREDO F A P, KONG L, et al. Secure vehicular communications through reconfigurable intelligent surfaces [J]. IEEE transactions on vehicular technology, 2021, 70(7): 7272–7276. DOI: 10.1109/TVT.2021.3088441
- [12] SRRC. NTT DoCoMo, Metawave test 5G mobile system in Tokyo [EB/OL]. (2018–12–06)[2023–10–15]. <http://www.srcc.org.cn/en/news4504.aspx>
- [13] ARUN V, BALAKRISHNAN H. RFocus: practical beamforming for small devices [EB/OL]. (2019–05–13) [2023–10–15]. <https://arxiv.org/abs/1905.05130.pdf>
- [14] DAI L L, WANG B C, WANG M, et al. Reconfigurable intelligent surface-based wireless communications: antenna design, prototyping, and experimental results [J]. IEEE access, 2020, 8: 45913–45923. DOI: 10.1109/ACCESS.2020.2977772
- [15] 中国移动通信有限公司研究院. 中国移动联合崔铁军院士团队完成6G智能超表面技术试验 [EB/OL]. (2021–07–21)[2023–10–15]. [https://www.sohu.com/a/478712728\\_121124379](https://www.sohu.com/a/478712728_121124379)
- [16] C114 通信网. 中兴通讯联合中国电信完成业界首个5G高频外场智能超表面技术验证测试 [EB/OL]. (2021–07–08) [2023–10–15]. <http://www.c114.com/news/127/a1167281.html>
- [17] 崔铁军, 金石, 章嘉懿, 等. 智能超表面技术研究报告 [R]. 2021
- [18] ZAPPONE A, DI RENZO M, SHAMS F, et al. Overhead-aware design of reconfigurable intelligent surfaces in smart radio environments [J]. IEEE transactions on wireless communications, 2021, 20(1): 126–141. DOI: 10.1109/TWC.2020.3023578
- [19] WEI X H, SHEN D C, DAI L L. Channel estimation for RIS assisted wireless communications—part I: fundamentals, solutions, and future opportunities [J]. IEEE communications letters, 2021, 25(5): 1398–1402. DOI: 10.1109/LCOMM.2021.3052822
- [20] WU Q Q, ZHANG S W, ZHENG B X, et al. Intelligent reflecting surface-aided wireless communications: a tutorial [J]. IEEE transactions on communications, 2021, 69(5): 3313–3351. DOI: 10.1109/TCOMM.2021.3051897
- [21] MA Q, BAI G D, JING H B, et al. Smart metasurface with self-adaptively reprogrammable functions [J]. Light: science & applications, 2019, 8(1): 1–12. DOI: 10.1038/s41377-019-0205-3
- [22] JIA S Q, YUAN X J, LIANG Y C. Reconfigurable intelligent surfaces for energy efficiency in D2D communication network [J]. IEEE wireless communications letters, 2021, 10(3): 683–687. DOI: 10.1109/LWC.2020.3046358
- [23] ZHENG B X, ZHANG R. Intelligent reflecting surface-enhanced OFDM: channel estimation and reflection optimization [J]. IEEE wireless communications letters, 2020, 9(4): 518–522. DOI: 10.1109/LWC.2019.2961357
- [24] LIU H, YUAN X J, ZHANG Y J A. Matrix-calibration-based cascaded channel estimation for reconfigurable intelligent surface assisted multiuser MIMO [J]. IEEE journal on selected areas in communications, 2020, 38(11): 2621–2636. DOI: 10.1109/JSAC.2020.3007057
- [25] ZAPPONE A, DI RENZO M, XI X J, et al. On the optimal number of reflecting elements for reconfigurable intelligent surfaces [J]. IEEE wireless communications letters, 2021, 10(3): 464–468. DOI: 10.1109/LWC.2020.3034686
- [26] LEE G, JUNG M, KASGARI A T Z, et al. Deep reinforcement learning for energy-efficient networking with reconfigurable intelligent surfaces [C]// Proceedings of ICC 2020 – 2020 IEEE International Conference on Communications (ICC). IEEE, 2020: 1–6. DOI: 10.1109/ICC40277.2020.9149380
- [27] XIE W W, ZOU J, XIAO J, et al. Quan-transformer based channel feedback for RIS-aided wireless communication systems [J]. IEEE communications letters, 2022, 26(11): 2631–2635. DOI: 10.1109/LCOMM.2022.3198793
- [28] JINDAL N. MIMO broadcast channels with finite-rate feedback [J]. IEEE transactions on information theory, 2006, 52(11): 5045–5060. DOI: 10.1109/TIT.2006.883550
- [29] SHEN W Q, DAI L L, ZHANG Y, et al. On the performance of channel-statistics-based codebook for massive MIMO channel feedback [J]. IEEE transactions on vehicular technology, 2017, 66(8): 7553–7557. DOI: 10.1109/TVT.2017.2656908
- [30] KUO P H, KUNG H T, TING P G. Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays [C]// Proceedings of 2012 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2012: 492–497. DOI: 10.1109/WCNC.2012.6214417
- [31] SHIN B S, OH J H, YOU Y H, et al. Limited channel feedback scheme for reconfigurable intelligent surface assisted MU-MIMO wireless communication systems [J]. IEEE access, 2022, 10: 50288–50297. DOI: 10.1109/ACCESS.2022.3174095
- [32] HU C, DAI L L, HAN S F, et al. Two-timescale channel estimation for reconfigurable intelligent surface aided wireless communications [J]. IEEE transactions on communications, 2021, 69(11): 7736–7747. DOI: 10.1109/TCOMM.2021.3072729
- [33] SHEN D C, DAI L L. Dimension reduced channel feedback for reconfigurable intelligent surface aided wireless communications [J]. IEEE transactions on communications, 2021, 69(11): 7748–7760. DOI: 10.1109/TCOMM.2021.3100428
- [34] CHEN Y B, WANG Y, WANG Z C, et al. Robust beamforming for active reconfigurable intelligent omni-surface in vehicular communications [J]. IEEE journal on selected areas in communications, 2022, 40(10): 3086–3103. DOI: 10.1109/JSAC.2022.3196095
- [35] HUANG C W, ZAPPONE A, ALEXANDROPOULOS G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communication [J]. IEEE transactions on wireless communications, 2019, 18(8): 4157–4170. DOI: 10.1109/TWC.2019.2922609
- [36] KIM J, HOSSEINALIPOUR S, MARCUM A C, et al. Learning-based adaptive IRS control with limited feedback codebooks [J]. IEEE transactions on wireless communications, 2022, 21(11): 9566–9581. DOI: 10.1109/TWC.2022.3178055
- [37] ZHONG R K, LIU Y W, MU X D, et al. AI empowered RIS-assisted NOMA networks: deep learning or reinforcement learning? [J]. IEEE journal on selected areas in communications, 2022, 40(1): 182–196. DOI: 10.1109/JSAC.2021.3126068
- [38] DU L S, ZHANG W, MA J H, et al. Reconfigurable intelligent surfaces for energy efficiency in multicast transmissions [J]. IEEE transactions on vehicular technology, 2021, 70(6): 6266–6271. DOI: 10.1109/TVT.2021.3080302
- [39] YOU L, XIONG J Y, NG D W K, et al. Energy efficiency and spectral efficiency tradeoff in RIS-aided multiuser MIMO uplink transmission [J]. IEEE transactions on signal processing, 2021, 69: 1407–1421. DOI: 10.1109/TSP.2020.3047474
- [40] WANG J H, TANG W K, JIN S, et al. Static power consumption modeling and measurement of reconfigurable intelligent surfaces [EB/OL]. (2023–03–01)[2023–10–15]. <https://arxiv.org/abs/2303.00299.pdf>
- [41] WANG J H, TANG W K, LIANG J C, et al. Reconfigurable intelligent surface: power consumption modeling and practical measurement validation (2022–11–01)[2023–10–15]. <https://arxiv.org/abs/2211.00323.pdf>
- [42] KIM J, YU H, KANG X, et al. Discrete phase shifts of intelligent reflecting surface systems considering network overhead [J]. Entropy, 2022, 24(12): 1753. DOI: 10.3390/e24121753
- [43] XIE Z Y, YI W Q, WU X L, et al. Downlink multi-RIS aided transmission in backhaul limited networks [J]. IEEE wireless communications letters, 2022, 11(7): 1458–1462. DOI: 10.1109/LWC.2022.3174448
- [44] JIN Y, ZHANG J Y, ZHANG X D, et al. Channel estimation for semi-passive reconfigurable intelligent surfaces with enhanced deep residual networks [J]. IEEE transactions on vehicular technology, 2021, 70(10): 11083–11088. DOI: 10.1109/TVT.2021.3109937

- [45] ELBIR A M, PAPAZAFEIROPOULOS A, KOURTESSIS P, et al. Deep channel learning for large intelligent surfaces aided mm-wave massive MIMO systems [J]. IEEE wireless communications letters, 2020, 9(9): 1447–1451. DOI: 10.1109/LWC.2020.2993699
- [46] JENSEN T L, DE CARVALHO E. An optimal channel estimation scheme for intelligent reflecting surfaces based on a minimum variance unbiased estimator [C]//Proceedings of ICASSP 2020 – 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020: 5000–5004. DOI: 10.1109/ICASSP40776.2020.9053695
- [47] YANG Y F, ZHENG B X, ZHANG S W, et al. Intelligent reflecting surface meets OFDM: protocol design and rate maximization [J]. IEEE transactions on communications, 2020, 68(7): 4522–4535. DOI: 10.1109/TCOMM.2020.2981458
- [48] WANG P L, FANG J, DUAN H P, et al. Compressed channel estimation for intelligent reflecting surface-assisted millimeter wave systems [J]. IEEE signal processing letters, 2020, 27: 905–909. DOI: 10.1109/LSP.2020.2998357
- [49] MAO Z D, LIU X Q, PENG M G. Channel estimation for intelligent reflecting surface assisted massive MIMO systems—a deep learning approach [J]. IEEE communications letters, 2022, 26(4): 798–802. DOI: 10.1109/LCOMM.2022.3147200
- [50] CHUNG H, KIM S. Efficient two-stage beam training and channel estimation for ris-aided mmwave systems via fast alternating least squares [C]//Proceedings of ICASSP 2022 – 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2022: 5188–5192. DOI: 10.1109/ICASSP43922.2022.9746094
- [51] GAN X, ZHONG C J, HUANG C W, et al. Multiple RISs assisted cell-free networks with two-timescale CSI: performance analysis and system design [J]. IEEE transactions on communications, 2022, 70(11): 7696–7710. DOI: 10.1109/TCOMM.2022.3208629
- [52] YANG S J, LYU W T, XIU Y, et al. Active 3D double-RIS-aided multi-user communications: two-timescale-based separate channel estimation via Bayesian learning [J]. IEEE transactions on communications, 2023, 71(6): 3605–3620. DOI: 10.1109/TCOMM.2023.3265115
- [53] WANG Z R, LIU L, CUI S G. Channel estimation for intelligent reflecting surface assisted multiuser communications: framework, algorithms, and analysis [J]. IEEE transactions on wireless communications, 2020, 19(10): 6607–6620. DOI: 10.1109/TWC.2020.3004330
- [54] WEI Y, ZHAO M M, ZHAO M J, et al. Channel estimation for IRS-aided multiuser communications with reduced error propagation [J]. IEEE transactions on wireless communications, 2022, 21(4): 2725–2741. DOI: 10.1109/TWC.2021.3115161
- [55] WEI X H, SHEN D C, DAI L L. Channel estimation for RIS assisted wireless communications—part II: an improved solution based on double-structured sparsity [J]. IEEE communications letters, 2021, 25(5): 1403–1407. DOI: 10.1109/LCOMM.2021.3052787
- [56] XU M, ZHANG S, MA J P, et al. Deep learning-based time-varying channel estimation for RIS assisted communication [J]. IEEE communications letters, 2022, 26(1): 94–98. DOI: 10.1109/LCOMM.2021.3127160
- [57] HE J G, WYMEERSCH H, DI RENZO M, et al. Learning to estimate RIS-aided mmWave channels [J]. IEEE wireless communications letters, 2022, 11(4): 841–845. DOI: 10.1109/LWC.2022.3147250
- [58] AN J C, XU C, WU Q Q, et al. Codebook-based solutions for reconfigurable intelligent surfaces and their open challenges [J]. IEEE wireless communications, 2022, 29(9): 1–8. DOI: 10.1109/MWC.010.2200312
- [59] JIA X, AN J C, LIU H, et al. Environment-aware codebook for reconfigurable intelligent surface-aided MISO communications [J]. IEEE wireless communications letters, 2023, 12(7): 1174–1178. DOI: 10.1109/LWC.2023.3265399
- [60] TANG W K, LI X, DAI J Y, et al. Wireless communications with programmable metasurface: transceiver design and experimental results [J]. China communications, 2019, 16(5): 46–61. DOI: 10.23919/jcc.2019.05.004
- [61] ZHENG B X, YOU C S, MEI W D, et al. A survey on channel estimation and practical passive beamforming design for intelligent reflecting surface aided wireless communications [J]. IEEE communications surveys & tutorials, 2022, 24(2): 1035–1071. DOI: 10.1109/COMST.2022.3155305
- [62] HU X L, WANG J W, ZHONG C J. Statistical CSI based design for intelligent reflecting surface assisted MISO systems [J]. Science China information sciences, 2020, 63(12): 222303. DOI: 10.1007/s11432-020-3033-3
- [63] JIA Y H, YE C C, CUI Y. Analysis and optimization of an intelligent reflecting surface-assisted system with interference [J]. IEEE transactions on wireless communications, 2020, 19(12): 8068–8082. DOI: 10.1109/TWC.2020.3019088
- [64] LIU J, ZHANG J, ZHANG Q, et al. Secrecy rate analysis for reconfigurable intelligent surface-assisted mimo communications with statistical CSI [J]. China communications, 2021, 18(3): 52–62. DOI: 10.23919/JCC.2021.03.005
- [65] BIE Q Y, LIU Y, WANG Y X, et al. Deployment optimization of reconfigurable intelligent surface for relay systems [J]. IEEE transactions on green communications and networking, 2022, 6(1): 221–233. DOI: 10.1109/TGCN.2022.3145026
- [66] LIU C, TIAN C, LIU P X. RIS-assisted secure transmission exploiting statistical CSI of eavesdropper [C]//Proceedings of 2021 IEEE Global Communications Conference (GLOBECOM). IEEE, 2022: 1–6. DOI: 10.1109/GLOBECOM46510.2021.9685746
- [67] SHI X, WANG J T, SONG J. Triple-structured sparsity-based channel feedback for RIS-assisted MU-MIMO system [J]. IEEE communications letters, 2022, 26(5): 1141–1145. DOI: 10.1109/LCOMM.2022.3147220
- [68] WU M H, GAO Z, HUANG Y, et al. Deep learning-based rate-splitting multiple access for reconfigurable intelligent surface-aided tera-hertz massive MIMO [J]. IEEE journal on selected areas in communications, 2023, 41(5): 1431–1451. DOI: 10.1109/JSAC.2023.3240781
- [69] CHEN W C, WEN C K, LI X, et al. Adaptive bit partitioning for reconfigurable intelligent surface assisted FDD systems with limited feedback [J]. IEEE transactions on wireless communications, 2022, 21(4): 2488–2505. DOI: 10.1109/TWC.2021.3112819
- [70] CHEN W C, WEN C K, LI X, et al. Channel customization for limited feedback in RIS-assisted FDD systems [J]. IEEE transactions on wireless communications, 2023, 22(7): 4505–4519. DOI: 10.1109/TWC.2022.3226442
- [71] SOKAL B, GOMES P R B, DE ALMEIDA A L F, et al. Reducing the control overhead of intelligent reconfigurable surfaces via a tensor-based low-rank factorization approach [J]. IEEE transactions on wireless communications, 2023, 22(10): 6578–6593. DOI: 10.1109/TWC.2023.3244487
- [72] YU X H, LI D, XU Y J, et al. Convolutional autoencoder-based phase shift feedback compression for intelligent reflecting surface-assisted wireless systems [J]. IEEE communications letters, 2022, 26(1): 89–93. DOI: 10.1109/LCOMM.2021.3123941
- [73] JAMALI V, ALEXANDROPOULOS G C, SCHOBER R, et al. Low-to-zero-overhead IRS reconfiguration: decoupling illumination and channel estimation [J]. IEEE communications letters, 2022, 26(4): 932–936. DOI: 10.1109/LCOMM.2022.3141206
- [74] LI J, LI X, BI Y J, et al. Energy-efficient joint resource allocation with reconfigurable intelligent surfaces in symbiotic radio networks [J]. IEEE transactions on cognitive communications and networking, 2022, 8(4): 1816–1827. DOI: 10.1109/TCNN.2022.3188986
- [75] SUN S Y, FU M, SHI Y M, et al. Towards reconfigurable intelligent surfaces powered green wireless networks [C]//Proceedings of 2020 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2020: 1–6. DOI: 10.1109/WCNC45663.2020.9120857
- [76] YANG Z H, CHEN M Z, SAAD W, et al. Energy-efficient wireless communications with distributed reconfigurable intelligent surfaces [J]. IEEE transactions on wireless communications, 2022, 21(1): 665–679. DOI: 10.1109/TWC.2021.3098632
- [77] YANG B, CAO X L, HUANG C W, et al. Intelligent spectrum learning for wireless networks with reconfigurable intelligent surfaces [J]. IEEE transactions on vehicular technology, 2021, 70(4): 3920–3925. DOI: 10.1109/TVT.2021.3064042
- [78] PAN Q Q, WU J, NEBHEN J, et al. Artificial intelligence-based energy efficient communication system for intelligent reflecting surface-driven VANETs [J]. IEEE transactions on intelligent transportation systems, 2022, 23(10): 19714–19726. DOI: 10.1109/TITS.2022.3152677
- [79] GUPTA K D, NIGAM R, SHARMA D K, et al. LSTM-based energy-efficient wireless communication with reconfigurable intelligent surfaces [J]. IEEE transactions on green communications and networking, 2022, 6(2): 704–712. DOI: 10.1109/TGCN.2021.3135437
- [80] GUO Y, FANG F, CAI D H, et al. Energy-efficient design for a NOMA assisted STAR-RIS network with deep reinforcement learning [J]. IEEE transactions on vehicular technology, 2023, 72(4): 5424–5428. DOI:

- 10.1109/TVT.2022.3224926
- [81] ZHANG Q, WANG Y H, LI H, et al. Resource allocation for energy efficient STAR-RIS aided MEC systems [J]. IEEE wireless communications letters, 2023, 12(4): 610–614. DOI: 10.1109/LWC.2023.3236411
- [82] FANG F, WU B B, FU S, et al. Energy-efficient design of STAR-RIS aided MIMO-NOMA networks [J]. IEEE transactions on communications, 2023, 71(1): 498–511. DOI: 10.1109/TCOMM.2022.3223706
- [83] LIU Y W, MU X D, XU J Q, et al. STAR: simultaneous transmission and reflection for 360° coverage by intelligent surfaces [J]. IEEE wireless communications, 2021, 28(6): 102–109. DOI: 10.1109/MWC.001.2100191
- [84] LIU R Q, DOU J W, LI P, et al. Simulation and field trial results of reconfigurable intelligent surfaces in 5G networks [J]. IEEE access, 2022, 10: 122786–122795. DOI: 10.1109/ACCESS.2022.3223447
- [85] ZHANG S W, ZHANG R. Intelligent reflecting surface aided multi-user communication: capacity region and deployment strategy [J]. IEEE transactions on communications, 2021, 69(9): 5790–5806. DOI: 10.1109/TCOMM.2021.3079128
- [86] YOU C S, ZHENG B X, ZHANG R. Wireless communication via double IRS: channel estimation and passive beamforming designs [J]. IEEE wireless communications letters, 2021, 10(2): 431–435. DOI: 10.1109/LWC.2020.3034388
- [87] WANG J H, TANG W K, HAN Y, et al. Interplay between RIS and AI in wireless communications: fundamentals, architectures, applications, and open research problems [J]. IEEE journal on selected areas in communications, 2021, 39(8): 2271–2288. DOI: 10.1109/JSAC.2021.3087259
- [88] ZHANG S, LI M Y, JIAN M N, et al. AIRIS: Artificial intelligence enhanced signal processing in reconfigurable intelligent surface communications [J]. China communications, 2021, 18(7): 158–171. DOI: 10.23919/JCC.2021.07.013
- [89] LIU W, PAN C H, REN H, et al. Low-overhead beam training scheme for extremely large-scale RIS in near field [J]. IEEE transactions on communications, 2023, 71(8): 4924–4940. DOI: 10.1109/TCOMM.2023.3278728
- [90] JIANG Y H, GAO F F, JIAN M N, et al. Reconfigurable intelligent surface for near field communications: beamforming and sensing [J]. IEEE transactions on wireless communications, 2023, 22(5): 3447–3459. DOI: 10.1109/TWC.2022.3218531
- [91] CUI M Y, WU Z D, LU Y, et al. Near-field MIMO communications for 6G: fundamentals, challenges, potentials, and future directions [J]. IEEE communications magazine, 2023, 61(1): 40–46. DOI: 10.1109/MCOM.004.2200136
- [92] 杨晓宇, 尉志青, 孟春伟. 可重构智能表面辅助的通信感知一体化系统 [J]. 中兴通讯技术, 2022, 28(5): 17–22
- [93] RIHAN M, ZAPPONE A, BUZZI S, et al. Passive vs. active reconfigurable intelligent surfaces for integrated sensing and communication: challenges and opportunities [J]. IEEE network, 2023, 37(9): 1. DOI: 10.1109/MNET.2023.3321542
- [94] LIU R, LI M, LUO H H, et al. Integrated sensing and communication with reconfigurable intelligent surfaces: opportunities, applications, and future directions [J]. IEEE wireless communications, 2023, 30(1): 50–57. DOI: 10.1109/MWC.002.2200206
- [95] 李斌, 刘文帅, 谢万城, 等. 智能超表面赋能移动边缘计算部分任务卸载策略 [J]. 电子与信息学报, 2022, 44(7): 2309–2316
- [96] BAI T, PAN C H, HAN C, et al. Reconfigurable intelligent surface aided mobile edge computing [J]. IEEE wireless communications, 2021, 28(6): 80–86. DOI: 10.1109/MWC.001.2100142

## 作者简介



陈天贝, 北京邮电大学信息与通信工程学院在读硕士研究生; 主要研究领域为智能反射面技术。



李娜, 北京邮电大学副教授、博士生导师, 主要研究领域为5G/6G无线组网关键技术、无线内生安全、物理层安全; 先后主持和参与国家重点研发计划、国家自然科学基金、北京市自然科学基金等国家级省部级项目10余项, 曾获1项科研成果奖励; 已发表论文40余篇, 获授权专利10余项, 出版专著1部。



陶小峰, 北京邮电大学教授、博士生导师, 移动互联网安全技术国家工程研究中心主任, 无线新技术研究所副所长, 英国工程技术学会会士, IEEE通信学会北京分会主席, 中国科协联合国信息咨询专业委员会委员; 主要研究方向为5G/6G; 曾主持国家杰出青年科学基金、国家“863”计划重大项目、北京市自然科学基金等项目, 被评为北京市科技新星、科技北京百名领军人才、“万人计划”科技创新领军人才, 曾2次获得国家技术发明奖; 已发表论文200余篇, 出版学术专著3部。



# 基于视距概率模型的6G室内 无线通信系统性能分析

## LoS Probability Model Based Performance Analysis for 6G Indoor Wireless Communication Systems

毛曦晨/MAO Xichen<sup>1,2</sup>, 王承祥/WANG Chengxiang<sup>1,2</sup>,  
杨松江/YANG Songjiang<sup>2</sup>, 黄杰/HUANG Jie<sup>1,2</sup>,  
黄晨/HUANG Chen<sup>1,2</sup>

(1. 东南大学 移动通信全国重点实验室, 中国 南京 210096;  
2. 网络通信与安全紫金山实验室, 中国 南京 211111)  
(1. National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China;  
2. Purple Mountain Laboratory, Nanjing 211111, China)

DOI: 10.12142/ZTETJ.202306007

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231212.1655.010.html>

网络出版日期: 2023-12-13

收稿日期: 2023-10-15

**摘要:** 工业物联网、大型商超、体育场馆等室内通信场景的重要性在6G无线通信系统中日益显著。研究室内视距(LoS)信号概率模型有利于提升室内通信系统性能分析准确性。针对室内环境, 考虑家具、设备等室内阻挡物的任意位置以及任意高度, 构建了视距信号概率模型。该模型可以准确刻画不同阻挡物场景下的视距概率。基于视距信号概率模型, 推导了信噪比分布, 并研究了室内覆盖性能和能量效率。相关研究成果有助于指导室内无线通信系统部署, 降低成本。

**关键词:** 室内无线通信; 视距概率; 覆盖概率; 能量效率; 绿色通信

**Abstract:** The importance of indoor communication scenarios, such as industrial Internet of things, superstores, and stadiums, is becoming increasingly significant in the sixth generation (6G) wireless communication systems. The investigation into indoor line-of-sight (LoS) signal probability is conducive to improving the accuracy of indoor communication performance analysis. In this paper, a LoS probability model is constructed for indoor environments, considering any position and any height of blockages. This model can accurately characterize the LoS probability in different scenarios. Based on the model, this paper derives the signal-to-noise ratio (SNR) distribution and investigates the coverage performance and energy efficiency. The research results are instrumental in instructing the deployment of indoor wireless systems and reducing costs.

**Keywords:** indoor wireless communication; LoS probability; coverage probability; energy efficiency; green communication

**引用格式:** 毛曦晨, 王承祥, 杨松江, 等. 基于视距概率模型的6G室内无线通信系统性能分析 [J]. 中兴通讯技术, 2023, 29(6): 39–45. DOI: 10.12142/ZTETJ.202306007

**Citation:** MAO X C, WANG C X, YANG S J, et al. LoS probability model based performance analysis for 6G indoor wireless communication systems [J]. ZTE technology journal, 2023, 29(6): 39–45. DOI: 10.12142/ZTETJ.202306007

随着城市居民密度的不断增加, 室内通信场景逐渐成为移动通信的重要场景, 工业物联网、大型商超等场景下的通信覆盖需求日益提升。中国联通在2019年发布白皮书《中国联通5G数字化室分技术白皮书》<sup>[1]</sup>, 预测室内流量将达到总体流量的80%以上。随着室内用户数量激增、传

**基金项目:** 国家重点研发计划项目(2021YFB2900300); 国家自然科学基金项目(61960206006、62271147); 中央高校基本科研业务费(2242022K60006、2242023K5003); 江苏省重点研发计划项目(BE2022067、BE2022067-1); 江苏省双创博士项目(JSSCBS20210082); 欧盟H2020 RISE TESTBED2项目(872172); 东南大学启动研究基金项目(RF1028623029)

输数据速率要求提升, 室内通信场景在未来通信系统的重要地位日益突出。全球学者均已开展针对6G的研究。国际电信联盟(ITU)在2023年6月发布了6G愿景建议书<sup>[2]</sup>, 强调室内通信在未来6G应用场景的重要地位, 其中海量通信、超高可靠低时延通信等均涉及室内用户。此外, 文献[3]和[4]同样阐述了室内通信系统在6G系统中的重要地位, 并提出光无线技术、超大规模天线阵列以及智能反射面等使能技术, 以提升室内通信性能。

室内通信场景的重要性逐渐凸显, 对此全球学者广泛开

展针对室内通信性能的研究。文献[5]研究室内覆盖概率模型，并推导信干噪比分布，但该模型仅考虑了视距（LoS）信号，未考虑遮挡效应和视距信号概率。文献[6]针对室内覆盖性能及能效提升，研究室内通信系统的优化问题，但在模型中同样未考虑视距概率模型的影响。文献[7]推导了用户与接入点的距离概率分布，并分析覆盖概率，但仍忽视了视距概率对于覆盖性能的影响。然而，视距概率模型是准确分析室内通信系统性能的重要基础。第3代合作伙伴计划（3GPP）在标准文档中给出基础视距概率模型<sup>[8]</sup>，使用指数函数表征视距概率与距离关系。文献[9]研究室内系统的干扰以及覆盖性能，同样直接采用指数函数表征视距概率。文献[10]和[11]研究智能反射面协助下的室外到室内系统性能，由于所考虑的室内传播场景较为简单，因此同样使用指数函数来表征视距概率。指数函数表征视距概率仅适用于较简单的室内场景。一旦考虑家具、设备、墙壁等阻挡物，指数函数将不再适用于室内视距概率分析。因此，一些研究已开展视距概率模型研究。文献[12]基于视距概率研究服务质量概率，但其模型中未考虑不同类型阻挡物的影响。此外，文献[13]和[14]研究了三维室内视距概率，但仍只考虑墙壁阻挡物的影响，忽视了其他类型阻挡物对于视距概率的影响。

综上，室内无线通信性能研究已广泛开展，但仍面临一些挑战：首先，现有性能分析研究忽略了视距概率模型的重要性，或单纯采用3GPP标准视距概率模型；其次，研究室内通信性能时，通信阻挡物考虑得较简单，未考虑阻挡物的高度等因素的影响。为弥补上述不足，本文开展了视距概率模型研究工作，并基于视距概率模型研究室内通信系统性能。本文的主要贡献如下：

- 1) 考虑除墙壁阻挡以外，场景中不同高度的阻挡物对信号传播的影响，为工业物联网、大型商超等场景的视距概率模型推导提供参考；
- 2) 构建适用于工业物联网、大型商超等6G场景的视距概率模型，该模型可以准确刻画不同阻挡物场景下的视距概率，并且考虑了基站位置、用户位置、基站数量等因素的影响；
- 3) 基于提出的视距概率模型和6G普适信道模型，研究了路径损耗、覆盖概率以及能量效率，总结分析室内无线通信性能主要特点，指导室内通信性能优化方案。

## 1 室内视距概率模型

### 1.1 室内通信场景

室内通信相较于室外通信，场景特征显著。首先，室内

通信的用户密度较高，且用户活动范围集中在固定区域。其次，室内传播阻挡物较为复杂，但阻挡物特征具有趋同性，例如：在工业物联网中，传播阻挡物多为大型生产设备，其形状、位置受到厂房设计和生产线限制，往往同区域阻挡物为同类型设备<sup>[14]</sup>。最后，室内通信信号传播场景较为固定，且一般具有各向同性的特征<sup>[13]</sup>。由于室内通信场景中阻挡物具有趋同性特征，因此在对室内通信场景建模时无须特别区分信号传播方向，可以认为其信号传播特性具有各向同性，仅受到阻挡物的影响。受文献[15]和[16]启发，本文归纳总结室内场景特征，将其建模为二维（2D）场景。具体场景如图1所示。

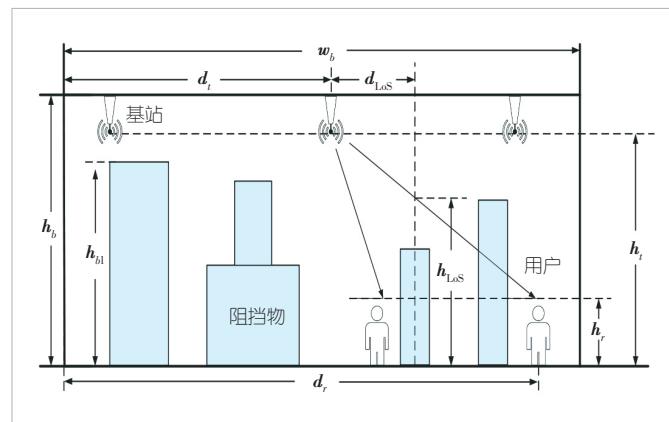
图1中 $h_b$ 和 $w_b$ 分别表示为房间的高度和最长直线传播距离。房间部署 $N$ 个基站，基站为同源基站且均部署在天花板上，基站间信号互无干扰。 $d_i$ 表示基站距离房间边缘的距离， $d_r$ 表示用户距离房间边缘的距离。 $h_t$ 和 $h_r$ 分别表示基站和用户的离地高度。基站端配备了 $N_T$ 根天线，用户端配备了 $N_R$ 根天线。 $h_{bl}$ 表示阻挡物的高度。假设在信号由某一基站向用户传播过程中，视距信号将经过 $N_{bl}$ 个阻挡物。 $d_{LoS}$ 表示信号传播途中障碍物距基站的距离。 $h_{LoS}$ 表示视距信号在障碍处的高度。

### 1.2 室内视距概率模型

文献[17]认为计算视距概率应基于传播过程阻挡物遮挡信号的概率。考虑室内空间具有有限性，且室内阻挡物往往相对固定，本文建立的视距概率模型如下。

如室内场景描述，对于第*i*个基站向用户传输信号过程中，信号将经过 $N_{bl}$ 个阻挡物，因此对于该信号来说，其视距概率为：

$$P_{LoS,i} = \prod_{i=1}^{N_{bl}} P\{h_{bl_i} < h_{LoS}\}, \quad (1)$$



▲图1 室内通信场景模型

即所有阻挡物均低于视距信号高度的概率。对于阻挡物处视距信号高度，可通过下式计算获得：

$$h_{\text{LoS}} = h_t - d_{\text{LoS}} \frac{|h_t - h_r|}{|d_t - d_r|} \quad (2)$$

考虑室内场景中阻挡物多为同类物体，例如办公室场景中的阻挡物为办公桌，大型商场的为货架，工厂的为设备。因此，阻挡物的高度、大小具有规律性，大部分集中在较小区间，少量阻挡物具有较高或较低高度。阻挡物高度可认为服从正态分布规律，即  $h_{bl} \sim N(\mu, \sigma^2)$ ， $\mu$  和  $\sigma^2$  分别表示阻挡物高度的均值和方差。其概率分布函数为

$$f(h_{bl}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(h_{bl} - \mu)^2}{2\sigma^2}\right) \quad (3)$$

因此，对于第  $i$  个基站来说，视距信号概率为

$$P_{\text{LoS}, i} = \prod_{i=1}^{N_{\text{bs}}} \int_0^{h_{\text{LoS}}} f(h_{bl}) dh_b = \prod_{i=1}^{N_{\text{bs}}} 0.5 \times \left[ \operatorname{erf}\left(\frac{h_{\text{LoS}} - \mu}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(-\frac{\mu}{\sigma\sqrt{2}}\right) \right] \quad (4)$$

由于室内场景部署了  $N$  个基站，因此对于每一个基站来说，其视距概率均如公式 (4) 所示，因此室内通信系统的总体视距概率为

$$P_{\text{LoS}} = 1 - \prod_{i=1}^N \left(1 - P_{\text{LoS}, i}\right) = 1 - \prod_{i=1}^N \left(1 - \prod_{i=1}^{N_{\text{bs}}} 0.5 \times \left[ \operatorname{erf}\left(\frac{h_{\text{LoS}} - \mu}{\sigma\sqrt{2}}\right) - \operatorname{erf}\left(-\frac{\mu}{\sigma\sqrt{2}}\right) \right]\right) \quad (5)$$

非视距概率为：

$$P_{\text{NLoS}} = 1 - P_{\text{LoS}} = \prod_{i=1}^N \left(1 - P_{\text{LoS}, i}\right) \quad (6)$$

## 2 室内无线通信性能分析

### 2.1 信道模型

#### 1) 路径损耗

本文采用3GPP TR 38.901标准文档中大规模室内场景的路径损耗公式。

$$PL_{\text{LoS}} = 32.4 + 17.3 \log_{10}(d) + 20 \log_{10}(f_c), \quad (7)$$

$$PL_{\text{NLoS}} = 32.4 + 31.9 \log_{10}(d) + 20 \log_{10}(f_c), \quad (8)$$

其中， $PL_{\text{LoS}}$  为视距信号路径损耗， $PL_{\text{NLoS}}$  为非视距信号路径损耗， $d$  为用户与基站之间的传播距离，即  $d = \sqrt{(h_t - h_r)^2 + (d_t - d_r)^2}$ ，单位为 m； $f_c$  为信号频率，单位为吉比特。

基于提出的视距概率模型，对于用户，平均路径损耗为：

$$\Lambda(h_r, d_r) = P_{\text{LoS}} PL_{\text{LoS}} + P_{\text{NLoS}} PL_{\text{NLoS}} \quad (9)$$

路径损耗取决于用户所处的位置。

#### 2) 小尺度信道模型

由于室内阻挡物分布极为密集，因此小尺度特性对于信号传播影响较大。文献[18]提出了6G普适几何随机信道模型，该模型可以较好地模拟真实室内信道环境。本文基于6G普适信道模型，结合室内通信场景进行必要简化。小尺度信道模型表达式如公式 (10) — (12) 所示。

$$h_{qp}(\tau) = \sqrt{\frac{K}{K+1}} h_{qp}^{\text{LoS}}(\tau) + \sqrt{\frac{1}{K+1}} h_{qp}^{\text{NLoS}}(\tau), \quad (10)$$

$$h_{qp}^{\text{LoS}}(\tau) = \begin{bmatrix} e^{j\theta_L^{VV}} & 0 \\ 0 & e^{j\theta_L^{HH}} \end{bmatrix} e^{j2\pi f_c \tau_{qp}} \delta(\tau - \tau_{qp}), \quad (11)$$

$$h_{qp}^{\text{NLoS}}(\tau) = \sum_{n=1}^{N_{qp}} \sum_{m=1}^{N_{bl}} \begin{bmatrix} e^{j\theta_{mn}^{VV}} & \sqrt{\mu\kappa^{-1}} e^{j\theta_{mn}^{HH}} \\ \sqrt{\kappa^{-1}} e^{j\theta_{mn}^{VV}} & \sqrt{\mu} e^{j\theta_{mn}^{HH}} \end{bmatrix} \sqrt{P_{qp,mn}} e^{j2\pi f_c \tau_{qp,mn}} \delta(\tau - \tau_{qp,mn}), \quad (12)$$

其中， $K$  为莱斯因子，由于其表征视距信号分量的占比，本文中  $K$  由视距概率代替。 $\theta_L^{VV}$ 、 $\theta_L^{HH}$ 、 $\theta_{mn}^{VV}$ 、 $\theta_{mn}^{VH}$ 、 $\theta_{mn}^{HV}$ 、 $\theta_{mn}^{HH}$  为随机相位， $\tau_{qp,mn}$  为多径信号时延， $P_{qp,mn}$  为信号功率， $\mu$  为共极性不平衡， $\kappa$  为交叉极化功率比， $f_c$  为载波频率。因此小尺度信道矩阵为：

$$\mathbf{H} = \begin{bmatrix} h_{11}(\tau) & \cdots & h_{1N_R}(\tau) \\ \vdots & \ddots & \vdots \\ h_{N_T 1}(\tau) & \cdots & h_{N_T N_R}(\tau) \end{bmatrix} \quad (13)$$

#### 2.2 覆盖概率

覆盖概率是衡量系统通信性能的重要指标，可以有效评估系统基站部署的有效性。一般用信噪比(SNR)来衡量覆盖概率，即当该位置的SNR超过设定阈值时，则认为该位置已被覆盖。

$$P_c = \text{Probability}\{\text{SNR} \geq \text{SNR}_{th}\}. \quad (14)$$

对于某一位置的用户来说，考虑路径损耗以及小尺度衰落等特性，其接收到的SNR为：

$$\text{SNR}(h_r, d_r) = \frac{|\mathbf{H}|^2 P_s}{\Lambda(h_r, d_r) P_n}, \quad (15)$$

其中， $P_s$ 代表信号发射功率， $P_n$ 表示噪声功率。SNR取决于用户所处位置。因此覆盖概率为：

$$P_c = \frac{\sum (h_{ri,\max} - h_{ri,\min})(d_{ri,\max} - d_{ri,\min})}{(h_{r,\max} - h_{r,\min})(d_{r,\max} - d_{r,\min})}, \quad (16)$$

其中， $h_{\max/\min}$ 表示用户的最高和最低高度， $d_{\max/\min}$ 表示用户活动的最大和最小位置范围。可覆盖区域用 $\vartheta$ 表示。

### 2.3 能量效率

能量效率（EE）是指每焦耳能量可传输的数据数量。随着通信系统能耗的日益上涨，EE逐渐成为一个重要的性能指标<sup>[1,3]</sup>。

基于公式（15）可得到系统平均SNR：

$$\text{SNR} = \mathbb{E}\left[\frac{|h_{qp}|^2 P_s}{\Lambda(h_r, d_r) P_n}\right], (h_r, d_r) \in \vartheta. \quad (17)$$

因此，系统能效为：

$$\text{EE} = \frac{\sum_{q=1, p=1}^{N_r, N_t} B \log_2(1 + \text{SNR})}{\text{Power}}, \quad (18)$$

其中， $B$ 表示系统带宽，Power为室内通信系统功耗。系统功耗采用文献[19]功耗模型，主要包含射频功耗、基带功耗、功率放大器功耗等功耗。

## 3 仿真结果分析

本节将分析提出的视距概率模型以及相应的仿真结果，并且分析不同基站部署下的系统性能。

### 3.1 仿真参数设置

本小节将总结仿真参数设置，后续如无特殊说明，则仿真参数依照本小节设置。

#### 1) 尺寸设置

考虑工业物联网、大型商超、办公室等场景，本文设置室内最长传播距离为100 m，高度为6 m。假设基站的高度为6 m，基站间距离为等间隔距离。此外，由于室内接收端多为手机、电脑以及工业自动化设备，因此接收端高度在同

一场景下可默认为相同高度。本文设置接收端高度为1.5 m。

#### 2) 收发端设置

根据文献[1]，5G时代室内微基站的单通道发射功率大于250 mW。250 mW、400 mW以及500 mW是常见的发射功率参数。因此本文中我们选择400 mW/每通道作为信号发射功率。设定发射端天线数为128，接收端天线数为4。

### 3.2 视距概率模型

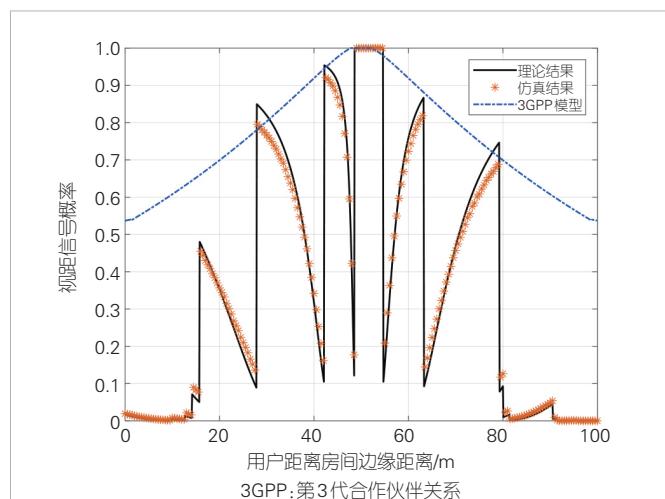
大多室内性能研究直接采用文献[8]中模型，其表达式为：

$$P_{\text{LoS}} = \begin{cases} 1 & , d \leq 5 \text{ m} \\ \exp\left(-\frac{d-5}{70.8}\right) & , 5 \text{ m} < d \leq 49 \text{ m} \\ \exp\left(-\frac{d-49}{211.7}\right) \times 0.54 & , 49 \text{ m} < d \end{cases} . \quad (19)$$

在公式（19）的视距概率所依据的场景模型中，基站分布为等间隔分布，且未考虑复杂阻挡物环境。本小节将对比该模型与本文提出的模型在复杂室内传播环境下的准确性。

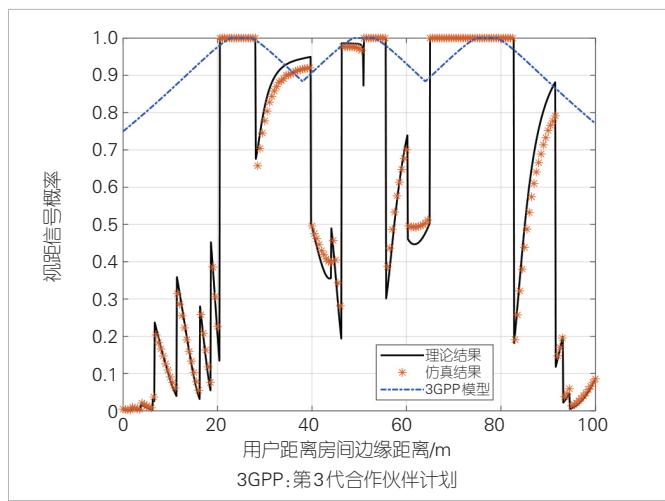
图2和图3分别展示了室内部署1个基站和部署3个基站的视距概率模型，并且对比了本文提出的视距概率模型与3GPP概率模型。从图中可以看出，3GPP模型可以基本刻画视距概率的变化趋势，但是由于室内阻挡物的原因，该模型在多数情况下高估了视距概率。本文所提出的视距概率模型在各种情况下均能较好地贴合场景仿真结果。使用本文模型能够更准确地分析室内通信系统性能。

图4和图5分别展示了在1个基站下视距概率与路径损耗、SNR的关系。从图中可以看出，随着视距概率减小，路径损耗增高，SNR降低。由于室内传播环境复杂，因此当基

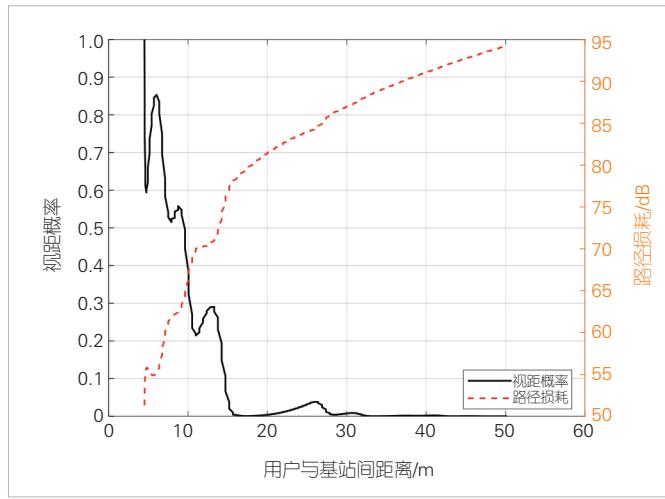


▲图2 1个基站下室内视距概率

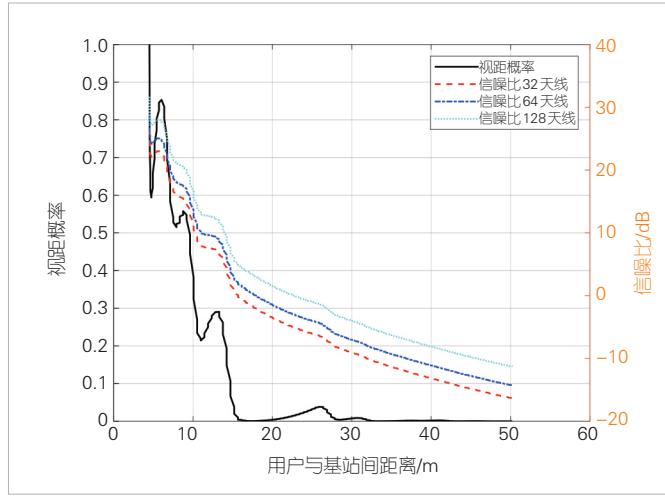
站与用户间距离超过30 m时，基本不存在视距信号分量，路径传播损耗已超过85 dB且信噪比低于-5 dB。此时，接收



▲图3 3个基站下室内视距概率



▲图4 视距概率与路径损耗的关系



▲图5 视距概率与信噪比的关系

信号质量差，难以维持正常通信。图5同时研究了不同发射端天线数情况下的SNR。可以发现，当天线数增加时，SNR有明显提升。因此，6G时代室内系统应用超大规模天线阵列将有助于提升系统性能。

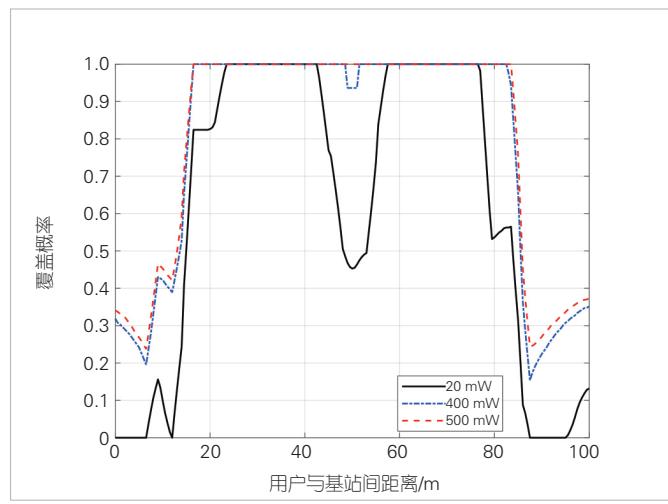
### 3.3 室内无线通信性能

图6展示了不同发射功率下的覆盖概率，此时室内部署2个基站。可以看出，增加信号发射功率可有效提高信号覆盖概率。但是由于遮挡效应，一味提升信号功率并不一定能够提升覆盖概率。例如，在房间边缘位置，400 mW与500 mW的覆盖概率近似。原因在于：障碍物形成覆盖死角，提升发射功率仅能缓解覆盖问题。因此，适当降低发射功率，可以获得近似的通信性能并且能降低成本。此外，优化建筑布局设计和基站部署、减少死角，都是提升性能的思路。

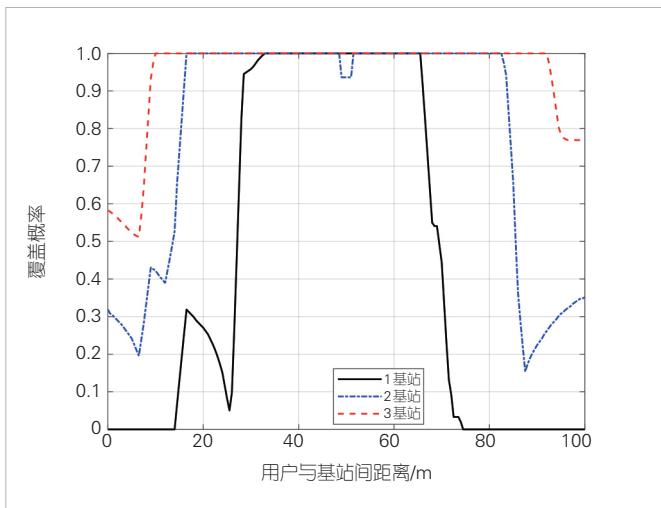
图7和图8分别展示了室内部署不同基站数下的覆盖概率和能量效率。图7说明增加基站数量可以有效改善覆盖概率。从图8中可以看出，当用户距离基站较近时，3基站的能效较高，原因在于用户可选择多个基站提升信号质量。然而，当距离增加时，2基站的能效较高，原因在于3基站中必然有1~2个基站无法服务该用户。图7和图8说明系统可在覆盖概率与能效中取得最优性能。

## 4 结论

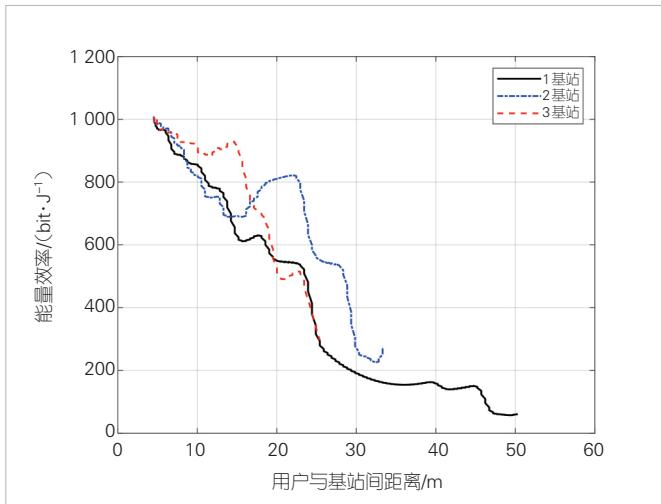
针对具有复杂传播环境的室内通信场景，本文提出了一种考虑障碍物高度的视距概率模型。该模型相较于3GPP视距概率模型可以更为准确地分析工业物联网、大型办公室、大型商场等6G室内场景视距概率。此外，本文基于6G普适信道模型，研究路径损耗、覆盖概率、能效等性能。结果证



▲图6 不同发射功率下的覆盖概率



▲图7 不同基站数量下的覆盖概率



▲图8 不同基站数量下的能量效率

明，视距概率伴随与基站间的距离的增加而降低，但由于障碍物存在，视距概率存在起伏。当用户与基站距离超过30 m时，几乎不存在视距信号，且信号质量较差。一味提升基站密度与信号功率可以提升覆盖概率，但效果并不显著，反而会降低系统能效。优化基站部署以及建筑室内布局同样是提升系统性能的可行方法。

## 参考文献

- [1] 中国联通. 中国联通5G数字化室分技术白皮书 [R]. 2019
- [2] ITU. Framework and overall objectives of the future development of IMT for 2030 and beyond: ITU-R 1-8 [S]. 2023
- [3] WANG C X, YOU X H, GAO X Q, et al. On the road to 6G: visions, requirements, key technologies, and testbeds [J]. IEEE communications surveys & tutorials, 2023, 25(2): 905–974. DOI: 10.1109/COMST.2023.3249835
- [4] YOU X H, WANG C X, HUANG J, et al. Towards 6G wireless communication networks: vision, enabling technologies, and new paradigm shifts [J]. Science China information sciences, 2021, 64(1): 110301. DOI: 10.1007/s11432-020-2955-6
- [5] YIN L, HAAS H. Coverage analysis of multiuser visible light communication networks [J]. IEEE transactions on wireless communications, 2018, 17(3): 1630–1643. DOI: 10.1109/TWC.2017.2782694
- [6] ABOAGYE S, IBRAHIM A, NGATCHED T M N, et al. Design of energy efficient hybrid VLC/RF/PLC communication system for indoor networks [J]. IEEE wireless communications letters, 2020, 9(2): 143–147. DOI: 10.1109/LWC.2019.2946144
- [7] LOU Z Y, BELMEKKI B E Y, ALOUINI M S. Coverage analysis of hybrid RF/THz networks with best relay selection [J]. IEEE communications letters, 2023, 27(6): 1634–1638. DOI: 10.1109/LCOMM.2023.3267900
- [8] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz: 3GPP TR 38.901 [S]. 2018
- [9] WU Y Z, KOKKONIEMI J, HAN C, et al. Interference and coverage analysis for terahertz networks with indoor blockage effects and line-of-sight access point association [J]. IEEE transactions on wireless communications, 2021, 20(3): 1472–1486. DOI: 10.1109/TWC.2020.3033825
- [10] LI Z Y, HU H N, ZHANG J L, et al. Coverage analysis of multiple transmissive RIS-aided outdoor-to-indoor mmWave networks [J]. IEEE transactions on broadcasting, 2022, 68(4): 935–942. DOI: 10.1109/TBC.2022.3196169
- [11] NEMATI M, MAHAM B, POKHREL S R, et al. Modeling RIS empowered outdoor-to-indoor communication in mmWave cellular networks [J]. IEEE transactions on communications, 2021, 69(11): 7837–7850. DOI: 10.1109/TCOMM.2021.3104878
- [12] LI J, BAO X, ZHANG W C, et al. QoE probability coverage model of indoor visible light communication network [J]. IEEE access, 2020, 8: 45390–45399. DOI: 10.1109/ACCESS.2020.2977936
- [13] YANG W F, ZHANG J L, GLAZUNOV A A, et al. Line-of-sight probability for channel modeling in 3-D indoor environments [J]. IEEE antennas and wireless propagation letters, 2020, 19(7): 1182–1186. DOI: 10.1109/LAWP.2020.2994392
- [14] ZHENG H, ZHANG J L, LI H, et al. Exact line-of-sight probability for channel modeling in typical indoor environments [J]. IEEE antennas and wireless propagation letters, 2018, 17(7): 1359–1362. DOI: 10.1109/LAWP.2018.2846748
- [15] AL-HOURANI A, KANDEEPAN S, LARDNER S. Optimal LAP altitude for maximum coverage [J]. IEEE wireless communications letters, 2014, 3(6): 569–572. DOI: 10.1109/LWC.2014.2342736
- [16] YANG S J, ZHANG J L, ZHANG J. Energy efficiency optimization in millimeter-wave air-to-ground links under UAV wobbling [C]// Proceedings of 2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, 2022: 572–576. DOI: 10.1109/PIMRC54779.2022.9977967
- [17] ITU. Propagation data and prediction methods required for the design of terrestrial broadband radio access systems operating in a frequency range from 3 GHz to 60 GHz: P.1410–6 [S]. 2023
- [18] WANG C X, LV Z, GAO X Q, et al. Pervasive wireless channel modeling theory and applications to 6G GBSMs for all frequency bands and all scenarios [J]. IEEE transactions on vehicular technology, 2022, 71(9): 9159–9173. DOI: 10.1109/TVT.2022.3179695
- [19] FU Y, WANG C X, MAO X C, et al. Spectrum-energy-economy efficiency analysis of 5G wireless communication systems with separated indoor/outdoor scenarios [J]. IEEE transactions on wireless communications, 2023, early access. DOI: 10.1109/TWC.2023.3273261

## 作者简介



毛曦晨，东南大学在读博士研究生；主要研究方向为6G无线通信系统性能优化、室内通信系统设计与优化。



**王承祥**, 东南大学首席教授、博士生导师, 移动通信全国重点实验室和网络通信与安全紫金山实验室教授, 东南大学信息科学与工程学院执行院长, 欧洲科学院院士, 欧洲科学与艺术院院士, 英国爱丁堡皇家学会院士, IEEE Fellow, IET Fellow, 中国通信学会会士; 主要研究方向为无线信道测量与建模、6G智能网络架构与关键技术、电磁信息论。



**杨松江**, 网络通信与安全紫金山实验室博士后; 主要研究方向为毫米波信道建模、无人机空对地信道模型、射线追踪信道建模方法、频谱效率分析和能源效率分析。



**黄杰**, 东南大学移动通信全国重点实验室副研究员、网络通信与安全紫金山实验室科研人员; 主要研究方向为毫米波、太赫兹、超大规模天线、智能反射面信道测量与建模、电磁信息论、6G无线通信关键技术。



**黄晨**, 网络通信与安全紫金山实验室研究员、东南大学校外导师、江苏省卓越博士后、中国通信学会高级会员、中兴通讯“促进产学研合作青年专家委员会”委员, 入选中国科协青年人才托举工程; 主要研究方向为6G智能无线通信、基于人工智能的无线信道特性分析与建模研究。

## 综合信息

### 《中兴通讯技术》2024年专题计划

期次	专题名称	策划人
1	下一代多址技术	北京交通大学教授 艾渤 北京交通大学教授 陈为
2	网络大模型	中国电信IP领域首席专家 唐宏 中兴通讯无线首席架构师 熊先奎
3	6G多天线技术	东南大学首席教授 金石 北京交通大学教授 章嘉懿 东南大学副研究员 韩瑜
4	6G无线系统技术	中国信息通信研究院副院长 王志勤 中国移动研究院院长 黄宇红 东南大学教授 王东明
5	卫星通信技术	哈尔滨工业大学(深圳)教授 张钦宇
6	数据通信新技术	中国电信研究院教授级高工 解冲锋 中国联通研究院副院长 唐雄燕



# 基于服务覆盖的无蜂窝边缘智能网络部署方法

Service Coverage–Oriented Design on Mobile Edge Computing Aided Cell–Free Massive MIMO Networks

姜静/JIANG Jing<sup>1</sup>, 乔东伟/QIAO Dongwei<sup>1</sup>,

Worakrin SUTTHIPHAN<sup>2</sup>

(1. 西安邮电大学陕西省信息通信网络与安全重点实验室, 中国 西安 710121; )

2. 泰国国家电信有限公司, 泰国 曼谷 10700)

(1. Shaanxi Key Laboratory of Information Communication Network and Security, Xi'an University of Posts and Telecommunication, Xi'an 710121, China;

2. National Telecommunication Public Company, Bangkok 10700, Thailand)

DOI:10.12142/ZTETJ.202306008

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231211.1119.006.html>

网络出版日期: 2023-12-12

收稿日期: 2023-10-10

**摘要:** 为实现未来网络绿色低碳发展, 以精准的按需服务提高网络运营和部署效率, 将服务覆盖作为网络覆盖能力的新度量。深入挖掘无蜂窝网络部署灵活的优势, 可以使每个覆盖单元达到最佳服务覆盖, 从而实现系统性能的差异化优化。结合无蜂窝边缘智能网络的通信链路质量对移动边缘计算(MEC)服务器计算和下载的时延、卸载概率的影响, 采用连续凸逼近算法, 引入松弛变量和辅助变量将接入点(AP)和MEC联合部署混合整数非线性规划问题转化为连续的凸优化问题进行求解。仿真结果证明, 基于服务覆盖的网络部署方法提升了现有网络能量效率, 并显著减少用户的计算时延和系统的能耗。

**关键词:** 无蜂窝大规模多输入多输出(MIMO); 服务覆盖; 移动边缘计算

**Abstract:** To achieve the green and low-carbon development of the future network, a new metric of network coverage performance is redefined, e.g. service coverage, to improve the efficiency of network operation by providing accurate on-demand services. Digging deeply into the advantages of flexible cell-less network deployment can optimize the system performance differentiation for each coverage unit, thus enabling overall optimization of system performance. The impact of the communication link quality is considered with the offloading delay and probability of the mobile edge computing (MEC) server. The continuous convex approximation algorithm to solve access point (AP) and MEC joint deployment is proposed, in which the slack variables and auxiliary variables are introduced to transform the mixed integer nonlinear programming problem into a continuous convex optimization problem. Simulation results show that the service coverage-based network deployment method could improve the energy efficiency of the existing network, and significantly reduce the computing delay of users and the energy consumption of the system.

**Keywords:** cell free massive multiple–input multiple–output (MIMO); service coverage; mobile edge computing

**引用格式:** 姜静, 乔东伟, Worakrin SUTTHIPHAN. 基于服务覆盖的无蜂窝边缘智能网络部署方法 [J]. 中兴通讯技术, 2023, 29(6): 46–53. DOI: 10.12142/ZTETJ.202306008

**Citation:** JIANG J, QIAO D W, SUTTHIPHAN W. Service coverage–oriented design on mobile edge computing aided cell–free massive MIMO networks [J]. ZTE technology journal, 2023, 29(6): 46–53. DOI: 10.12142/ZTETJ.202306008

随着5G网络的规模化部署及新一代信息通信技术的高速发展, 信息通信行业面临着日益增长的能耗和碳排放压力。预计到2025年, 通信行业将消耗全球20%的电力。

基金项目: 国家自然科学基金(62371392、61901370、61971027); 陕西省国际科技合作重点项目(2023-GHZD-37); 陕西省重点产业链项目(2023-ZDLGY-49)

为了实施国家的“双碳”战略, 工业和信息化部等7个部门联合发布了《信息通信行业绿色低碳发展行动计划(2022—2025年)》, 强调未来网络应实现绿色、低碳和高质量发展。运营商希望充分利用各种场景下的业务特征, 实现无线服务的按需精准投放, 减少网络建设和运营成本, 提高资源利用效率和成本效益<sup>[1-2]</sup>。为了实现该目标, IMT-2030(6G)推进组发布了《6G典型场景和关键能力》白皮书, 提

出了6G网络应具备按需服务能力。即通过已有业务数据的分析和挖掘，实现个性化和精准的服务供给<sup>[3]</sup>。因此，绿色低碳发展和按需服务将成为6G网络设计和部署的关键要素，打造更环保、高效和可持续发展的移动通信网络。

无蜂窝大规模多输入多输出（MIMO）于2017年由瑞典林雪平大学的E. G. LARSSON教授和美国纽约大学的T. L. MARZETTA教授在分布式天线系统（DAS）的基础上共同提出，旨在解决5G部署中传统蜂窝通信网络小区间干扰严重、越区切换频繁等问题<sup>[4]</sup>。无蜂窝大规模MIMO系统在覆盖范围内分布部署大量的无线接入点（AP），通过回程链路与中央处理器（CPU）链接，多个AP形成协作集合为用户提供服务。6G网络在架构层面将通信网络与算力、算法和数据深度融合，构建网络内完整的智能体系。如图1所示，无蜂窝大规模MIMO边缘智能网络在CPU或AP部署移动边缘计算（MEC）服务器，通过MEC的缓存功能显著降低前/后向链路的流量负载和系统能耗，从而解决无蜂窝网络部署时前/后向链路容量受限的瓶颈问题；MEC借助无蜂窝均匀、一致的覆盖性能，提高计算和缓存下载的成功概率<sup>[5]</sup>。无蜂窝大规模MIMO边缘智能网络结合了MEC和无蜂窝大规模MIMO系统的性能优势，融合通信、计算、存储等功能，具有高效、可靠、低延迟与低成本的优势，成为6G的重要研究方向之一<sup>[6-7]</sup>。

传统的无蜂窝大规模MIMO系统假定AP随机、均匀分布在覆盖区域。该假设较为理想化，未充分挖掘无蜂窝网络部署灵活的性能优势。文献[8]首次研究无蜂窝大规模MIMO中的AP位置优化问题，证明了基于最大吞吐量的AP位置优化方法与AP随机部署的传统方案相比，系统总吞吐量提高了3倍；与Lloyd算法（和距离最小）部署方法相比，系统

总吞吐量提高了1.25倍。文献[9]将遗传算法（GA）与射线追踪（RT）信道模拟方法相结合，以最大化系统能效为目标生成最佳AP部署方案。上述工作证明了AP部署的优化将显著提高无蜂窝大规模MIMO系统性能。但是，现有研究中无蜂窝大规模MIMO技术的研究通常以系统整体的吞吐量最大、能效最优等作为优化目标<sup>[10-12]</sup>，未考虑用户和移动业务在不同区域的非均匀分布特性，如何通过无蜂窝网络AP和CPU部署等实现6G精准、个性化的服务供给尚未有研究。

为了匹配非均匀的移动业务分布特征，实现按需精准投放，提高资源投入的性价比，服务覆盖作为一种新的网络性能度量被提出<sup>[13]</sup>。服务覆盖的定义为每覆盖单元的传输容量与业务流量需求的比率，表征了6G网络的服务能力与业务需求之间的匹配程度。文献[14]以最佳服务覆盖为目标，实现覆盖单元的总吞吐量与平均速率需求的比率之和的最大化，以此来优化接入点的部署。构建基于服务覆盖的目标模型可以使无蜂窝大规模MIMO网络达到按需服务的能力。同时，基于多元线性回归模型的梯度下降算法可以求解接入点的位置问题，实现无蜂窝网络业务需求的精准匹配，显著提升系统的能量效率与频谱效率。

针对传统蜂窝网络中MEC服务器的部署问题，文献[15]提出在通信和计算时延受限的条件下，推导MEC系统模型下通信延迟和计算延迟与网络负载和网络资源参数的比例关系，为AP、MEC服务器和中心服务器（CS）的部署提供了理论依据。文献[16]以最大化服务范围内的资源利用率为目地，求解了基于随机部署的MEC服务器最优部署密度。文献[17]利用非均匀泊松点过程模拟MEC服务器的部署过程，求解满足预期业务需求的基站和MEC服务器位置和数量。文献[18]针对由不同传输功率和计算能力的多层次异构无线网络，根据MEC服务器的用户类型和AP分层设计求解最优成功边缘计算概率，为MEC的优化配置提供指导。针对无蜂窝大规模MIMO边缘智能网络的MEC服务器部署，文献[6]证明在无蜂窝大规模MIMO边缘智能网络的MEC部署中，CS成功接收到用户的卸载任务的概率随着AP覆盖范围的增大而降低。因此，在给定的成功卸载概率情况下，配置更多的AP将会更节能。但是上述MEC服务器的部署研究尚未解决在考虑通信和计算深度耦合条件下，无蜂窝边缘智能网络如何提供满足多样化业务需求的通信传输能力和算力。

本文将研究基于服务覆盖的无蜂窝网络AP和MEC的联合部署，首先建立多样化业务需求的服务覆盖理论模型，并利用元学习的学习能力，实现多种应用场景和多样化业务需求的服务覆盖建模；然后根据非均匀分布的业务需求，以每个覆盖单元的最佳服务覆盖为目标，研究大规模MIMO无蜂



▲图1 基于服务覆盖的无蜂窝大规模多输入多输出智能网络

窝网络 AP 和 MEC 在多资源约束条件下的联合部署和构型，通过松弛引入辅助变量将非凸离散问题转化为凸连续问题，求解接入点和 MEC 的联合部署位置，从而实现无蜂窝大规模 MIMO 网络中与实际业务相匹配的传输能力与协同算力。

## 1 基于服务覆盖的无蜂窝边缘网络设计

针对多样化的业务需求和应用场景，根据服务覆盖定义结合 6G 典型应用场景建立多样化的服务覆盖理论模型，进而以最大服务覆盖为目标，设计和优化 AP、CPU 以及 MEC 服务器的部署、网络构型、算力资源和频谱资源，实现精准按需服务的无蜂窝大规模 MIMO 边缘智能网络的传输能力、计算缓存能力和存储能力。

### 1.1 基于服务覆盖需求统计预测和设计

6G 业务丰富多样，每种业务对传输速率、流量密度、空口时延、同步和抖动、移动性、计算时延、计算能力等性能指标的需求有所不同。为精确统计不同业务需求，本文将覆盖区域划分为  $N$  个覆盖单元，如图 1 所示。与文献[13]中服务覆盖定义为传输速率和业务需求速率之比不同，针对多样化的业务需求，本文中我们假设场景的关键服务能力由多个元素组成，例如：传输速率、时延和可靠性要求等，则服务覆盖定义如下：

$$\vartheta_n = a_{n1} \frac{C_1}{T_1} + a_{n2} \frac{C_2}{T_2} + a_{n3} \frac{C_3}{T_3}, \quad (1)$$

其中， $a_{ni}$  为每种关键能力的加权系数，其取值与该区域的特定业务占比、用户模型等有关，所有关键能力的加权系数之和为 1。 $C_i, i = 1, 2, 3$  为覆盖单元内可实现的通信和计算能力， $T_i, i = 1, 2, 3$  该区域内不同类型的业务需求。

根据上述定义，我们利用元学习的学习能力不断学习或获取典型场景中的业务特征和关键能力表达，实现多种应用场景和多样化业务需求的服务覆盖建模分析。具体流程如图 2 所示：首先，定义关键业务需求和服务能力的数据标签和训练样本，嵌入 6G 典型场景的训练样本；其次，通过实际场景的少量样本，训练和获取不同移动业务的关键技术需求；再次，借助线性分类器产生不同业务的权重系数，平衡多样化业务需求对网络关键服务能力的要求，确定网络的关键服务能力；最后，依据不

同覆盖区域的业务需求和关键服务能力，建立服务覆盖模型。利用测试样本完成模型的验证和优化，可以提高服务覆盖的准确性、系统性和科学性。

### 1.2 基于服务覆盖的无蜂窝大规模 MIMO 部署设计

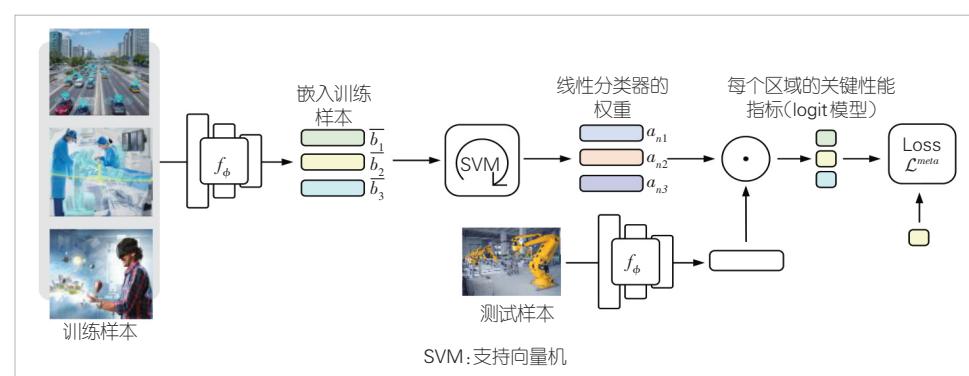
在无蜂窝网络中，AP 部署的密度、位置、服务范围决定了 AP 与用户间的路径损耗，直接影响容量、能耗和通信时延等系统性能。基于服务覆盖的无蜂窝大规模 MIMO 部署设计将以每个覆盖单元的最佳服务覆盖为目标，通过构建基于 AP 数量和位置的无蜂窝网络可达速率的闭合表达式，设计大规模 MIMO 无蜂窝系统的拓扑结构，优化无蜂窝网络的系统性能。在优化的过程中，让每一个覆盖单元的业务需求和传输能力相匹配，可以最小化总覆盖范围内的匹配能力差之和，从而优化每个覆盖单元的精细化服务能力迭代，最终实现无蜂窝网络的按需、精准服务能力。

### 1.3 基于服务覆盖的无蜂窝边缘智能网络设计

在无蜂窝边缘智能网络中，AP 或 CPU 配置了 MEC 服务器，网络边缘能够实时存储、下载、处理和分析数据。面对不同应用场景下丰富多样的业务需求，基于服务覆盖的无蜂窝边缘智能网络不仅需要考虑通信需求，还需要考虑算力、计算时延等业务需求的匹配程度。为实现无蜂窝边缘智能网络通信、计算、存储的联合部署，不仅需要优化 AP 个数和密度、AP 部署位置，还需要优化 CPU 和 MEC 的算力、MEC 部署个数和容量、MEC 和 AP 的连接关系，从而实现与业务需求相匹配的计算、存储和通信能力。进一步地，CPU 根据流量负载、缓存和计算需求对通信资源进行动态分配和调度，匹配用户的需求，实现通信和计算能力的协同优化，提升资源利用率和通信计算质量。

## 2 系统架构与目标函数

针对无蜂窝边缘智能网络中非均匀的通信与计算需求，



▲图2 基于元学习的服务覆盖建模分析

本文将通过AP和MEC的部署优化，实现与每个覆盖单元业务需求相匹配的计算与通信能力。首先，构建无蜂窝边缘智能网络通信与计算系统模型。其中，通信模型建模了覆盖单元内用户的可达传输速率；计算模型用于表征该覆盖区域内MEC的计算处理能力，用于匹配覆盖单元内用户非均匀的计算任务。其次，分析了无蜂窝边缘智能网络中的通信与计算能力之间的关系。最后，以最大化每个覆盖单元内的服务覆盖为目标，每个覆盖单元获得相匹配的通信计算能力。

## 2.1 系统模型

在无蜂窝边缘智能网络系统中，我们假设所有AP均配备独立的MEC服务器并通过回程链路连接到CPU，并分布在N个覆盖单元内，如图1所示。假设第m个AP和第k个用户之间的信道表示为： $g_{mk} = \beta_{mk}^{\nu_2} h_{mk}$ ，其中 $\beta_{mk}$ 和 $h_{mk}$ 分别为大尺度衰落和小尺度衰落， $h_{mk} \sim \mathcal{CN}(0,1)$ 。相干周期 $\tau_c$ 划分为 $\tau_p$ 个导频传输间隔和 $\tau_c - \tau_p$ 个上行链路数据传输间隔。第k个用户在相干时间开始同时发送导频序列 $\psi_k \in \mathbb{C}^{\tau_c \times 1}$ ， $\|\psi_k\|_2^2 = 1$ ， $\tau_p = K$ 。因此，第m个AP处接收到的信号为：

$$\gamma_m^u = \sum_{k=1}^K \sqrt{p_k} g_{mk} x_k + \omega_m^p, \quad (2)$$

其中， $p_k = \eta_k p_k^{\max}$ 为发射功率， $\eta_k$ 和 $p_k^{\max}$ 分别表示第k个用户的功率控制系数和最大上行链路功率。 $x_k$ 表示第k个用户的上行链路数据， $\omega_m^p$ 是一个 $\tau_p$ 维的加性噪声，且满足 $\omega_m^p \sim \mathcal{CN}(0,1)$ 。

第k个用户的信号通过AP经由CPU被组合成如下解码形式：

$$\hat{x}_k = \sum_{m=1}^M \hat{g}_{mk}^* \gamma_m^u, \quad (3)$$

其中， $\hat{g}_{mk} = \psi_k^H \mathbf{y}_m^p / \sqrt{\tau_p p_k}$ 为最小二乘信道估计， $\mathbf{y}_m^p \in \mathbb{C}^{\tau_p \times 1}$ 为第m个AP接收到的导频矢量。我们假设AP和MEC只能部署在L个备选位置，并将第n个覆盖单元定义为 $U_n$ ，则第n个单元的传输能力可以表示为单元内所有用户的总上行传输容量：

$$C_n(\mathbf{z}) = \sum_{k \in U_n} W \log_2 \left( 1 + \sum_{l=1}^L z_l \gamma_k \right) = \\ \sum_{k \in U_n} W \log_2 \left( 1 + \frac{p_k \sum_{l=1}^L z_l |\hat{g}_{lk}^* g_{lk}|^2}{\sum_{l=1}^L z_l \left( \sum_{k'=1}^K p_{k'} |\hat{g}_{lk}^* g_{lk'}|^2 + \hat{g}_{lk}^2 \right)} \right), \quad (4)$$

其中， $\mathbf{z} \in \mathbb{R}^{L \times 1}$ 为0和1组成的决策向量，表示为 $\mathbf{z} = \{z_1, z_2, \dots, z_L\}^T$ ， $z_l = 1$ 表示在第l个位置联合部署AP和MEC。 $\gamma_k$ 为第k个用户的信干噪比， $k \in U_n$ 表示第k个用户位于第n个单元。

假设第k个用户有大小为 $Z_k$ bits的计算密集型任务，由于移动设备的计算能力和电池容量有限，用户端无法承受高强度的计算任务，因此假设第k个用户将 $Z_k$ bits任务全部卸载至边缘服务器而不考虑在本地并行计算。当计算任务卸载到边缘服务器时，经历的延迟可以分解为卸载数据的传输延迟和边缘服务器处理的计算延迟。

因此，第n个单元的时延表示为第n个覆盖单元内用户的总时延，具体为：

$$t_n(\mathbf{z}) = t_n^{\text{comp}} + \sum_{l=1}^L z_l t_n^{\text{tr}} = \sum_{k \in U_n} \left( \frac{T_k N_{cpb}}{\sum_{m \in U_n} f_{mk}} + \sum_{l=1}^L z_l \frac{T_k}{C_n(\mathbf{z})} \right), \quad (5)$$

其中， $t_n^{\text{tr}}$ 表示第n个覆盖单元内用户卸载数据到边缘服务器所需的传输延迟； $t_n^{\text{comp}}$ 表示第n个单元内边缘服务器处理用户任务的总计算延迟； $f_{mk}$ 为连接第m个AP的MEC用于执行第k个用户的处理速度，单位为cycles/s，且 $f_{mk} \leq f^{\max}$ ， $f^{\max}$ 为每个边缘服务器处理能力上限； $N_{cpb}$ 为处理1 bit信息所需的周期数，单位为cycles/bit。

## 2.2 目标函数

本文以实现最大服务覆盖为目标，根据差异化的业务需求设计了无蜂窝边缘智能网络的AP和算力部署，包括MEC的计算能力 $f_{mk}$ 、MEC服务器个数M、AP和MEC服务器联合部署位置 $\mathbf{z}$ 。假定服务覆盖的关键能力要素包括网络算力、系统时延和通信可达速率，则服务覆盖区域定义如下：

$$\vartheta_n = a_{n1} \frac{T_n(\mathbf{z})}{\bar{T}_n} + a_{n2} \frac{t_n(\mathbf{z})}{\bar{t}_n} + a_{n3} \frac{C_n(\mathbf{z})}{\bar{C}_n}, \quad (6)$$

其中， $a_{nt} \in [0,1]$ 为每种关键能力的加权系数且 $\sum_{t=1}^3 a_{nt} = 1$ ，其取值与该区域的特定业务占比、用户模型等有关。 $T_n(\mathbf{z})$ 为第n个覆盖区域系统可提供的计算能力， $\bar{T}_n$ 为该区域的算力需求； $t_n(\mathbf{z})$ 为第n个覆盖区域内用户处理密集型任务可达到的最小时延， $\bar{t}_n$ 为第n个覆盖区域内应满足的最小时延； $C_n(\mathbf{z})$ 为第n个覆盖区域系统可提供的传输容量， $\bar{C}_n$ 为第n个位置区域业务传输所需的平均容量。

在无蜂窝边缘智能网络中，针对不同覆盖单元非均匀的通信与计算需求，通过合理的算力资源和AP部署，可以为用户提供低时延、高可靠的通信和计算需求，并满足系统最

小计算和通信时延的约束条件。该问题建模为P1:

$$\min_{\mathbf{z}, t_n^u, \gamma_k} \frac{1}{N} \sum_{n=1}^N (\vartheta_n(z) - \sigma)^2, \quad (7)$$

$$\text{s.t. } z_l = 0 \text{ 或 } 1, l = 1, \dots, L, \quad (8)$$

$$t_n^{comp} + t_n^{tr} \leq t_n, \quad (9)$$

第1个限制条件下,  $z$  表示 AP 和 MEC 的部署方案,  $z_l$  为 0 或 1 组成的 AP 和 MEC 部署位置元素, 通过联合部署实现每个覆盖区域算力和通信需求的精准匹配; 第2个限制条件为第  $n$  个覆盖单元内, 传输和计算时延应小于该单元内移动业务允许的最小时延。 $\sigma$  为冗余因子, 为网络业务突发留有合理的传输能力冗余。

### 3 基于服务覆盖的无蜂窝边缘智能网络部署

P1 是一个多目标多约束的混合整数非线性规划问题, 首先将覆盖区域内通信和计算能力转化为约束条件, 然后利用拉格朗日松弛法和罚函数法对问题中的整数约束进行松弛得到连续非凸问题, 最后通过引入辅助变量将非凸目标函数和约束条件进行近似, 再利用连续凸逼近算法迭代求解。

首先, 问题 P1 转化为 P2:

$$\min_{\mathbf{z}, t_n^u, \gamma_k} F(\mathbf{z}, t_n^u, \gamma_k) = \min_{\mathbf{z}, t_n^u, \gamma_k} \frac{1}{N} \sum_{n=1}^N (\vartheta'_n(z) - \sigma)^2, \quad (10)$$

$$\text{s.t. } z_l = 0 \text{ 或 } 1, l = 1, \dots, L, \quad (11)$$

$$t_n^{comp} + z_l t_n^{tr} \leq t_n, \quad (12)$$

$$T_n(z) \leq a_{n1} \sigma \bar{T}_n, \quad (13)$$

$$C_n(z) \leq a_{n3} \sigma \bar{C}_n, \quad (14)$$

$$t_n^{comp} \geq t_{\min}^{comp}, t_n^{tr} \geq t_{\min}^{tr}, f_{mk} \geq 0, \forall n \in [1, N]. \quad (15)$$

其中,  $\vartheta'_n = a_{n2} \frac{t_n(z)}{t_n}$ 。约束 (15) 是由于用户的发射功率存在上限  $p_k \leq p_k^{\max}$ , 那么可得传输延迟存在下限  $t_{\min}^{tr}$ ; 边缘服务器处理能力存在上限  $f_{mk} \leq f^{\max}$ , 那么可得到计算时延存在下限  $t_{\min}^{comp}$ 。由于约束条件以及目标函数存在变量相乘的非凸结构, 因此先对 0~1 整数约束 (11) 进行松弛, 等价为下面两个约束条件:

$$z_l - z_l^2 \geq 0, \forall l \in [1, L], \quad (16)$$

$$0 \leq z_l \leq 1, \forall l \in [1, L]. \quad (17)$$

其中, 离散变量  $z_l$  被松弛为取值范围从 0~1 的连续变量, 然后利用罚函数法使约束条件 (16) 成为目标函数的一部分, 优化问题 P2 则转换为 P3:

$$\min_{\mathbf{z}, t_n^u, \gamma_k} F(\mathbf{z}, t_n^u, \gamma_k) + \mu \sum_{l=1}^L (z_l - z_l^2), \quad (18)$$

$$\text{s.t. } (12) - (15), (17).$$

其中,  $\mu$  是惩罚因子, 在  $z_l$  不等于 0 或 1 时惩罚目标函数, 迫使松弛后的连续变量  $z_l$  逼近 0 或 1。通过松弛, 约束条件 (12) — (15) 和目标函数中离散变量乘连续变量的结构变成了连续变量乘连续变量的结构, 但仍是非凸的, 通过引入辅助变量  $r$  和新的非凸约束条件, 可得:

$$F(\mathbf{z}, t_n^u, \gamma_k) + \mu \sum_{l=1}^L (z_l - z_l^2) \leq r. \quad (19)$$

这时可将 P3 等价转换为  $\min r$ 。我们设  $\min r$  可行域为  $\Omega = \{z, t_n^u, \gamma_k, r\}$ , 根据文献[19]先求得 (12) — (15)、(17) 和 (19) 这些非负约束条件在可行解  $\Omega(k)$  处的近似函数, 从而得到连续的凸优化问题 P4:

$$\min_{\mathbf{z}, t_n^u, \gamma_k, r} r, \quad (20)$$

$$\text{s.t. } (15), (17), (21), (22), (23), (24).$$

其中, (21) — (24) 为非凸相乘转化为在可行解  $\Omega(k)$  处的 DC 函数近似形式:

$$t_n^{comp} + z_l t_n^{tr} \leq t_n^{comp} + \tilde{g}_{1,n}(z_l, t_n^u; z_l(k), t_n^{tr}(k)) \leq t_n, \quad (21)$$

$$T_n(z) = f_n(z_l, f_{mk}) \leq \tilde{f}_n(z_l, f_{mk}; z_l(k), f_{mk}(k)) \leq a_{n1} \sigma \bar{T}_n, \quad (22)$$

$$C_n(z) = g_n(z_l, \gamma_k) \leq \tilde{g}_n(z_l, \gamma_k; z_l(k), \gamma_k(k)) \leq a_{n3} C_n(z), \quad (23)$$

$$F(z_l, t_n^u, \gamma_k) + \mu \sum_{l=1}^L \tilde{g}_{l,1}(z_l; z_l(k)) \leq r. \quad (24)$$

凸优化问题 P4 可以通过内点法或 MATLAB 的凸优化工具箱 (CVX) 等方法求解, 从而得到最优解  $\hat{\Omega} = \{\hat{\mathbf{z}}, \hat{t}^u, \hat{\gamma}_k, \hat{r}\}$  对应  $\min r$  在约束 (19) 和可行解  $\Omega(k)$  处近似函数的最小值。根据连续凸逼近, 我们令  $\Omega(k+1) = \hat{\Omega}$ , 并将  $\Omega(k+1)$  作为新的可行解代入问题 P4, 对新的近似函数进行下一轮凸优化求解, 直到结果收敛到局部最优解或不可行解。基于连续凸逼近的迭代算法如算法 1。

#### 算法 1: 基于连续凸逼近的迭代求解算法

**初始化:** 定义精度  $\varepsilon$ , 最大迭代次数  $I_{\max}$ , 迭代序号  $n=0$ , 惩罚因子  $\eta$ , 惩罚因子倍增系数  $c=2$ , 初始最优值  $r^{(0)} = \inf$ , 设置

初始可行解  $\Omega(0) = \{z(0), t^r(0), \gamma(0), r(0)\}$ ;

**Repeat:**

将  $\Omega(n)$  和  $\eta$  代入问题 P4;

更新迭代序号、惩罚因子:  $n \leftarrow n + 1, \eta = c \cdot \eta$ ;

求解问题 P4, 得到最优解  $\hat{\Omega}$ , 最优值  $r^{(n)}$ ;

更新可行解:  $\Omega(n) \leftarrow \hat{\Omega}$ ;

**Until**  $|r^{(n)} - r^{(n-1)}| \leq \varepsilon$  或者  $n \geq I_{\max}$ ;

对  $z$  中各元素四舍五入取整得到  $\hat{z}$ ;

将  $\hat{z}^*$  代入问题 P2, 求解得到其他优化变量  $\{\gamma_k^*, t^{tr*}\}$ ;

**输出:**  $\{\hat{z}^*, \gamma_k^*, t^{tr*}, r\}$

#### 4 仿真结果与分析

在仿真过程中, 我们首先生成非均匀的用户业务流量场景, 并得到该区域的平均流量需求。然后, 利用文献[21]中的热点模型生成非均匀分布的用户位置, 使用户在热点区域分布得更加密集, 在低业务区域分布得更加稀疏, 并在该模型下, 通过调整热点区域的数量来使用户分布更符合真实场景。我们设置了3个热点区域, 利用偏移因子  $\mu_\delta \in [0, 1]$  来表示用户在热点区域中的聚类程度。当  $\mu_\delta = 0$  时, 用户均匀分布在覆盖区域内; 随着  $\mu_\delta$  的增加, 用户更紧密地分布在热点区域。对于业务流量的产生, 假设流量到达遵循泊松过程, 那么可以通过设置平均到达率  $v$  来模拟各种不均匀的流量分布。

本文中, 我们考虑了一个由  $M$  个配备了 MEC 的 AP 和  $K$  个用户组成的  $1 \text{ km} \times 1 \text{ km}$  的无蜂窝大规模 MIMO 边缘智能网络系统。我们将覆盖区域划分为  $N$  个覆盖单元, 然后根据服务覆盖对区域内的  $L$  个部署点进行  $M$  个位置的部署选择。第  $k$  个用户和第  $m$  个 AP 之间的小尺度衰落服从  $h_{mk} \sim \mathcal{CN}(0, 1)$ , 大尺度衰落系数建模为:

$$\beta_{mk} = 10^{\frac{PL_{mk}}{10}} 10^{\frac{\sigma_{sh} z_{mk}}{10}}, \quad (25)$$

其中,  $PL_{mk}$  表示自由空间路径损耗,  $10^{\frac{\sigma_{sh} z_{mk}}{10}}$  表示具有标准差为  $\sigma_{sh}$  以及  $z_{mk} \sim \mathcal{CN}(0, 1)$  的阴影衰落。 $PL_{mk}$  采用三斜率模型<sup>[20]</sup>表示:

$$PL_{mk} = \begin{cases} -L - 35 \log_{10}(d_{mk}) & \text{if } d_{mk} > d_1 \\ -L - 10 \log_{10}(d_{mk}^2 d_1^{1.5}) & \text{if } d_0 < d_{mk} \leq d_1 \\ -L - 10 \log_{10}(d_0^2 d_1^{1.5}) & \text{if } d_{mk} \leq d_0 \end{cases}, \quad (26)$$

其中,

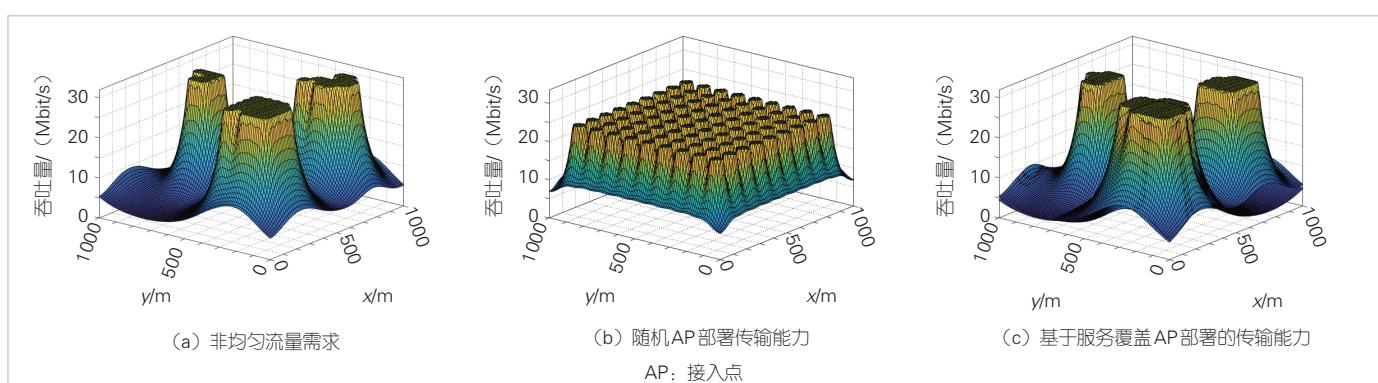
$$L = 46.3 + 33.9 \log_{10}(f) - 13.82 \log_{10}(h_{AP}) - (1.1 \log_{10}(f) - 0.7) h_u + 1.56 \log_{10}(f) - 0.8 \text{ dB}。 \quad (27)$$

系统的部分仿真参数如表1所示。

本文中, 我们以不同的方式部署 AP, 以此来评估和比较基于服务覆盖的 AP-MEC 部署方法和随机部署方法的系统性能。图3为在非均匀业务需求分布情况下, 不同部署 AP 方法下系统可达容量对比图。图3中的横坐标  $x$ 、 $y$  表示小区覆盖区域的地理位置, 纵坐标  $z$  表示可实现的吞吐量, 用户的业务需求呈非均匀分布, 如图3(a)所示。然而, 传统的随机部署 AP 方法实现了均匀一致的覆盖, 如图3(b)所示, 这造成传输能力在低流量需求区域过剩, 高流量需求区域不足。

▼表1 仿真参数

参数	取值
系统带宽	10 MHz
处理 1 bit 所需的周期数 $N_{cpb}$	500 cycles/bit
边缘服务器计算能力 $f^{mec}$	100 GHz
载频 $f$	1.9 GHz
阴影衰落的标准差 $\sigma_{sh}$	10 dB
AP 天线高度 $h_{AP}$	15 m
用户天线高度 $h_u$	1.65 m
三斜率模型参数 $d_1$	50 m
三斜率模型参数 $d_0$	10 m

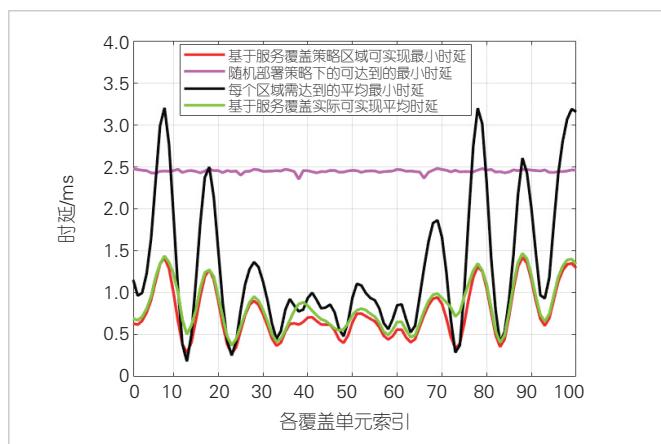


▲图3 非均匀业务需求和不同部署 AP 方法下系统的可达容量

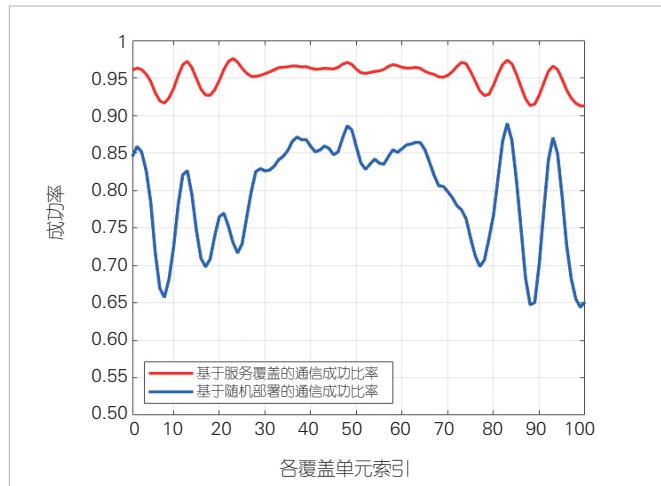
域明显不足。图3(c)可以实现与图3(a)流量需求相一致的吞吐量,这证明基于服务覆盖的AP部署方法能够满足热点模型区域所需的业务需求。

图4比较了随机AP部署和基于服务覆盖的AP-MEC部署的平均时延分布图。在随机部署方法下,各覆盖单元可实现的时延均值约为2.5 ms。这是因为随机部署算法由于部署AP位置的随机性,导致某些区域无线网络服务能力不足,产生了业务拥塞,而另一些区域覆盖能力过剩。这种不均衡网络资源分布更容易导致用户的时延超出业务允许的最小时延。而在基于服务覆盖部署AP-MEC方法下,各覆盖单元内可实现的处理时延均值约为0.7 ms,低于各覆盖单元所需的时延均值1.8 ms。这是因为在基于服务覆盖部署方法中,我们在业务需求高的覆盖区域部署了更多的AP,缩短了用户的接入和传输时延,从而提高系统性能和用户体验。

图5为不同AP-MEC部署方法下,每个覆盖单元的通信成功率,即覆盖单元内满足最低通信需求和最低时延的概



▲图4 随机部署移动边缘计算和基于服务覆盖部署移动边缘计算的可达最小时延分布



▲图5 随机部署和服务覆盖的通信成功率

率。AP-MEC随机部署方法的成功概率均值约为0.754 3,曲线抖动明显,在数据传输过程中会导致数据丢失或重传增加通信的延迟和不稳定性;而基于服务覆盖的通信成功率的均值约为0.945 2,曲线平稳,在高通信成功率区域常伴随着较低的丢包率,有助于提高实时通信应用的质量,证明了基于服务覆盖部署方法能够提供高效稳定的通信。

## 5 结束语

为解决现有网络以系统整体性能优化为目标,出现的部分服务区域资源浪费和能量损失问题,本文首先引入服务覆盖这一新型网络性能度量,然后针对无蜂窝边缘智能网络提出了一种基于服务覆盖的AP-MEC联合部署方法。通过构建基于元学习的多样化业务需求服务覆盖模型,我们将以每个覆盖单元的最佳服务覆盖为目标来实现系统性能的差异化优化,按需部署网络资源,实现精准的按需服务供给。仿真结果表明,基于服务覆盖的AP-MEC联合部署方法,能够实现系统整体性能的显著提升,同时在不同覆盖单元中实现了精准的按需服务供给,实现高能效、低时延、低能耗的无蜂窝边缘智能网络系统。

在未来的研究过程中,我们将进一步以最佳服务覆盖为目标,提供多场景、多维资源的精准服务,以满足多样化的通信需求,构建高质量、智能化的无蜂窝边缘智能网络系统。

## 参考文献

- [1] LIU G Y, HUANG Y H, LI N, et al. Vision, requirements and network architecture of 6G mobile network beyond 2030 [J]. China communications, 2020, 17(9): 92–104. DOI: 10.23919/jcc.2020.09.008
- [2] SAAD W, BENNIS M, CHEN M Z. A vision of 6G wireless systems: applications, trends, technologies, and open research problems [J]. IEEE network, 2020, 34(3): 134–142. DOI: 10.1109/mnet.001.1900287
- [3] IMT-2030(6G)推进组. 6G典型场景和关键能力 [R]. 2021
- [4] NGO H Q, ASHIKHMIN A, YANG H, et al. Cell-free massive MIMO versus small cells [J]. IEEE transactions on wireless communications, 2017, 16(3): 1834–1850. DOI: 10.1109/twc.2017.2655515
- [5] INTERDONATO G, BJÖRNSON E, QUOC NGO H, et al. Ubiquitous cell-free Massive MIMO communications [J]. EURASIP journal on wireless communications and networking, 2019, 2019(1): 1–13. DOI: 10.1186/s13638-019-1507-0
- [6] MUKHERJEE S, LEE J. Edge computing-enabled cell-free massive MIMO systems [J]. IEEE transactions on wireless communications, 2020, 19(4): 2884–2899. DOI: 10.1109/twc.2020.2968897
- [7] QIN L T, LU H C, WU F. When the user-centric network meets mobile edge computing: challenges and optimization [J]. IEEE communications magazine, 2023, 61(1): 114–120. DOI: 10.1109/mcom.006.2200283
- [8] NAYEBI E, RAO B D. Access point location design in cell-free massive MIMO systems [C]//Proceedings of 2018 52nd Asilomar Conference on Signals, Systems, and Computers. IEEE, 2018: 985–989. DOI: 10.1109/acssc.2018.8645382
- [9] SHEN K, DE PESSEMIER T, MARTENS L, et al. Genetic algorithm combined with ray tracer for optimizing cell-free mMIMO topology in a confined environment [C]//Proceedings of 2021 15th European Conference on Antennas and Propagation (EuCAP). IEEE, 2021: 1–5. DOI: 10.23919/

- eucap51087.2021.9410934
- [10] JIANG J, WANG J C, CHU H Y, et al. Whale swarm reinforcement learning based dynamic cooperation clustering method for cell-free massive MIMO systems [J]. IEEE transactions on vehicular technology, 2023, 72 (3): 4114–4118. DOI: 10.1109/tvt.2022.3222756
- [11] SONG H C, GOLDSTEIN T, YOU X H, et al. Joint channel estimation and data detection in cell-free massive MU-MIMO systems [J]. IEEE transactions on wireless communications, 2022, 21(6): 4068–4084. DOI: 10.1109/twc.2021.3126646
- [12] TUAN H D, NASIR A A, NGO H Q, et al. Scalable user rate and energy-efficiency optimization in cell-free massive MIMO [J]. IEEE transactions on communications, 2022, 70(9): 6050–6065. DOI: 10.1109/tcomm.2022.3194046
- [13] SHENG M, ZHOU D, BAI W G, et al. 6G service coverage with mega satellite constellations [J]. China communications, 2022, 19(1): 64–76. DOI: 10.23919/jcc.2022.01.006
- [14] JIANG J, YAN F Y, YE Y H, et al. Traffic demand-oriented cell-free massive MIMO network [J]. IEEE wireless communications letters, 2023, 12(11): 1861–1865. DOI: 10.1109/lwc.2023.3296527
- [15] KO S W, HAN K F, HUANG K B. Wireless networks for mobile edge computing: spatial modeling and latency analysis [J]. IEEE transactions on wireless communications, 2018, 17(8): 5225–5240. DOI: 10.1109/twc.2018.2840120
- [16] HUI M, CHEN J, ZHOU Y C, et al. Server deployment and load balancing in stochastic mobile edge computing networks [J]. IEEE communications letters, 2022, 26(5): 1194–1198. DOI: 10.1109/comm.2022.3151467
- [17] MARTIN-PEREZ J, COMINARDI L, BERNARDOS C J, et al. Modeling mobile edge computing deployments for low latency multimedia services [J]. IEEE transactions on broadcasting, 2019, 65(2): 464–474. DOI: 10.1109/tbc.2019.2901406
- [18] PARK C, LEE J. Mobile edge computing-enabled heterogeneous networks [J]. IEEE transactions on wireless communications, 2021, 20(2): 1038–1051. DOI: 10.1109/twc.2020.3030178
- [19] SCUTARI G, FACCHINEI F, LAMPARIELLO L. Parallel and distributed methods for constrained nonconvex optimization—part I: theory [J]. IEEE transactions on signal processing, 2017, 65(8): 1929–1944. DOI: 10.1109/tsp.2016.2637317
- [20] TANG A, SUN J X, GONG K. Mobile propagation loss with a low base station antenna for NLOS street microcells in urban area [C]//Proceedings of IEEE VTS 53rd Vehicular Technology Conference, Spring 2001. Proceedings (Cat. No.01CH37202). IEEE, 2001.: 333–336. DOI: 10.1109/vetecs.2001.944859
- [21] MIRAHSAN M, SCHOENEN R, YANIKOMEROGLU H. HetHetNets: heterogeneous traffic distribution in heterogeneous wireless cellular networks [J]. IEEE journal on selected areas in communications, 2015, 33 (10): 2252–2265. DOI: 10.1109/jsac.2015.2435391

## 作者简介



姜静, 西安邮电大学科研处处长、教授, “新一代宽带无线移动通信网”国家科技重大专项评审组专家, 工业和信息化部IMT-2020推进工作组专家, 中国通信学会物联网委员会智慧生态组委员, 陕西创新驱动共同体先进通信技术专委会秘书长; 主要研究领域为Massive MIMO、人工智能无线通信技术等; 先后主持国家自然科学基金、国家重大专项等基金项目10余项; 已发表论文100余篇, 授权专利67项。



乔东伟, 西安邮电大学在读硕士研究生; 主要研究领域为Massive MIMO。



Worakrin SUTHIPHANT, 泰国国家电信移动网络规划经理; 主要研究方向为无线网络规划、资源配置优化等。

# 6G 绿色网络发展趋势和关键技术



## Development Trends and Key Technologies of 6G Green Networks

李福昌/LI Fuchang, 郭希蕊/GUO Xirui

(中国联合网络通信集团有限公司研究院, 中国 北京 100048)  
(Research Institute of China United Network Communications Co., Ltd,  
Beijing 100048, China)

DOI: 10.12142/ZTETJ.202306009

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231211.1113.002.html>

网络出版日期: 2023-12-11

收稿日期: 2023-10-16

**摘要:** 6G网络将驱动无线网络进入万物智联,但也面临能耗巨大的挑战。提出了构建泛在智简、绿能降碳的6G绿色网络的愿景,并分析了6G绿色网络的关键技术问题,包括网络架构和协议智简技术、先进天线技术、智能资源管理技术和动态关断技术等。介绍了实现6G绿色低碳设备的一系列技术,包括高效芯片技术、高效信号处理技术、高效功放技术以及新型材料技术,并针对6G网络与绿色能源融合技术以及分布式微电网协同技术进行了分析。

**关键词:** 绿色网络; 绿色设备; 绿色站点; 绿色能源

**Abstract:** The 6G network will drive wireless networks into the era of intelligent connection of all things, but it also faces the challenge of huge energy consumption. The vision of building a 6G green network that is ubiquitous, intelligent, simple, green, and carbon-reducing is proposed, and the key technical issues of the 6G green network are analyzed, including network architecture and protocol intelligent simplification technology, advanced antenna technology, intelligent resource management technology, and dynamic shutdown technology. A series of technologies for realizing 6G green and low-carbon devices are introduced, including efficient chip technology, efficient signal processing technology, efficient power amplifier technology, and new material technology. Finally, the integration technology of the 6G network and green energy and the collaborative technology of distributed microgrid are analyzed.

**Keywords:** green network; green equipment; green site; green energy

**引用格式:** 李福昌, 郭希蕊. 6G绿色网络发展趋势和关键技术 [J]. 中兴通讯技术, 2023, 29(6): 54–59. DOI: 10.12142/ZTETJ.202306009

**Citation:** LI F C, GUO X R. Development trends and key technologies of 6G green networks [J]. ZTE technology journal, 2023, 29(6): 54–59.  
DOI: 10.12142/ZTETJ.202306009

随着移动通信近10年的快速发展,“4G改变生活,5G改变社会”带动了数字经济的高速发展,5G更是作为新型数字基础设施实施的重要技术在网络强国、数字中国、智慧社会中发挥了重大作用。基于移动通信发展红利,人民对于6G技术愿景充满了更高的期待。6G通信网络面向用户,将更加智能化、数字化,为工业互联网和物联网提供可靠的网络通信。

在移动通信技术发展历程中,虽然网络设备能效不断提升,但随着网络频率升高、带宽增大、规模增长,网络能耗呈现持续增长的趋势,未来6G网络面临严重的能耗挑战。如何在网络服务能力持续提升的基础上实现碳达峰、碳中和是6G网络发展中的关键问题。本文从绿色网络技术、绿色低碳设备、绿色站点技术、绿色能源技术等角度提出了构建泛在智简、绿能降碳的6G绿色网络<sup>[1]</sup>。

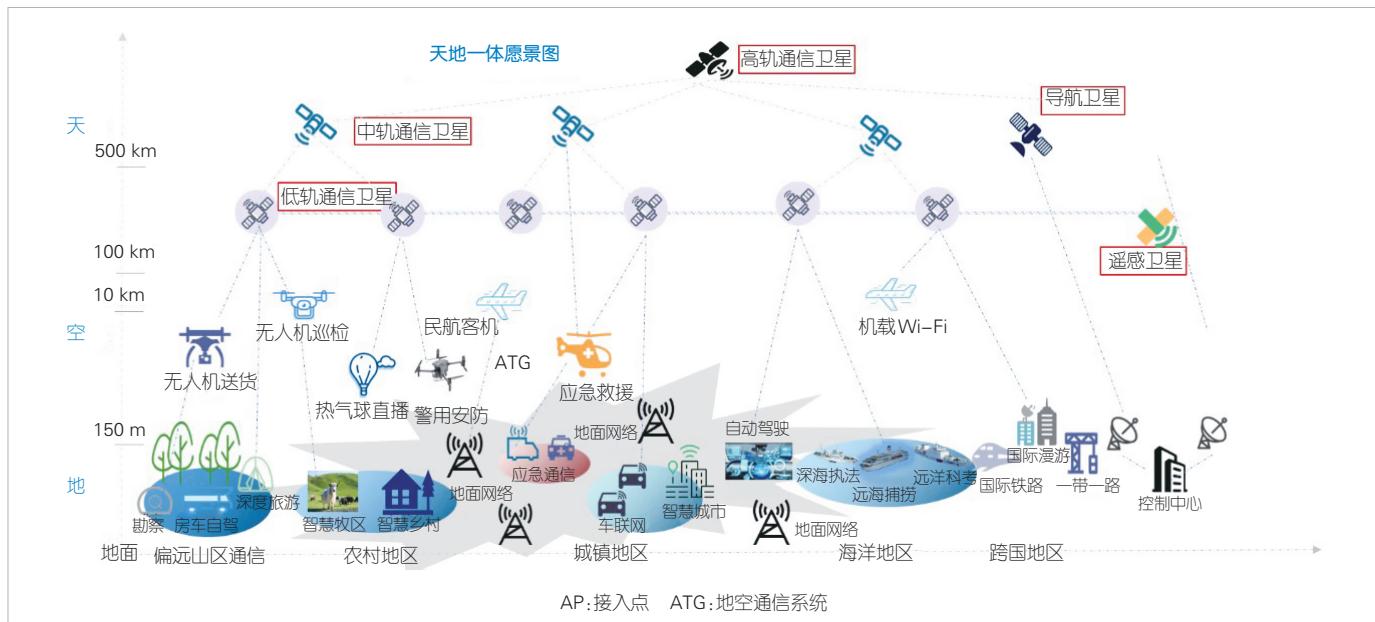
## 1 6G绿色网络技术

6G网络覆盖涵盖空天地海等全球全域场景,网络架构如图1所示。6G网络需要在“泛在连接”的基础上通过架构智简、协议智简、组网智简构建绿色低碳的网络体系<sup>[2]</sup>。

### 1.1 架构智简技术

为了实现万物智联、绿色低碳的发展目标,6G网络架构将发生颠覆式重构。无线网将打破传统的有边界的、烟囱式的架构,支持融合泛在、绿色节能的新型网络架构。

6G绿色网络架构将支持空天地一体化组网,通过卫星网络、空基网络、低空及地面网络3层组网,形成以地面网络为基础、以非地面网络为拓展的立体全域覆盖绿色节能网络。地面网络与非地面网络互联互通、深度融合,采用统一的协议栈,支持海量用户无感知、极简的泛在接入。基于内



▲图1 6G绿色网络架构

生智能，6G网络将充分利用网络节点的通信、计算和感知能力，通过分布式学习、群智式协同，实现端到端智能编排调度、跨域智能管控，保障用户为中心的业务感知和最优的网络能效。

6G绿色网络架构中地面与非地面网络层可采用超蜂窝、无蜂窝等符合绿色通信发展趋势的新型组网方式。

在超蜂窝架构下，基站控制面与用户面解耦，控制基站与业务基站可以独立按需部署。前者提供用户接入以及控制信号的传递，可采用大区覆盖模式；后者为用户提供高速数据传输，可按需灵活部署。同一个控制基站覆盖范围内可以部署多个业务基站，且业务基站可根据业务负荷变化动态休眠。在该架构下，网络覆盖可跟随业务需求动态调整，在不影响覆盖性能的前提下通过控制业务基站适时进入休眠状态，实现更为灵活的休眠，提升网络节能效果。

无蜂窝架构以用户为中心，部署多个分布式接入点以及一个与所有接入点相连接的中央处理单元。通过中央处理单元的集中信号处理，广泛分布的接入点可以实现高水平的协作，形成一个“超级基站”覆盖整个区域。每个用户接入一组特定的接入点，可以利用空间宏分集和低路径损耗提升网络的频谱效率和能量效率。当区域内用户较少时，可以关断部分接入点，进一步节省系统能耗。

## 1.2 协议智简技术

协议智简是通过极简信令优化降低信令开销实现能耗的降低，主要方法包括优化信令接口，减少冗余信息，采用压

缩编码技术等。协议智简技术有助于提高通信系统的效率和稳定性，同时降低能耗和资源占用。

在小区低负荷时，减少公共信号的传输次数，可以使小区进入关断或休眠的节能状态，降低系统能耗。

在多载波场景下，基于多载波之间的协同，可以将多个载波的系统消息汇聚在一个载波上传输。不发送系统消息的载波则可在低负荷时关断，在降低公共信号开销的同时降低能耗。

面向海量连接需求，6G绿色网络结合多域多用户接收技术、稀疏导频技术等，使终端无须复杂信令交互即可发起传输，适时进入休眠状态，从而实现零信令开销、低功耗的极简传输。

## 1.3 先进的天线技术

6G网络将结合超大规模多输入多输出（MIMO）、分布式MIMO、智能超表面等先进的天线技术和设备实现网络智能高效精准的网络覆盖，提升网络覆盖效能。

### 1) 超大规模MIMO技术

MIMO技术作为提升频谱效率最有效的技术手段，已在现有网络中得到了广泛应用，基站的收发通道数从3G支持的1T1R扩展为5G设备的64T64R。未来6G网络随着频段升高，电磁波波长也变短，将采用超大规模的天线阵列，可以产生宽度更小、能量更加集中的波束，并精确指向目标用户，以应对超高频带来的大路损。这不仅提升了覆盖和容量，而且大幅度减少了每比特能耗。

## 2) 分布式 MIMO 技术

分布式 MIMO<sup>[3-4]</sup>支持以用户为中心的智能接入网架构，利用多天线带来的空间自由度，及对信号空时频特性的联合优化，增强通信系统干扰抑制能力，使用户获得良好的信道条件，更好地保障业务质量在地理分布上的均衡，实现用户体验速率、边缘覆盖性能的明显改善，同时实现设备能效提升。

## 3) 智能超表面技术

可重构的智能超表面（RIS）技术<sup>[5-6]</sup>采用了可编程的新型电磁散射材料，通过调控超表面天线单元的相位、幅度或极化方向，大幅改善电磁波传播环境，从而达到扩展无线信号覆盖，提升小区吞吐量的效果<sup>[7]</sup>。

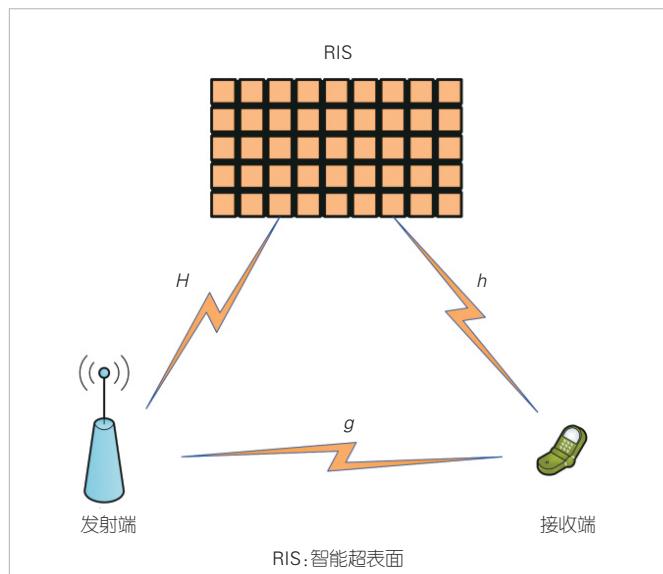
图2为智能超表面辅助的通信系统模型，其中接收端接收到的信号 $y$ 为：

$$y = \sqrt{\beta} (h^H H + g)s + n。 \quad (1)$$

从发射端经过 RIS 到达接收端的等效信道  $h^H H$  为 RIS 与接收端间信道  $h$ 、RIS 的可调相移对角矩阵  $H$  以及发送端与 RIS 间信道  $H$  的乘积， $g$  为接收端和发射端之间的直达信道。 $s$  是发送端发送的信号， $n$  为高斯白噪声。当使用 RIS 辅助通信时，RIS 单元反射的信号可以表示为入射信号与该单元反射系数的乘积。

RIS 技术的特点是以低功耗的方式智能调控大规模低成本的无源天线单元，从而重构无线环境，实现更佳的覆盖、更优的用户体验。

## 4) 轨道角动量技术



▲图2 智能超表面辅助的通信系统模型

轨道角动量（OAM）技术<sup>[8]</sup>可以产生具有不同 OAM 状态的电磁波，为移动通信系统提供了一种新型信道复用技术。利用不同模态数的电磁涡旋波间的正交性，可以在同一载波上将信息加载到具有不同轨道角动量的电磁波上，提高系统容量和提高频谱利用率。系统容量可以写为电场强度信道容量  $C^{(E)}$  和 OAM 信道容量  $C^{(O)}$  之和<sup>[9-10]</sup>即：

$$C = C^{(E)} + C^{(O)} = \sum_{n=1}^N B_n \log(1 + n_{\text{SNR},E}^{(n)}) + \sum_{m=1}^M W_m \log(1 + n_{\text{SNR},O}^{(m)}), \quad (2)$$

其中， $N$  和  $M$  分别为电场强度和 OAM 的信道个数， $B_n$ （或  $W_m$ ）分别为第  $n$ （或  $m$ ）个信道的电场强度带宽（或 OAM 带宽）， $n_{\text{SNR},E}^{(n)}$ （或  $n_{\text{SNR},O}^{(m)}$ ）分别为第  $n$ （或  $m$ ）个信道的信噪比。

## 1.4 智能资源管理技术

无线网络的能耗与资源使用效率、业务负荷变化紧密相关。如何在满足用户服务质量要求的前提下通过智能资源管理提高网络资源利用率，是6G绿色网络构建需要重点关注的问题。随着大数据、人工智能（AI）等技术的快速发展，网络资源管理将更加实时化、智能化，资源随业务负荷变化灵活调整，达到提高网络资源和能源利用率的目标。

在无线设备层面，借助 AI 技术，可以在网络业务量、资源使用率、能耗等指标感知与评估的基础上，基于业务精准预测，对空域、频域、时域等多域资源进行智能化协同调度，实现以用户为中心的资源与能源灵活配置，达到能耗、网络性能、业务体验的最佳平衡。例如，时分复用（TDM）帧结构是一种可以实现网络侧和终端的节能帧结构设计，AI 可以根据实时的网络状态和用户需求，自动调整使用合适的帧结构等。此外，在无蜂窝、超蜂窝绿色网络架构下，每个分布式基站可实时监测网络状态，控制基站或中央处理单元对多个基站或接入点进行协同的资源调度，形成以用户为中心的覆盖与资源配置，充分利用网络资源，降低能耗。例如，通过 AI 技术可以对计算任务进行分析和预测，确定哪些任务可以被卸载，并将计算任务卸载到合适的设备上。根据设备的能耗和性能等因素，AI 还可以动态调整计算任务的分配，以实现最佳的节能效果。

在站点基础设施层面，需要基于能源管理数字化转型，将能源基础设施打造成一张数字化和智能化的网络。一方面，基于大数据技术实现自动开站、自动联网、智能运维、全生命周期资源管理，提升能效和运维效率；另一方面，根据能耗、碳排放数据以及环境数据的实时感知，智能控制电源、空调等设施的工作状态，节省能源消耗。

此外，AI技术与数字孪生等技术结合，可以实现数据驱动的决策与优化。孪生网络可以模拟网络的服务、环境、负载和网络设备的性能、能耗情况，实时监测网络的运行状态和能耗情况。同时，数字孪生可以对网络进行模拟优化，为AI提供参数输入和决策支持，从而帮助AI提出更有效的节能方案。

### 1.5 基于L1/L2的新型动态关断节能技术

目前，传统的节能关断技术主要是基于高层和半静态的负荷预测和用户连接数预测，存在省电状态切换慢、基于历史负荷而非当前实时负荷的预测等问题，无法在不牺牲容量的条件下大幅度降低基站功耗。

新型动态关断技术主要基于基站的能效感知的动态调度器<sup>[11]</sup>，可以根据实时负荷水平、无线信道质量和实时可用资源做最优调度，实现时域关断、通道关断和下行功率控制毫秒级的快速状态切换，可以在几乎不牺牲容量的条件下大幅度降低基站功耗。动态通道关断示意如图3所示。

## 2 绿色低碳设备

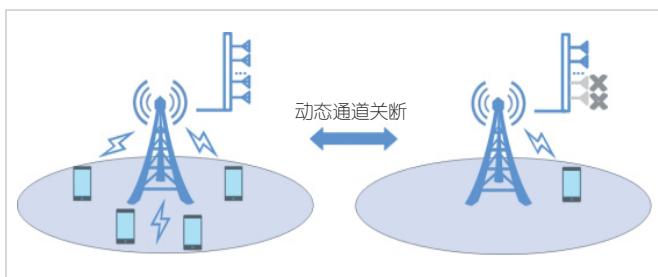
未来6G网络，在信号瞬时带宽方面将会超过1 GHz以上，工作频率也会达到100 GHz频段甚至太赫兹。针对高频大带宽设备，芯片技术、信号处理技术和器件在能耗方面将面临诸多新问题和新挑战。

### 2.1 高效芯片技术

6G网络将驱动无线网络进入万物智联，基站系统要实时处理达到Tbit/s的数据流量。这对芯片性能提出了更高的要求，芯片的能耗将成倍提升。下文中通过对芯片制程、封装工艺、异构芯片等技术研究，我们探索芯片的节能降耗技术。

#### 1) 先进制程

制程节点越小意味着晶体管越小，速度越快，那么能耗表现就越好。超高集成度的6G基带及数字中频芯片，配合



▲图3 动态关断示意图

使用超高集成度的射频前端，可以全面降低6G基站的整机功耗。

#### 2) 基于Chiplet的芯片封装集成技术

采用跨工艺、跨封装的Chiplet多芯粒互联技术和面向通用小芯片互连通道（UCIE）标准的Chiplet接口IP设计等先进技术，可以实现更高的系统集成度，在同等单位面积硅片上集成更多数量的晶体管，提升晶体管密度，实现芯片的节能。

#### 3) 专用芯片和通用芯片技术

采用领域定制（DSA）的异构计算，将专用芯片和通用芯片进行融合，并采用中央处理器（CPU）/通用处理器（GPU）/网络处理器（NPU）/现场可编程门阵列（FPGA）/专用集成芯片（ASIC）来完成对特定业务的加速，从而达到提升系统算力的目的。

### 2.2 高效信号处理技术

在6G网络高能目标牵引下，6G系统的基带处理、数字中频、射频信号算法处理需要进行新的设计和优化。

在信道编码方面，新的6G信道编码可以考虑具有内在并行特征的低密度奇偶校验（LDPC）码，降低信道编码的复杂度，同时考虑在牺牲少量性能的条件下追求吞吐量、功耗、灵活性、成本的综合最优，实现最佳的基带芯片节能效果。

在采样带宽方面，欠采样条件下的射频算法处理可以缓解大带宽对模数转换器（ADC）采样芯片造成的功耗压力。但同时要注意欠采样算法模型和计算复杂度不能过高，否则同样会带来算法处理芯片的功耗压力，因此需要在射频器件能效提升和基带处理芯片功耗提升之间做折中考虑。

在降低射频信号峰均比方面，我们可以采用以下两种技术：一是直接设计低峰均比信号，例如，通过改变信号调制方式或帧结构参数，使得处理信号本身就具有低峰均比特性；二是基于传统的峰值因子消减（CFR）算法向大带宽方向进行技术迭代和优化。

在数字预失真（DPD）算法方面，我们采用6G超大带宽和多频段的DPD算法，提升功率放大器（PA）的线性度，从而提升功放效率。同时，还需要解决由相控阵波束成形带来的多个并行PA使用单个DPD算法进行线性化的问题。

在物理层技术增强方面，6G网络还可采用极化编码、概率成形调制、基于免调度的非正交多址接入以及全双工等潜在的新技术，提升系统频谱效率，降低接入开销，在满足6G全业务场景、全类型终端的接入需求的同时提升系统能效。

另外，通信与感知融合技术和反向散射技术也可以提升6G网络的能效。通信与感知的融合使得无线系统可以深度感知无线环境，优化/简化信道测量和干扰测量，从而提升通信的能效。反向散射技术可以有效降低数据传输功耗。终端还可以通过能量采集技术从环境的能量源中收集能量，满足近零功耗终端的通信需求。

### 2.3 高效功放技术

有统计指出<sup>[12]</sup>，无线通信系统近40%功耗来自PA。目前业界采用第3代半导体氮化镓（GaN）技术解决了5G大带宽、高频率的问题，但随着毫米波、太赫兹技术的发展与应用，需要继续提升GaN器件的功放效率。可以通过包络跟踪技术以及Doherty功放+异相Outphasing<sup>[13]</sup>发射机技术来提升末端功放效率。在6G PA材料方面，GaN、锗化硅（SiGe）、磷化铟（InP）等材料的使用可以有效提高功率放大器的功率和效率。

### 2.4 新型材料技术

对于6G超大规模天线的MIMO，新型材料提供了良好的宽带性能和卓越的功率密度及功耗效率，满足严格的热规范，同时为紧密集成的超大规模天线阵列节省了宝贵的芯片空间。

低损绿色天线<sup>[14]</sup>通过低损耗新型天线罩、塑料金属化新型阵子、一体化低损耗移相器、免电缆天线设计等技术提高天线电磁转换效能，提高天线覆盖增益，实现相同覆盖范围内发射功率的降低，从而实现整网能效的提升。

高效散热材料采用更高效的散热技术，可降低设备工作温度，从而提升芯片工作能效。目前业界正在开发全新的高导热压铸铝合金材料。该材料较传统压铸铝合金导热系数提升30%，同等产品散热能力提升5%。有源天线处理单元（AAU）及射频拉远单元（RRU）设备散热能力提升可使处理器及功放等器件工作于更低能耗的状态，从而提升设备能效。

## 3 绿色站点技术

### 3.1 站点级节能

为有效降低6G网络配套设施能耗与碳排放，大幅减少基站机房及配套设备的建维成本，我们需重构传统建维体系，从系统架构、供电系统、空调系统、绿色能源、智能维等多个方面着手，提升站点基础设施整体能效和碳效。

1) 构建极简系统架构。坚持“极简建设+低碳高效”

发展模式，基于集成化、标准化、预制化、模块化等核心要素，以智能供电系统、多模式空调系统、新能源及备储一体系统等打造新型系统架构，满足6G网络室内/室外多场景低碳、灵活部署的需求。

2) 提高供电系统效能。聚焦供电设备器件级、设备级、系统级的能效提升，推动氢燃料电池、固态电池及备储一体等新型电池技术落地，加速6G网络能源清洁化，实现电业技术协同，并结合网络级、业务级能耗/碳数据精准采集，全面提升供电系统效能。

3) 提升散热能力。6G站点应因地制宜地充分利用自然冷源，引入AI，实现空调运行状态的智能控制，使空调系统始终运行在最佳能效区间；面向6G设备高功率需求，推动液冷技术在6G站点的应用，提升站点整体散热能力。

4) 打造智能建维。基于动态资源管理技术，构建、运行、维护、再分配及报废的全生命期的建维体系。引入数字孪生技术，以数字化、智能化搭建建维技术体系，实现与新型基础设施结合、自主智能化运行和故障的自我维护等，最大限度降低6G网络建维成本。

### 3.2 数据中心节能

数据中心的绿色节能低碳发展不再单纯追求极低的电能使用效率（PUE），需要综合考虑可再生能源、算力、信息通信（IT）设备能耗、制冷及供配电设备能效等多方面关键因素，重点关注各系统的高效协同工作，并通过高效制冷、余热回收、多维度能效提升、AI智能控制等节能新技术的应用，配合精细化的运行维护管理，最终实现数据中心的绿色节能。

## 4 绿色能源技术

随着6G的到来，绿色能源相关产业逐渐成熟，将成为基站系统供电的主要来源之一。6G网络技术将推动绿色能源系统优化配置、控制管理、智能运维等方面的能力提升，增强绿色能源的发电效率、稳定性和安全性。从6G网络能源供给“开源”入手，丰富基础设施的绿色化转型模式，力争打造电力算力协同、资源利用充分、能效优化合理、运行稳定可靠的6G网络绿色能源供给体系。

### 4.1 6G设备与绿色能源深度融合技术

为解决6G网络耗电量大、增速快、碳排放量高的问题，6G网络需要提高清洁能源供给效率，因此可在6G前端设备中融合光伏电池、新型储能电池、热电转换器等技术，并内置传感器采集绿色能源相关的参数数据，如光照、风速、温

度等，增强6G网络对绿色能源的感知与调节能力，提升6G设备清洁能源的就地消纳能力，减少6G设备高碳排能源输入占比，直接降低6G设备运营阶段碳排放。

## 4.2 6G网络与分布式微电网协同技术

6G网络时代，会出现大量分布式形态的算力资源。为持续推进节能降碳，我们需要构建相适应的高度分布式形态的绿色能源网络，因此需要将6G网络与绿色能源物联网连接，形成通信-电力系统能量平衡优化机制，提供通信与电力调度需求耦合的通信网-电力网互补支撑，在边缘侧实现源-网-荷-储灵活互动以及海量分布式资源的协同运行，形成电力网络与通信网络的灵活性共享机制，实现跨网的资源优化配置，达到大幅降低6G通信系统碳排放效果，从而提升6G网络与电力系统的能效与碳效。

## 5 结束语

2030年前实现碳达峰、2060年前实现碳中和的双碳目标与中国6G技术与网络发展时间正好契合。本文中，我们提出了在网络架构、关键技术、协议简化、绿色设备、绿能融合等方面的技术演进方向，打造绿色低碳的6G绿色网络。未来，我们希望6G网络碳排放强度（单位信息流量碳排放）相比2020年下降90%以上，为2030年实现通信领域碳达峰夯实基础。

## 参考文献

- [1] IMT-2030(6G)推进组. 6G总体愿景与潜在关键技术白皮书 [R]. 2021
- [2] 中国移动研究院. 2030+网络架构展望(2020年) [R]. 2020
- [3] 尤肖虎, 王东明, 王江舟. 分布式MIMO与无蜂窝移动通信 [M]. 北京: 科学出版社, 2019
- [4] WANG D M, ZHANG C, DU Y Q, et al. Implementation of a cloud-based cell-free distributed massive MIMO system [J]. IEEE communications magazine, 2020, 58(8): 61–67. DOI: 10.1109/mcom.001.2000106
- [5] 智能超表面技术联盟(RISTA). 智能超表面技术白皮书 [R]. 2023
- [6] ZHANG L, CHEN MING ZHENG, TANG W K, et al. A wireless communication

scheme based on space- and frequency-division multiplexing using digital metasurfaces [J]. Nature electronics, 2021, 4(3): 218–227. DOI: 10.1038/s41928-021-00554-4

- [7] IMT-2030(6G)推进组. 智能超表面技术研究报告 [R]. 2021
- [8] 杨航, 郑史烈, 张红旗, 等. 太赫兹轨道角动量通信关键技术与挑战 [J]. 移动通信, 2023, 47(5): 20–25. DOI: 10.3969/j.issn.1006-1010.20230322-0002
- [9] 张超, 王元赫. 论涡旋电磁波轨道角动量传输新维度 [J]. 通信学报, 2022, 43(6): 211–222. DOI: 10.11959/j.issn.1000-436x.2022087
- [10] WANG Y H, ZHANG C. Enhanced Shannon capacity with orbital angular momentum dimension [C]//Proceedings of 2022 IEEE International Conference on Communications Workshops (ICC Workshops). IEEE, 2022: 1–6. DOI: 10.1109/iccworkshops53468.2022.9814548
- [11] BORYLO P, TORNATORE M, JAGLARZ P, et al. Latency and energy-aware provisioning of network slices in cloud networks [J]. Computer communications, 2020, 157: 1–19. DOI: 10.1016/j.comcom.2020.03.050
- [12] GSMA Intelligence. Going green: benchmarking the energy efficiency of mobile [EB/OL]. [2023-10-13]. <https://data.gsmaintelligence.com/api-web/v2/research-file-download?file=300621-Going-Green-efficiency-mobile.pdf>
- [13] 陈艺. 宽带高回退Doherty-Outphasing功放的研究 [D]. 杭州: 杭州电子科技大学, 2020
- [14] 孔善球, 林学进, 王昊, 等. “双碳战略”给移动通信天线产业带来的机遇和挑战 [J]. 移动通信, 2022, 46(3): 99–102. DOI: 10.3969/j.issn.1006-1010.2022.03.017

## 作者简介



李福昌, 中国联通研究院教授级高级工程师; 主要从事移动通信及固网移动融合等标准制定、测试验证、课题研究等工作; 发表论文100余篇。



郭希蕊, 中国联通研究院高级工程师; 主要从事无线通信研究相关的工作; 发表论文10余篇。

# 零信任关键技术与产业发展研究



## Key Technologies and Industrial Development of Zero Trust

张云畅/ZHANG Yunchang, 柴瑶琳/CHAI Yaolin,  
穆瑮博/MU Yubo  
(中国信息通信研究院, 中国 北京 100083)  
(China Academy of Information and Communications Technology, Beijing 100083, China)

DOI: 10.12142/ZTETJ.202306010

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231212.1442.004.html>

网络出版日期: 2023-12-13

收稿日期: 2023-10-16

**摘要:** 零信任架构作为一种网络安全的新理念、新架构、新技术，基于“持续验证，永不信任”的核心思想，通过融合软件定义边界、身份识别与访问管理、微隔离三大技术，将重塑现有网络安全架构和网络安全设施，并深刻改变关键基础设施的部署与应用模式，带来网络安全领域的一场新变革。从零信任演进与技术发展情况入手，分析零信任当前产业发展现状，通过对比全球零信任战略部署、技术创新、产业规划研提中国发展建议。同时剖析中国零信任发展中的问题，研提中国在零信任顶层设计、技术革新发展、产业生态闭环的针对性建议。

**关键词:** 零信任；零信任架构；网络安全战略

**Abstract:** As a new concept, new architecture, and new technology of network security, zero trust architecture is based on the idea of "continuous verification, never trust". By integrating software-defined borders, identity and access management, and micro-isolation technologies, zero trust architecture will reshape the existing network security architecture and network security facilities, and furthermore, profoundly change the deployment and application method of key infrastructure, leading to a new revolution in the field of network security. The current status of the zero trust industry from the perspective of zero trust evolution and its technological development is analyzed, and the suggestions on further development are provided by comparing global zero trust strategic deployment, technological innovation, and industrial planning. At the same time, the problems in the development of zero trust in China are discussed. Finally, originating from the micro level, targeted suggestions on the top-level design of zero trust, technological innovation and development, and closed-loop industrial ecology are proposed.

**Keywords:** zero trust; zero trust architecture; cyber security strategy

**引用格式:** 张云畅, 柴瑶琳, 穆瑮博. 零信任关键技术与产业发展研究 [J]. 中兴通讯技术, 2023, 29(6): 60–65. DOI: 10.12142/ZTETJ.202306010

**Citation:** ZHANG Y C, CHAI Y L, MU Y B. Key technologies and industrial development of zero trust [J]. ZTE technology journal, 2023, 29(6): 60–65. DOI: 10.12142/ZTETJ.202306010

网络安全是产业数字化升级中的关键保障。近10年来层出不穷的网络安全问题不仅阻碍了各行业发展，更是减缓了产业革新的速度，迫使全球经济发展受到制约。零信任作为破解当前困境的强有力手段，在其理念形成的初期便受到了业界的广泛关注。各国高度重视零信任领域的战略布局和创新研究，抢抓国际新技术主导权。中国也在不断推进零信任领域的技术创新应用。如何借助零信任重塑网络安全架构，已成为中国网络安全领域的研究热点。

本文将围绕总体研究、核心技术、产业发展3个方面，梳理零信任的发展态势，深入解析当前中国零信任产业发展存在的问题及其演进趋势，并提出发展建议。

## 1 零信任整体发展态势

### 1.1 国际零信任战略部署加快

零信任成为美国重塑政府整体网络安全架构的重要手段。2021年5月12日，美国发布总统行政命令<sup>[1]</sup>“（强制）要求联邦机构制定零信任安全架构的实施计划”。2021年9月7日，美国公布《推动美国政府朝向零信任网络安全演进的基本准则》<sup>[2]</sup>，明确指出各级政府部署零信任安全架构的时间表和发展目标。2021年11月，美国正式推出以零信任实施方案为核心的Thunderdome<sup>[3]</sup>（雷霆穹顶）项目，这标志着全球首个零信任国家级战略部署计划正式启动。2022年，美国白宫管理与预算办公室（OMB）发布备忘录（编

号为 M-22-09)<sup>[4]</sup>, 要求所有行政部门于 2024 年实现零信任全面部署。此外, 美国 2023 财年年度预算<sup>[5]</sup>显示, 总额 8133 亿美元的年度预算中将有 112 亿美元用于网络战建设, 其中包括“实施”部门的零信任架构。2024 财年网络安全预算备忘录<sup>[6]</sup>再次将零信任在联邦民事行政部门机构的实施升为第一优先级。

新加坡<sup>[7]</sup>、加拿大<sup>[8]</sup>等在美国一系列零信任战略规划文件出台后快速做出响应。2021 年 4 月 12 日, 加拿大公布的《网络与安全战略》<sup>[8]</sup>详细介绍了零信任的概念及架构, 并指出其对未来网络服务和安全的支撑性作用。新加坡于 2021 年 10 月 5 日宣布发布《网络安全战略 2021》<sup>[7]</sup>以强调零信任应用的必要性, 同时明确将积极在全国范围推行零信任落地实践。此外, 欧盟于 2022 年 3 月 22 日公布《为联盟的机构、机关、办公室和机构制定高水平的网络安全措施》<sup>[9]</sup>, 指出要明确朝向零信任架构迈进的具体步骤。根据多家第三方机构发布的数据, 日本、澳大利亚、印度及一些中东国家均已在本土企业开展零信任模型的部署应用, 并持续关注该领域的发展。

## 1.2 全球 SASE 技术标准体系不断完善

零信任技术发展趋近成熟, 国际化组织标准建设进程不断加快。2020 年 8 月, 美国国家标准与技术研究院 (NIST) 加紧推动零信任架构研究, 经过多轮修订发布了 SP 800-207《零信任架构》<sup>[10]</sup>。2021 年 9 月, 网络安全基础设施安全局 (CISA) 发布《零信任成熟度模型》征求意见稿<sup>[11]</sup>, 这标志着一种针对军政商三方零信任架构成熟度的评测方案正式形成。

全球各国/地区针对零信任领域的标准制定速度不断加快。作为一个权威的国际产业组织, 云安全联盟 (CSA) 率先针对零信任三大技术之一的软件定义边界, 进行了标准化研究, 并在 2014 年发布《SDP 标准规范 V1.0》<sup>[12]</sup>, 接着在 2022 年发布《软件定义边界 (SDP) 标准规范 V2.0》<sup>[13]</sup>。与此同时, 中国也在产业各方的推动下于 2022 年形成《零信任能力成熟度模型》, 以指导中国零信任网络安全架构的落地。

## 1.3 零信任产业全面高速发展

全球零信任产业生态已初步形成, 供需双方均在积极推进规模化部署进程。以谷歌、微软、思科为代表的美国网络安全领军企业对零信任市场进行了全面布局, 与美国政府联动引领产业活力化发展。以电信、金融、能源为首的各垂直行业也在积极采用零信任网络安全架构, 从试点

试验到大面积落地部署应用, 逐步从需方立场支撑零信任产业的发展。

Gartner<sup>[14]</sup>在 2020 年曾预测, 2022 年在面向生态合作伙伴开放的新型数字业务应用程序中, 80% 的业务将通过零信任网络访问 (ZTNA) 进行访问。据 IDC 预测, 到 2024 年, 安全远程访问解决方案将以 260 亿美元价值占据全球网络安全市场 12.5% 的份额<sup>[15]</sup>, 其中零信任相关产品和解决方案将占据重要地位。2022 年 6 月, Markets and Markets<sup>[16]</sup>在到 2027 年的全球预测报告中表示: 全球零信任证券市场规模预计从 2022 年的 274 亿美元增长到 2027 年的 607 亿美元。

## 2 零信任网络安全架构及关键技术

### 2.1 零信任网络安全架构

零信任是一种新型网络安全理念, 秉承着“持续验证, 永不信任”的原则, 拒绝隐式授予信任, 持续进行安全评估, 其体系架构是一种端到端的企业资源和数据安全方法, 包括身份 (人类和非人类的实体)、凭证、访问管理、操作、端点、宿主环境和基础设施, 旨在通过利用网络分段、防止横向移动、提供第 7 层威胁预防和简化精细用户访问控制来保护现代数字环境。因此, 零信任总体功能架构围绕技术和管理两个维度, 涵盖八大关键能力, 即身份安全、基础设施、网络安全、数据安全、应用/负载安全、网络可持续安全检测、评估和网络安全可视化以及综合安全管理, 如图 1 所示。

**身份安全:** 身份 (人员、设备、应用、进程) 是资源访问的入口, 是零信任的基础。

**网络安全:** 建立可信、可靠的网络链路是数据访问的重要环节。

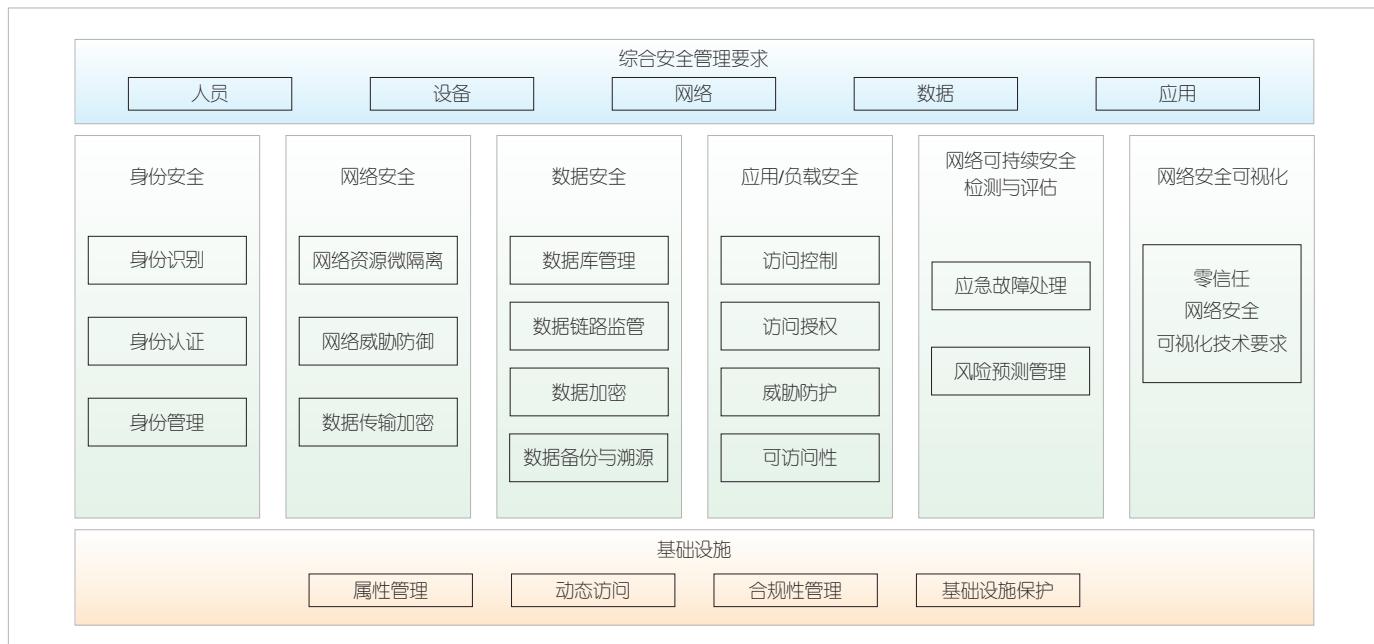
**数据安全:** 数据是安全的核心, 应被分类、标记和加密, 并基于这些属性有条件地访问。

**应用/负载安全:** 应用程序和应用程序编程接口 (API) 提供了数据访问接口, 同时负载安全保障了资源交互的稳定性。

**基础设施:** 建立安全的基础设施 (本地服务器、云端虚拟机、容器、微服务) 是减少风险的有效措施。

**网络可持续安全检测与评估:** 网络可持续安全检测与评估是零信任架构实施的保障, 其着眼于应急事件处理和风险管理两个维度, 形成流程化处理模式。

**网络安全可视化:** 网络安全可视化是零信任架构的坚实基础, 其借助终端、应用、行为和事件这四大主体实现。



▲图1 零信任总体功能架构图

安全管理：安全管理是网络安全架构部署实践的保障。

## 2.2 零信任关键技术

零信任的关键技术包括软件定义边界、身份识别与访问管理、微隔离。

### 1) 软件定义边界

作为零信任理念的实践方案，软件定义边界实现了对基于网络的攻击行为的阻断，分离了控制平台与数据平台，最小化攻击面；通过预验证、预授权，拒绝未经验证授权业务的端口访问；隐藏关键资产的同时对应用访问实现可视化，实现实时监督；可集成安全架构，大幅提升产品的兼容性；可代替部分传统网络安全的人力资源，节约成本。

### 2) 身份识别与访问管理

作为零信任模型的应用基础，身份识别与访问管理是企业应用零信任的第一步，同时也是零信任项目取得信任的关键一步。身份识别与访问管理涵盖了用户身份、规则、身份验证管理软件以及访问管理策略和协议，实现了持续的动态认证与动态授权，融合了多因子身份认证、单点登录和用户行为分析，提供了静态密码以外的针对用户凭据的安全保护，对相关但独立的多个系统实施一种访问机制，可检查用户行为并自动应用算法和分析，以监测潜在的安全威胁。

### 3) 微隔离

作为零信任概念的技术实践，微隔离实现了数据中心内

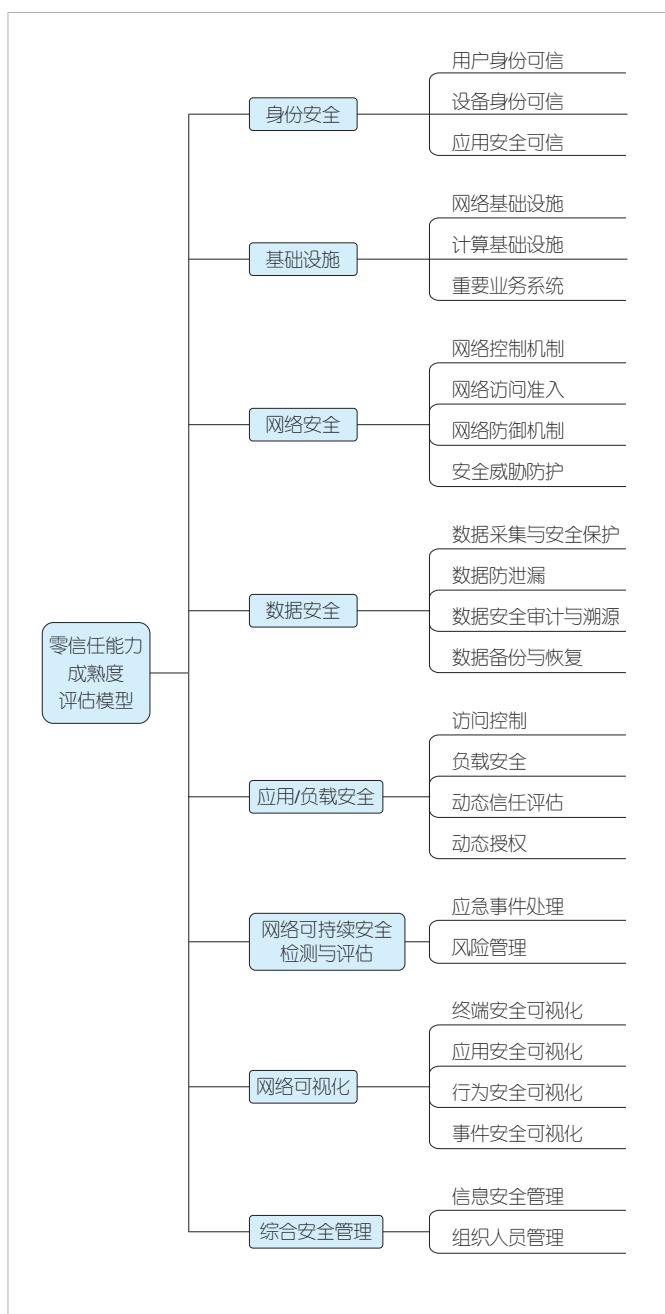
工作负载间流量可视化与访问控制，在保留传统防火墙的安全边界防御的同时，限制工作负载流量间通信。攻击面的可视化加强了应用程序活动的管理细粒度。无须基于硬件的防火墙，微隔离可将安全性集成到虚拟化工作负载中，其安全策略与虚拟网络（虚拟机、操作系统或者其他安全目标）同步，实现纯软件方式的安全模型部署，改善监管合规态势，同时隔离信息技术（IT）基础设施，在混合云模式下支持跨平台的数据流量识别及统一管理。

### 4) 零信任能力成熟度评估模型

本文中，我们参考《零信任能力成熟度模型》标准，遵循零信任的定义和原则，依据零信任总体功能架构，提出零信任能力成熟度评估模型，将零信任总体架构细分为8个技术模块与26个能力域，形成零信任关键能力图谱，如图2所示。

为使零信任能力成熟度判定标准详细且规范化，本文将零信任能力成熟度划分为无零信任阶段、传统阶段、初级阶段、优化阶段和持续安全阶段五大成熟度等级，旨在以全局或专业领域视角从零信任核心功能子组件、技术交互方式等方面展示零信任安全能力成熟度，助力零信任技术优化升级，如表1所示。

以身份安全技术模块为例，我们对该模块下四大能力域涵盖的能力要点进行量化打分。用户身份可信域的计算公式为：用户身份可信子领域4级能力符合度=（达到4级的安全能力数量+达到更高等级的安全能力数量）/用户身份可信



▲图2 零信任关键能力图谱

▼表1 零信任能力成熟度五大阶段

零信任能力成熟度阶段	阶段描述
无零信任阶段	不具备零信任安全技术能力的原始阶段
传统阶段	具备基本网络安全技术能力、概念级的零信任系统能力
初级阶段	具备基础零信任安全技术能力、部分零信任功能模块的安全技术能力
优化阶段	具备系统级别的零信任安全技术能力，支持主动防御能力
持续安全阶段	具备标准级别、可持续提升的零信任安全技术能力

子领域内安全措施总数  $\times 100\%$ 。能力符合度达到 80% 即可判定用户身份可信域达到“优化阶段”。

零信任能力成熟度评估模型是《零信任能力成熟度模型》标准落地的重要抓手，更是零信任技术规范化发展的基石。对各厂商零信任产品进行成熟度测试，有助于各企业查漏补缺，不断完善零信任产品的功能，提升零信任安全服务的质量，推动零信任产业规范化、规模化发展。

### 3 中国零信任发展面临三大挑战

#### 1) 战略统筹布局未开展，规划前瞻性不足

中国已将网络安全建设视为国家长远发展的核心，明确“网络安全牵一发而动全身，深刻影响政治、经济、文化、社会、军事等领域安全”<sup>[17]</sup>。“十四五”期间出台《网络安全产业高质量发展三年行动计划（2021-2023年）》<sup>[8]</sup>等文件，逐步推进零信任相关研究与试点实验项目落地。但是，与美政府持续加大零信任技术投资力度相比，中国仍缺乏整体发展路线图与具体配套支持资金，政府机构、行业企业、社会组织等未形成落地应用的引导措施，缺乏自主技术引领的解决方案和应用实例支撑，零信任战略布局相对滞后。

#### 2) 技术体系建设待完善，自主创新能力不足

当前，中国零信任技术研究紧跟国际发展趋势。2020年8月全国信息技术标准化技术委员会决定对《信息安全技术零信任参考体系架构》标准进行立项。2021年6月中国通信学会组织开展《零信任能力成熟度》标准的研制工作，并于2022年7月完成文稿的发布。虽然已有部分标准成功落地，但是中国零信任网络安全架构尚未形成自主技术体系，创新力仍然较低，仅仅对美国已发布的部分标准进行对标立项，标准化建设还不足，整体架构仍停留在概念、原型阶段，形成的产品细节、规格等方面也存在着诸多差异，难以适配垂直行业多样化应用部署，整体技术发展缓慢且成熟度低，并且尚未开展垂直行业的零信任架构升级和应用部署。

#### 3) 产业生态闭环未形成，供需协同性不足

近两年，中国网络安全龙头企业已推出部分零信任产品及方案，在医疗、电信等行业快速实践落地。但中国零信任产业发展尚处于初期阶段，缺乏全面实施零信任架构的标杆企业。2021年，第三方咨询机构 Forrester<sup>[19]</sup>从推出时间、成熟度和客户情况3个方面筛选出34家国际零信任方案供应商代表，中国仅有3家厂商入选。2022年，第三方咨询机构 Gartner 发布的《零信任网络访问市场指南》<sup>[20]</sup>共调研了42家全球零信任代表性供应商，中国仍是仅3家上榜。此外，

中国缺少产业服务平台支撑，零信任需求侧与供给侧的发展步调不一，市场流通产品方案存在覆盖面不广、适配性不高、服务性不足等基础问题。

## 4 零信任产业发展建议

### 1) 全面统筹规划，强化零信任整体战略布局

强化政策引导，完善中国零信任体系构建。依托财政资金、国家专项等，全面跟进和研判美国在参考架构、关键技术、融合应用等方面的最新进展，加大零信任架构关键技术的科研投入。抓住行业数字化升级转型机遇，推动网络安全架构向零信任架构迁移，鼓励组织研究并出台零信任创新应用试点示范、专项项目等支持政策，并在财政投资的网络安全项目中同步配套零信任架构和关键网络安全基础设施建设措施，明确零信任高质量发展路线图。广泛征集零信任应用实践案例，并对独具行业代表性和可推广性的优秀案例进行宣传，鼓励零信任产品供应商借鉴学习，不断优化产品升级，促进各行业需求方与供应商的长期稳定合作。鼓励传统信息安全提供商、新型云安全提供商与行业用户深度融合，打造一批有影响力的零信任行业应用标杆。在党政、金融、能源、交通等领域开展零信任安全试点示范，提升零信任网络安全架构在实际应用中的市场影响力。

### 2) 重视培育创新，发展零信任自主创新技术

深化核心自主研究，打造中国零信任创新引擎。加快构建基于软件定义边界、身份安全、微隔离三大技术路线的零信任标准体系，推进跨云跨网智能安全管控、基于数字身份的细颗粒度访问控制、可持续数据安全监测和评估等产品升级。同步构建零信任安全测评体系，深入剖析零信任技术能力与功能需求，助力零信任应用部署，依托“零信任能力成熟度”的评估模型，分级分类分步骤差异化部署零信任。推进电信运营商、互联网企业、行业用户、安全企业、科研机构、高等院校等建立零信任联合创新中心、联合实验室等，打造创新试验床，促进研究成果应用转化，保障零信任产品创新能力持续供给，同时定期开展零信任技术专题研讨，助力零信任技术创新以及产品研发。

### 3) 增强产业协同，构建零信任健康产业生态

积极推动供需联动，释放中国零信任集群效应。针对中国大型网络安全供应商提供政策优化保障，打造可以带动跨领域、跨行业能力整合的龙头企业，带动中小型企业蓬勃发展。依托算网融合产业及标准推进委员会等第三方产业平台，提供交流论坛、项目推介、实践比赛、案例评优等多举措多层次供需对接服务，切实强化零信任产业合作。聚焦零信任产业链上下游，绘制零信任产业图谱，适配党政、金

融、能源、交通等领域建立多维度零信任能力评估体系，从核心技术功能、应用解决方案、关键能力成熟度等多维度，引导以优秀产品服务能力为导向的零信任市场良性发展。

## 参考文献

- [1] The White House. Improving the nation's cybersecurity [EB/OL]. (2021-05-12)[2023-10-25]. <https://www.whitehouse.gov/briefing-room/presidential-actions/2021/05/12/executive-order-on-improving-the-nations-cybersecurity>
- [2] The White House. Draft federal strategy for moving the U.S. government towards a zero trust architecture [EB/OL]. (2021-09-07) [2023-10-25]. <https://www.whitehouse.gov/omb/briefing-room/2021/09/07/office-of-management-and-budget-releases-draft-federal-strategy-for-moving-the-u-s-government-towards-a-zero-trust-architecture>
- [3] Defense Information Systems Agency. Thunderdome realizing zero trust [EB/OL]. [2023-10-25]. <https://dd80b675424c132b90b3-e48385e382d2e5d17821a5e1d8e4c86b.ssl.cf1.rackcdn.com/external/fact-sheetthunderdometemplatefinal.pdf>
- [4] The White House. Moving the U. S. government toward zero trust cybersecurity principles [EB/OL]. (2022-01-26)[2023-10-25]. <https://www.whitehouse.gov/wp-content/uploads/2022/01/M-22-09.pdf>
- [5] U. S. Department of Defense. The department of defense releases the president's fiscal year 2023 defense budget [EB/OL]. (2022-03-28)[2023-10-25]. <https://www.defense.gov/News/Releases/Release/Article/2980014/the-department-of-defense-releases-the-presidents-fiscal-year-2023-defense-budget>
- [6] The White House. Administration cybersecurity priorities for the FY 2024 budget [EB/OL]. (2022-07-22)[2023-10-25]. <https://www.whitehouse.gov/wp-content/uploads/2022/07/M-22-16.pdf>
- [7] CSA. The Singapore cybersecurity strategy 2021 [EB/OL]. (2021-10-05)[2023-10-25]. <https://www.csa.gov.sg/Tips-Resource/publications/2021/singapore-cybersecurity-strategy-2021>
- [8] Government of Canada. Network and security strategy [EB/OL]. (2021-04-12) [2023-10-25]. <https://www.canada.ca/en/shared-services/corporate/publications/network-security-strategy.html>
- [9] European Commission. Laying down measures for a high common level of cybersecurity at the institutions, bodies [EB/OL]. (2022-03-22)[2023-10-25]. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022PC0122>
- [10] National Institute of Standards and Technology. Zero trust architecture [EB/OL]. [2023-10-25]. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-207.pdf>
- [11] Cybersecurity and Infrastructure Security Agency. Zero trust maturity model draft [EB/OL]. [2023-10-25]. <https://www.cisa.gov/zero-trust-maturity-model>
- [12] Cloud Security Alliance. SDP specification v1.0 [EB/OL]. (2014-04-30)[2023-10-25]. [https://downloads.cloudsecurityalliance.org/initiatives/sdp/SDP\\_Specification\\_1.0.pdf?\\_ga=2.228130653.1951536048.1669190328-1229248686.1669036181](https://downloads.cloudsecurityalliance.org/initiatives/sdp/SDP_Specification_1.0.pdf?_ga=2.228130653.1951536048.1669190328-1229248686.1669036181)
- [13] Cloud Security Alliance. Software-defined perimeter (SDP) specification v2.0 [EB/OL]. (2022-03-10) [2023-10-25]. <https://cloudsecurityalliance.org/artifacts/software-defined-perimeter-zero-trust-specification-v2>
- [14] Gartner. Market guide for zero trust network access [EB/OL]. (2019-04-29)[2023-10-25]. <https://www.gartner.com/en/documents/3912802>
- [15] 白杨. 零信任安全的2022：产业迎来爆发期，“去虚向实”将成行业共识 [EB/OL]. (2021-04-19) [2023-10-25]. <https://finance.sina.com.cn/chanjing/cyxw/2022/03-03/doc-imcwipih6443414.shtml>
- [16] Market Research. Zero trust security market by solution type, deployment mode, authentication type, organization size, vertical and region – global forecast to 2027 [EB/OL]. (2019-04-29) [2023-10-25]. <https://www.marketresearch.com/MarketsandMarkets-v3719/Zero-Trust-Security-Solution-Type-31844180>
- [17] 中共中央党史和文献研究院. 习近平关于网络强国论述摘编 [M]. 北京: 中央文献出版社, 2021

- [18] 中华人民共和国工业和信息化部. 网络安全产业高质量发展三年行动计划(2021—2023年) [R]. 2021
- [19] Forrester. New tech: zero trust network access, Q2 2021 [EB/OL]. (2021-04-19)[2023-10-25]. <https://www.forrester.com/report/new-tech-zero-trust-network-access-q2-2021/RES161768?objectid=RES161768>
- [20] Gartner. Market guide for zero trust network access [EB/OL]. (2023-09-25)[2023-10-25]. <https://www.gartner.com/en/documents/4773531>

### 作者简介



张云畅，中国信息通信研究院技术与标准研究所技术专家；主要从事零信任、算网安全等方面的研究工作；起草并参与多个零信任白皮书撰写、标准研制工作，目前拥有专著1部，起草行业白皮书1个，参与研制行标/团标5项。



柴瑶琳，中国信息通信研究院技术与标准研究所高级项目主管；主要从事SD-WAN、零信任、算网安全等相关技术研究工作；起草并参与20多项行业/团体标准的制订，发表论文10余篇，授权/申请技术专利6项，拥有软著7项，组织发布行业白皮书3个。



穆研博，中国信息通信研究院技术与标准研究所互联网中心副主任，高级工程师；主要从事算网融合、未来网络、云计算等方面的研究工作；牵头多个技术体系的标准研制，发表论文10余篇，拥有软著12项。

# 集成电路产业技术发展趋势探讨



## Development Trends of Integrated Circuit Industry Technology

刘新阳/LIU Xinyang, 晁沛荫/CHAO Peiyin,  
李婷宇/LI Tingyu  
(深圳市中兴微电子技术有限公司, 中国 深圳 518055)  
(Sanechips Technology Co., Ltd, Shenzhen 518055, China)

DOI: 10.12142/ZTETJ.202306011

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231213.1626.002.html>

网络出版日期: 2023-12-14

收稿日期: 2023-10-12

**摘要:** 集成电路产业已成为全球科技竞争的焦点之一。产业链地图上的主要成员均在尝试强链补链, 加强自身竞争优势, 降低供应链风险。数十年来的集成电路全球化发展趋势发生逆转, 进一步增大了需求和供应的不确定性。中国建设自立自强的集成电路产业更加迫切, 需要走可持续发展路线, 支撑各行业的产业数字化, 提升全球竞争力。从技术发展趋势来看, 中国要坚持双循环, 在先进工艺、先进封装、架构创新等方面进行多路径创新, 坚持多样化发展, 摆脱路径依赖。

**关键词:** 宏观环境; 产业链; 先进工艺; 先进封装; 领域定制架构; 异构设计

**Abstract:** The integrated circuit industry has become one of the focuses of global competition in science and technology. Major members in the industry chain map are trying to strengthen their competitive advantages and reduce supply chain risks. Decades of integrated circuit globalization have reversed, further increasing the uncertainty of demand and supply. It is more urgent for China to build a self-reliant integrated circuit industry, which will take the path of sustainable development, to support the digitalization and improve the global competitiveness of domestic industries. From the perspective of technology trends, China should adhere to dual-circulation and make multi-path innovations in advanced processing, advanced packaging, and architecture innovation. We should persist in diversified development to get rid of path dependency.

**Keywords:** macro-environment; industry chain; advanced processing; advanced packaging; domain-specific architecture; heterogeneous design

**引用格式:** 刘新阳, 晁沛荫, 李婷宇. 集成电路产业技术发展趋势探讨 [J]. 中兴通讯技术, 2023, 29(6): 66–70. DOI: 10.12142/ZTETJ.202306011

**Citation:** LIU X Y, CHAO P Y, LI T Y. Development trends of integrated circuit industry technology [J]. ZTE technology journal, 2023, 29(6): 66–70. DOI: 10.12142/ZTETJ.202306011

## 1 宏观环境变化带来集成电路产业新形势

### 1.1 国际集成电路产业形势

集成电路是一个全球化的产业, 参与全球化分工的各个国家、地区依托自身的优势, 历经数十年市场竞争, 形成了当前的产业格局。美、欧、日、韩、中国是集成电路产业链的主要参与方。其中, 美、欧、日在集成电路上游领域较强。美国整体优势最大, 例如美国在电子设计自动化(EDA)、IP、装备等领域均有明显优势, 欧洲在IP、装备领域较强, 而日本在材料、装备方面占优。2018年以来, 中美科技竞争愈演愈烈, 人工智能<sup>[1–5]</sup>、5G<sup>[6–9]</sup>等未来数字化、智能化产业升级的关键技术成为竞争焦点。作为算、存、传、管、智数字化基础设施的底层关键技术, 集成电路更是受到前所未有的关注。围绕集成电路产业的竞争逐渐从企业层面扩展到产业链甚至国家/地区层面。集成电路全球产业链的发展越来越具有不确定性。

2021年2月, 拜登政府执政之初就“美国供应链安全”签署了14017号行政命令。同年6月, 白宫发布了一份《供应链百日评估》报告, 指出美国供应链存在的一系列漏洞, 并要求“建立有弹性的供应链, 振兴美国制造业, 促进广泛的增长”。在这一报告的指导下, 美国随后密集出台相关法令法规, 一方面通过政策、金融、市场等多种手段加强其在前沿技术上的研发和市场优势, 推动芯片生产制造端重新回归美国国内, 建设本土半导体制造产业链; 另一方面在其政治盟友的基础上构建产业链盟友圈, 在降低自生供应链风险的同时遏制战略对手获得半导体相关技术和装备。其中, 影响力较大的法令法规包括:

- 2022年8月9日, 美国通过《2022年芯片与科学法案》, 向新兴技术提供超过2000亿美元的投资, 其中对美国本土芯片产业提供超过500亿美元补贴, 涵盖半导体制造激励、商业研发和劳动力发展、教育基金、国防基金、创新基金等, 同时要求任何接受补贴的公司10年内禁止在“特定

国家”扩大和新建芯片产能。

- 2022年8月12日，美国商务部工业和安全局（BIS）出台新规定，限制全栅场效应晶体管（GAAFET）电子计算机辅助软件（EDA/ECAD）、宽禁带半导体材料氧化镓（Ga<sub>2</sub>O<sub>3</sub>）和金刚石、压力增益燃烧技术（PGC）等前沿技术出口。

- 2022年10月8日，BIS发布新规定，对出口管制条例（EAR）进行修订，增加对先进计算芯片、超级计算机等的出口限制，并增加外国直接产品FDP脚注的公司名单，以及对美国人/美国公司的出口合规约束。本次修订对中国企业获取高性能计算芯片、超级计算机、半导体制造设备进行了严格限制。

- 2023年9月22日，美国商务部国家标准与技术研究所（NIST）发布了《芯片法案》（“CHIPS Act”）护栏条款实施细则，将所有位于中国的企业（不限于国有企业）列为受关注外国实体。

- 2023年10月18日，BIS修改出口管制条款，进一步收紧对先进计算和先进工艺的限制，展现了保持甚至拉大中美技术差距的战略决心。

作为集成电路产业地图上的重要成员，欧洲和日本在今年也相继发布了相关法案。2023年1月24日，《欧洲芯片法案》草案和修正案通过，其配套的430亿欧元资金与美国芯片法案接近。该法案主要目标是解决集成电路供应短缺问题，加强欧洲的技术领先地位，并制定措施，与成员国和国际合作伙伴一起应对未来供应链中断的挑战。仅3天后，就传出美国已和荷兰、日本就限制向中国出口一些先进的芯片制造设备达成协议的消息。2023年7月23日，日本修订外汇与外贸法相关法令，对先进半导体制造所需的23个品类设备追加出口管制的措施开始正式生效。

可以看出，各国家/地区正在通过宏观政策强化自身在集成电路产业链上的力量，这使全球化发展趋势面临更大的不确定性。在此背景下，加强集成电路产业链自立自强是中国必由之路。

## 1.2 中国集成电路产业形势

在“十四五”产业数字化转型政策的推动下，各行各业对人工智能、5G等技术的需求持续扩大。作为数字化底座的基础技术，中国集成电路产业一方面受到市场需求的推动，对芯片特别是高端芯片的需求越来越大<sup>[10]</sup>，另一方面由于全球化产业链的发展面临不确定性，供应端风险在增大。

习近平总书记说过，“关键核心技术是要不来、买不来、讨不来的”。面对宏观环境的不确定性，中国最具确定性的出路，就是打出政策、产业、技术和人才体系的全方位组合

拳，对中国集成电路产业链形成有效推动力。

近年来，中国集成电路产业链在各个环节都取得长足进步：

- 装备和材料实现从无到有，对成熟制程形成整体供给支撑能力，部分产品进入先进制程供应体系；
- 制造工艺技术取得长足进步，逻辑电路、三维闪存、内存均有突破，并实现量产；
- 封装集成从中低端进入高端，传统封装规模世界第一，先进封装达到国际先进水平，技术种类覆盖90%；
- 设计能力得到大幅提升，处理器、现场可编程门阵列（FPGA）、通信系统级芯片（SoC）等高端芯片取得重大突破。

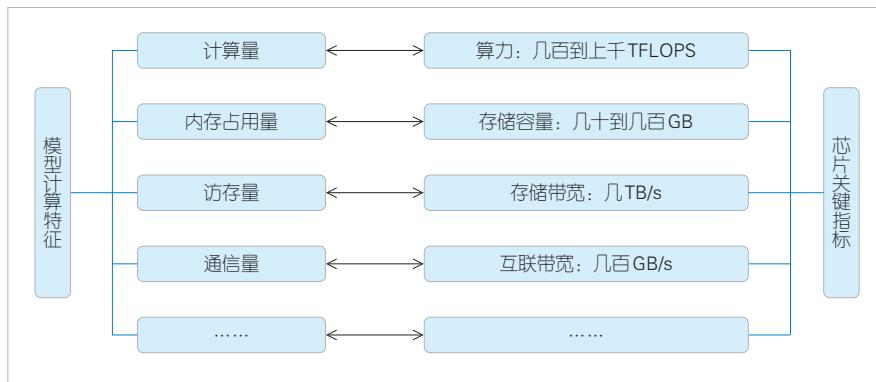
中国集成电路产业不断攻克技术难题，产业链持续向上突破，并向产业生态体系化建设发展。

中国的集成电路产业在快速发展，但也存在发展时间短、关键领域处于落后位置的问题。中国的服务器、通用处理器在技术和市场上均取得快速提升，而国际领先者Intel、NVIDIA等公司也不断推出新产品，他们不仅在先进工艺、先进封装领域持续发展，还在领域定制、异构计算、Chiplet<sup>[11]</sup>等架构创新的推动下获得性能领先。这使得中国同类产品与国际领先者的差距进一步拉大。中国迫切需要有竞争力的设计企业以实现产品竞争力的突破，需要在技术路线上坚持多元化发展，立足内循环，坚持双循环，积极吸收和应用外部的先进技术和能力，以便更好地支撑千行百业的数字化升级。

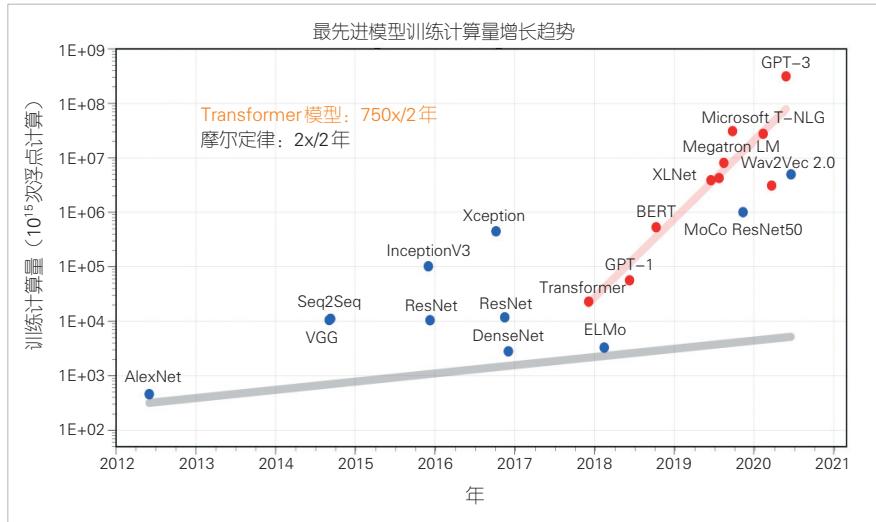
## 2 未来技术发展趋势

集成电路技术创新对国民经济的发展非常重要，是经济增长的触发器。信息社会数字化转型对数据获取、存储、处理需求的不断增加，导致算力需求指数持续上升。集成电路产业技术发展的目标，是单位面积、单位功耗或单位成本下计算密度、存储密度、连接密度的不断提高。自2018年起，大模型蓬勃发展<sup>[12]</sup>。2023年ChatGPT的火爆更是为该领域发展按下加速键，全球科技企业与研究院校等纷纷推出自己的大模型。由此带来的新一轮AI芯片对算力、存力、运力均提出更高需求，如图1所示。

在图2中蓝色为传统卷积神经网络（CNN）、循环神经网络（RNN）等传统模型训练算力，红色为大模型训练算力，灰色为摩尔定律下硬件算力。该图表明，随着大模型参数量规模增长至千亿甚至万亿量级，大模型训练算力每2年提升750倍，芯片算力随摩尔定律每2年提升2倍，大模型训练所需算力增速已远远超过摩尔定律带来的算力供应能



▲图1 模型计算特征与芯片关键指标

▲图2 模型训练计算量与摩尔定律随时间发展增速对比<sup>[13]</sup>

力。随着未来越来越多的大模型应用落地，更大的算力需求将被释放。由于芯片算力=晶体管密度×芯片面积×利用率，因此算力的提升需要从3个方向同时发力，在先进工艺、先进封装、架构创新等方向多路径创新。

## 2.1 先进工艺

晶体管密度的提升来自于先进工艺的不断演进。数十年来，摩尔定律是集成电路技术发展的主旋律：每18~24个月实现性能、功耗、面积(PPA)性能翻倍，推动着集成电路从技术、产品到产业持续指数发展。尽管摩尔定律现在有所放缓，但是在光刻技术、材料、晶体管结构等的推动下仍在持续演进。如表1所示，工艺制程对PPA提升的贡献是其他路径所无法替代的。提高晶体管密度可以使产品集成度增加，从而可在单位面积上实现更多的功能；提升晶体管速度可以提高工作频率，进而提升产品性能；降低晶体管功耗有助于减少整机功耗，从而让产品变得更加节能。

因此，国际领先的晶圆厂均在持续提高工艺水平。

TSMC在晶圆代工制造领域仍处于绝对领先地位，主节点2年一代，保持持续领先优势。可以预见，装备技术和材料科学的持续进步，将为集成电路的持续微缩提供技术基础，推动先进工艺制程在未来数年里继续向1 nm节点发展。

## 2.2 先进封装

在半导体制程工艺提升放缓的趋势下，使用先进封装实现更高集成度，可以扩展芯片面积，从而提升算力，这成为业界越来越关注的技术方向。先进封装分为2.5D封装和3D封装。前者在二维平面扩展芯片面积，后者在三维竖直方向堆叠扩展芯片面积。硅通孔(TSV)技术是先进封装的基础技术之一。高密度高深宽比的TSV制备效果直接关系到集成技术的可靠性和良率等问题。

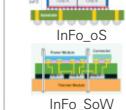
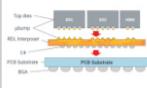
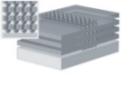
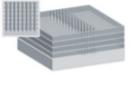
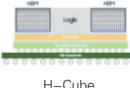
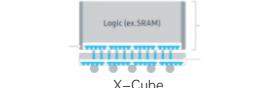
2.5D封装可以通过扇出、硅桥、硅中介层、重分布中介层、混合中介层等方式实现。目前业界主流且已实现量产的代表技术包括CoWoS(TSMC)、InFO\_oS(TSMC)、EMIB(Intel)等，如图3所示。

2.5D封装需求强劲，尺寸和结构持续发展，其中大尺寸是发展趋势。单芯片制造的最大面积上限由光刻机的光罩尺寸(reticle)决定。当前2.5D封装尺寸可达 $3.3 \times$  reticle面积( $3000 \text{ mm}^2$ 左右)，未来2年预计可达 $6 \times$  reticle面积( $5000 \text{ mm}^2$ 左右)。与此同时，晶圆级封装也得到越来越多的应用。中国主流晶圆厂和封测厂等封装产业链各环节厂商纷纷建立2.5D先进封装能力。

相比于2.5D封装，3D封装可以提供更高带宽、更低延迟、更低功耗、更强系统集成能力。与此同时，3D封装的封装复杂度、成本也更高。目前，3D封装在中央处理器/图形处理器、存储、光电合封芯片等领域都已经实现应用。同

▼表1 TSMC典型工艺节点对性能、功耗、面积(PPA)的提升

工艺节点	归一到28 nm		
	晶体管密度	性能	功耗
28 nm	1.00 ×	1.00 ×	1.00 ×
16 nm	1.56 ×	1.42 ×	0.50 ×
7 nm	4.68 ×	1.91 ×	0.17 ×
5 nm	8.56 ×	2.20 ×	0.14 ×
3 nm	13.36 ×	2.53 ×	0.10 ×

	2.5D封装					3D封装	
	扇出	硅桥	硅中介层	重分布中介层	混合中介层	微凸点	混合键合
							
							
							

注：图片分别来自TSMC、Intel及Samsung公司

▲图3 业界典型2.5D封装和3D封装技术

时，3D封装从微凸点向混合键合演进，通过更高的互连密度来大幅提升带宽。中国3D封装在存储领域已实现量产，逻辑芯片也已进入研发导入流程。

综上所述，先进封装有多种细分技术方向，在带宽、封装面积、信号完整性(SI)、电源完整性(PI)、热管理、工艺成熟度和成本等方面各有不同，需要结合相应产品需求选择匹配的封装方案。

### 2.3 架构创新

图灵奖获得者J. M. HENNESSY和D. A. PATTERSON认为，CPU等通用架构处理器的性能增加越来越缓慢，未来10年将是体系结构发展的黄金年代。特定领域架构(DSA)的设计可以进一步提升系统性能。Google、Tesla等基于特定应用设计他们的DSA芯片。即便当前占据绝对市场的Nvidia GPGPU芯片也越来越多地通过融合DSA设计来提升性能。通过分析需求场景和算法特点，DSA可以发现不能被当前架构较好支持的部分，并能用指令粗粒度固化该部分进行加速计算，从而提高系统性能。然而DSA架构在软件方面仍存在巨大挑战，例如指令的碎片化会导致软件适配的工作难以被下一代芯片复用。因此，架构设计的关键在于：根据算法趋势和应用场景，在硬件极致能效和通用可编程之间找到平衡。

除DSA外，另一种重要的架构创新是被称为“乐高式”架构的异构集成Chiplet架构技术，即小芯片或芯粒。Chiplet技术将多个可模块化芯片(主要形态为裸片“Die”)通过Die to Die互连技术集成在一个封装内，构成专用功能异构芯片。Chiplet能够从成本、性能、最小化资源需求和灵活搭配4个角度对现有架构进行优化，逐渐成为未来算力方案的

基础架构。

Chiplet技术不是简单的“拆/分”“拼接”方案，要想最大化产品的性能、功耗、面积、成本竞争力，需要从先进封装、Die to Die互连技术、系统级协同设计3个维度协同创新。先进封装是实现2.5D封装的工艺基础。Die to Die互连技术是系统能够高效组合芯粒的关键技术。互连协议标准应具备兼容性与开放性。目前国际上BoW(开放计算项目提出的并行互联标准)、AIB(Intel提出的并行互联标准)等并行方案可能趋向收敛于UCIE(Intel提出的通用Chiplet并行互联标准)，但仍存在超短距离SerDes串行互联标准(XSR)和极短距离SerDes串行互联标准(USR)等少量串行方案。与此同时，中国相关联盟也在建立标准，积极推动Chiplet生态建设。系统级协同设计是指如何根据目标，构建最优的芯粒分解-组合-集成方案，同时在整个前端-后端-封装的开发流程中进行适配。此外，Chiplet技术还面临着散热、供电、多物理场仿真、可测性等方面的挑战，需要芯片设计公司、OSAT、Foundry、EDA、IP、原始材料供应商等产业链各方共同培育完整的生态。

### 3 结束语

集成电路产业对人工智能、5G等产业升级起到关键基础作用，是全球科技竞争的焦点之一。各个国家/地区在通过政策手段加强自身竞争优势、降低供应链风险的同时，向竞争对手施加出口管制，给产业的供需端带来更大的不确定性。在外部宏观环境的变化和自身产业升级需求的推动下，中国致力于建设自立自强、可持续发展的集成电路产业，为各行业的产业数字化和全球竞争提供支撑。从发展趋势

来看，中国既要在先进工艺、先进封装、架构创新等方面进行多路径创新，同时也要坚持多样化发展，进行独特的路径创新，建立可持续发展生态。

## 致谢

感谢深圳市中兴微电子技术有限公司王健、庞健和吴枫对本论文做出的贡献！

## 参考文献

- [1] 白冰, 裴丽, 左晓燕. 用于人工智能的硅基光电子芯片 [J]. 中兴通讯技术, 2021, 27(1): 77–82. DOI:10.12142/ZTETJ.202101015
- [2] 韩炳涛, 刘涛, 唐波. 深度学习的10年回顾与展望 [J]. 中兴通讯技术, 2022, 27(6): 75–84. DOI:10.12142/ZTETJ.202206013
- [3] ZHANG C C, ZHANG N, CAO W, et al., AI-based optimization of handover strategy in non-terrestrial networks [J]. ZTE communications, 2021, 19(4): 98–104. DOI: 10.12142/ZTECOM.202104011
- [4] LIU W C, SHEN M Q, ZHANG A D, et al. Artificial intelligence rehabilitation evaluation and training system for degeneration of joint disease [J]. ZTE communications, 2021, 19(3): 46–55. DOI: 10.12142/ZTECOM.202103006
- [5] LU F, GU L, TIAN X H, et al. Federated learning based on extremely sparse series clinic monitoring data [J]. ZTE communications, 2022, 20(3): 27–34. DOI: 10.12142/ZTECOM.202203004
- [6] 陆平, 欧阳新志, 高雯雯. 5G行业虚拟专网能力提升与实践 [J]. 中兴通讯技术, 2022, 27(2): 68–74. DOI:10.12142/ZTETJ.202202011
- [7] 张启明, 郑兴明, 张寿勇. 5G TSN技术的创新研究 [J]. 中兴通讯技术, 2022, 27(3): 78–83. DOI:10.12142/ZTETJ.202203013
- [8] LIU Z, GAO Y, LI D P, et al. Enabling energy efficiency in 5G network [J]. ZTE communications, 2021, 19(1): 20 – 29. DOI: 10.12142/ZTECOM.202101004
- [9] XIAO K, LIU X, HAN X H, et al. Flexible multiplexing mechanism for coexistence of URLLC and eMBB services in 5G networks [J]. ZTE communications, 2021, 19(2): 82 – 90, DOI: 10.12142/ZTECOM.202102011
- [10] 刘新阳. 5G半导体产业发展和创新趋势思考 [J]. 中兴通讯技术, 2021, 27(4): 51–52. DOI:10.12142/ZTETJ.202104010
- [11] 李乐琪, 刘新阳, 庞健. Chiplet关键技术与挑战 [J]. 中兴通讯技术, 2022, 27(5): 57–62. DOI:10.12142/ZTETJ.202205011
- [12] YANG J F, JIN H Y, TANG R X, et al. Harnessing the power of llms in practice: a survey on chatgpt and beyond [EB/OL]. (2023–04–26) [2023–10–15]. <https://arxiv.org/abs/2304.13712>
- [13] GitHub. AI and memory wall [EB/OL]. [2023–10–15]. [https://github.com/amirgholami/ai\\_and\\_memory\\_wall](https://github.com/amirgholami/ai_and_memory_wall)

## 作者简介



**刘新阳**, 深圳市中兴微电子技术有限公司副总经理、中国通信学会集成电路专委会副主任委员、中兴通讯股份有限公司多媒体与移动网络国家重点实验室IC研发中心主任, 曾任中兴通讯硬件研究所副所长、微电子研究院副院长、终端事业部整体解决方案产品线产品经理；主要负责中兴微电子芯片战略与技术规划、运营工作, 拥有20余年通信集成电路规划、研发、经营管理经验。



**晁沛荫**, 深圳市中兴微电子技术有限公司战略规划总监；具备丰富的电子、通信、芯片等ICT行业领域经验, 研究方向为集成电路宏观环境和中长期规划。



**李婷宇**, 深圳市中兴微电子技术有限公司技术规划工程师；主要从事人工智能技术规划及芯片架构的研究工作。

# 软件定义广域网络(SD-WAN) 2.0 测试方法



## Test Methodology of Software Defined Wide Area Network (SD-WAN) 2.0

党小东/DANG Xiaodong, 柴瑶琳/CHAI Yaolin,  
穆琰博/MU Yubo

(中国信息通信研究院, 中国 北京 100191)  
(China Academy of Information and Communications Technology, Beijing  
100191, China)

DOI: 10.12142/ZTETJ.202306012

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20231212.1441.002.html>

网络出版日期: 2023-12-13

收稿日期: 2023-10-12

**摘要:** 以产业共性需求为基准,结合技术更迭的具体特征,构建了设备、方案和服务3个维度的软件定义广域网络(SD-WAN)2.0测试规范体系。该体系从不同维度抽象出测试工作的方向和范围,明确了相应的测试评估指标。该测试方法体系将进一步的SD-WAN 2.0测试研究工作提供重要参考。

**关键词:** SD-WAN 2.0; 测试方法; 设备; 方案; 服务

**Abstract:** Based on the common needs of the industry and combined with the specific characteristics of technological change, the software defined wide area network (SD-WAN) 2.0 test specification system of three dimensions of equipment, solution, and services is constructed. The system abstracts the test direction and test scope of the test work from different dimensions, and clarifies the corresponding test evaluation indicators. This test methodology provides an important reference for further SD-WAN 2.0 test research work.

**Keywords:** SD-WAN 2.0; testing method; equipment; solution; service

**引用格式:** 党小东, 柴瑶琳, 穆琰博. 软件定义广域网络(SD-WAN)2.0测试方法 [J]. 中兴通讯技术, 2023, 29(6): 71-75. DOI: 10.12142/ZTETJ.202306012

**Citation:** DANG X D, CHAI Y L, MU Y B. Test methodology of software defined wide area network (SD-WAN) 2.0 [J]. ZTE technology journal, 2023, 29(6): 71-75. DOI: 10.12142/ZTETJ.202306012

**伴** 随着中国各垂直行业数字化转型的工作推进,信息通信网络的数字化、算力化、智能化将成为产业高质量发展的必然选择。近年来,软件定义广域网络(SD-WAN)已经成为云网融合实践应用的新热点。IDC最新产业调研数据显示<sup>[1]</sup>,2021年中国SD-WAN市场规模达1.9亿美元,增速近90%;预计未来5年以每年55.3%的年复合增长率增长,到2026年市场规模将达到17.2亿美元。SD-WAN应用已经覆盖电信、金融、能源、零售、医疗、教育、工业、政务、移动互联网、交通等10多个垂直方向,已经形成广泛而深刻的影响。

随着5G、人工智能、物联网等新业务的加速部署<sup>[2-4]</sup>,广域网应用又迎来诸多新型挑战:基础网络架构封闭、无边界安全隐患、多终端移动性管理复杂、算力灵活调度困难等。局限于Overlay网络互联的SD-WAN基本功能特性已经

无法适配企业新业务需求,SD-WAN逐渐向融合内生安全、IPv6+Underlay确定性体验保证、多云访问的2.0阶段发展演进。

系统规范的测试方法是技术发展的必要因素。当前全球对SD-WAN 2.0的测试方法还没有形成规范化体系。本文将从设备、方案、服务3个维度对SD-WAN 2.0测试方法进行研究,为SD-WAN 2.0技术发展提供可靠、可用的测试依据。

### 1 测试方法体系的总体设计

#### 1.1 三大测试维度

随着软件定义技术的不断成熟,SD-WAN 2.0逐渐发展为融合多种信息技术(主要包括软件定义网络、网络功能虚

拟化、网络编排与探测等)的新型广域网架构(如图1所示),能够以平台或托管方式提供基础网络连接、广域网加速、安全防御、智能运维等多种网络服务。面对以业务与应用为导向的SD-WAN 2.0新型架构,传统单一的网络能力测试方法已经难以满足系统评估和分析的需要。开展SD-WAN 2.0的测试工作急需建立新的系统级测试方法。

当前SD-WAN市场交付形态分为3种:设备交付方式、方案交付方式、服务交付方式。针对这3种交付方式,本文研究设备、方案以及服务3个维度的测试方法。设备维度的研究对象主要是基于设备(包含SD-WAN终端设备、SD-WAN控制器和SD-WAN编排器)层面的功能、安全、可靠以及性能方面的测试;方案维度的研究对象主要基于SD-WAN编排器、SD-WAN控制器、SD-WAN终端设备一套可整体部署的网络解决方案,但不包括网络资源;服务维度的研究对象主要基于在传统网络运营商提供的Underlay连接服务(包括以太网服务、IP服务、L1连接服务、互联网接入服务等)基础上集成SD-WAN能力,为用户提供可直接使用的基础通信服务和增值服务。

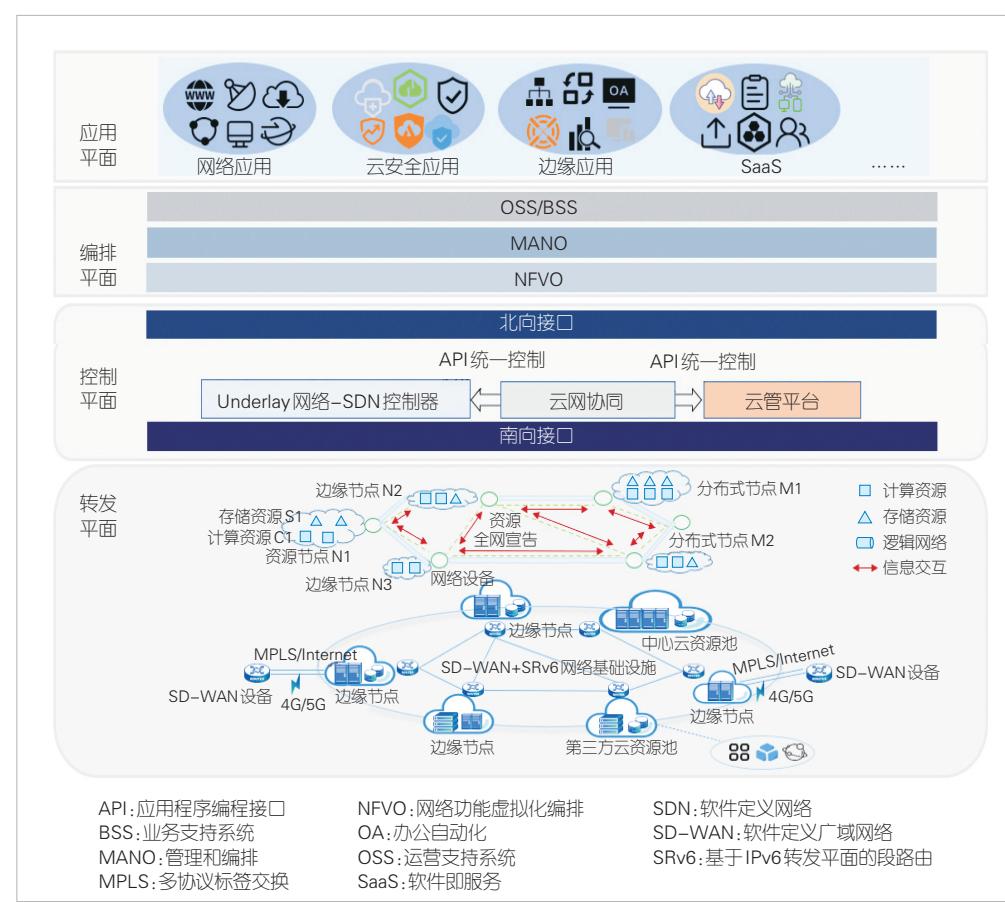
## 1.2 测试视角与测试思路

目前,已有多个标准化组织和开源社区陆续开展了SD-WAN 2.0的测试方法探究。开放网络用户组织(ONUG)定义了面向企业多云集成应用场景SD-WAN 2.0架构,并总结了4个典型用例:PCI Compliance、Guest Access、Direct Cloud Access、Direct Internet Access。互联网工程任务组(IETF)定义了SD-WAN服务数据模型以及开放式软件定义广域网交换(OSE)数据模型草案。城域以太网论坛(MEF)定义了两个典型用例Hybrid WAN(SD-WAN Service over Internet and MPLS WANs)和Dual Internet WAN(SD-WAN Service over Multiple ISPs),并开展了SD-WAN服务质量评测工作。针对SD-WAN相关的测试方法主要集

中在服务质量方面,缺乏对网元层面和组网层面的考虑,存在测试片面的问题。

从测试视角出发<sup>[5]</sup>,本测试方法体系由设备、方案、服务三级构成,设备和方案测试秉承SDN的“转控分离”原则,可分为控制面测试和转发面测试两部分。其中,控制面测试致力于策略管理和网络控制,向下通过南向接口协议对接SD-WAN终端设备,以实现协议/功能的集中控制和定制,向上则通过北向接口为上层业务应用以及资源管理系统提供灵活的网络资源抽象,并开放多个层次的可编程能力。转发面测试专注于系统接口能力和整体功能/性能指标,灵活定义网络策略和行为,便捷地实现任何新业务和新功能。

按照从局部到全局的测试思路<sup>[6]</sup>,应当建立设备级别、方案级别和服务级别的三级评测方法体系。设备维度首先应当从转发网元入手,这是因为转发网元的功能/性能参数是评测网络能力最基本的技术指标;然后在此基础上扩展测试范围,此时应当围绕与转发网元连接的控制网元和编排网元实体展开测试。这些网元实体是完整实现具体网络能力的最小单元。考虑到不同应用场景下SD-WAN技术体系存在的差异化特点,我们需要深入研究结合场景部署的技术要求,



▲图1 SD-WAN 2.0整体架构图

形成方案级别的评测方法；针对上层服务，结合应用场景部署的技术要求，形成服务级别的评测方法。有针对性地开展测试评估工作。设备、方案以及服务测试的结合，更容易形成系统的测试方法体系。这对指导产业发展、推动SD-WAN创新技术商业化进程具有重大意义。

## 2 设备维度的测试方法分析

中国传统测试方法基于网元本身的功能和性能，主要针对2—3层的普通流量进行测试，系统模块化、标准化基础较差，较少考虑开放性、通用性、综合能力，在可测试性方面缺乏系统考虑和规范化要求。与传统测试方法不同，SD-WAN 2.0 测试方法注重自上而下成体系的测试，依托标准具备完善的顶层设计规范采用全球专业仪表做流量仿真，真实模拟视频、语音、OA等4—7层的应用流量，具有一致性和可重复性，且能够控制网元，更适合做兼容性测试。

当前SD-WAN设备维度技术架构大体上可以抽象为3个功能层级：基础设施层、业务控制层和管理编排层。基础设施层要求转发面网元具备网络可编程能力，通过业务控制层的南向接口开放网络设备的控制功能；业务控制层要求能够管控基础设施，形成全局视角，借助业务控制层的北向接口向管理编配层提供网络服务能力；编排管理层是对SD-WAN业务的抽象和管理，对业务中各个网元和网元行为特征的抽象在技术后台的管理功能层中实现。

### 2.1 基础设施层能力测试

基础设施层是SD-WAN技术架构的最底层。作为重要的转发单元，基础设施层测试主要围绕功能测试、安全测试、可靠测试三大类，从协议、功能、策略各个方面展开系统的评测。

#### 1) 功能测试

- 自动化部署：边缘设备通过动态主机配置协议(DHCP)、U盘、邮件、短信等方式，进行自动化部署、远程部署、虚拟化部署、版本检测与升级。
- 动态选路：根据链路服务等级协议(SLA)、应用带宽、优先级、权重动态选择链路，根据动态路由协议支撑Overlay组网策略和规则。
- 应用识别<sup>[7]</sup>：基于端口协议(POP3-110、SMTP-25、NetBIOS-138、HTTP-443)、负载协议(ECMP协议)、特征库识别、首包识别、深度包检测(DPI协议)等识别机制。
- WAN优化：自适应前向纠错(FEC)以及包令校正(POC)等技术，应用压缩和重复数据消除。
- 网络地址转换(NAT)穿越：测试设备支持NAT-D负

载、NAT会话穿越应用程序(STUN)技术，边缘设备支持与控制面、管理面进行NAT穿越互通。

#### 2) 安全测试

- URL过滤：自定义设置统一资源定位符(URL)黑名单和白名单，进行网页过滤，提供超文本传输协议(HTTP)和超文本传输安全协议(HTTPS)URL过滤，进行恶意URL防护。

• 防火墙策略：包括基于IPv4源/目的地址、IPv6源/目的地址、目的端口、源端口、媒体访问控制(MAC)地址访问控制策略以及静态NAT策略。

• 隧道加密：使用IP网络安全协议(IPSec)/传输层安全协议(TLS)/数据包传输层安全协议(DTLS)等加密技术来保护传输中的数据，借助相关国密算法SM2/SM3/SM4/SM9或业界公认可算法AES256等构建数据传输隧道。

#### 3) 可靠测试

- 主备切换：边缘设备主备切换，切换期间业务流量转发无中断。

### 2.2 业务控制层能力测试

控制器层测试向上承接编排器层传递的业务逻辑测试(主要为南向接口和北向接口测试)，向下对转发网元实现管理测试，包含设备管理(SD-WAN终端设备、SD-WAN安全设备)、网络管理(Underlay网络、Overlay网络)、日志管理。

• 设备管理：部署设备接收控制器策略的下发，并和控制器进行实时信息共享。

• 网络管理：Overlay网络拓扑管理和网络路由的集中控制。

• 日志管理：控制器管理设备的信息采集和流量信息的采集与分析。

• 北向接口：控制器提供给上层应用的调用接口以及协议一致性测试。

• 南向接口：功能一致性和互操作性。

### 2.3 编排管理层能力测试

编排器层上接应用层接收用户的操作命令，并对网络业务进行抽象和建模，下接控制器层在实际的网络中实现业务的需求。编排管理层测试主要围绕网络服务生命周期管理、资源调度、跨域编排展开。

• 网络服务生命周期管理：对部署过程、网络服务的执行和终止全部生命周期的管理。

• 资源调度：网络资源模型映射、虚拟化网元管理等。

• 跨域编排：跨域统一策略下发和执行、跨域资源共享等。

### 3 方案维度的测试方法分析

方案维度的测试是SD-WAN独特的测试方法。该方法将SD-WAN的众多组件进行连接组网，集编排、控制、转发于一体，不对单一网元展开测试，而是从网元的功能、链路的负载、SD-WAN控制器的联动多个方面展开整体组网方案的系统评测。这种测试主要从基本方案功能、管理运维功能、安全防护功能3个方面展开测试方法的研究。

#### 1) 基本方案功能测试

- 负载均衡：根据目标地址、源地址、服务、时间等要素的链路负载均衡，基于自动探测链路质量的主动链路均衡算法，基于应用的负载均衡等。

- 按需组网：支持full-mesh拓扑、hub-spoke拓扑等，支持full-mesh拓扑变更为hub-spoke拓扑，支持hub-spoke拓扑变更为full-mesh拓扑。

- 流量调度：将不同类型（地址、端口、协议）的应用在不同的WAN口隧道上进行调度，根据时延、抖动、丢包率设置相应的切换阀值进行流量调度。

- 路由交换<sup>[6]</sup>：测试通过静态路由协议和动态路由协议，例如开放最短路径优先（OSPF）、外部边界网关协议（EBGP）、内部边界网关协议（IBGP）等，在网络拓扑中进行数据包的交换转发。

- 服务质量功能：根据优先级队列、低时延队列等队列调度，流量整形等流量优先级进行网络流量调度，以确保业务转发可以高效运行。

#### 2) 管理运维功能测试

- 管理运维：通过可视化界面来集中管理实时网络资源监控、日志搜集和告警上送、版本升级等，保障网络业务基本运维。通过图形界面来实现下发设备跨大版本的平滑升级，提供日志文件一键收集功能、控制器管理整个设备组网情况和日志。

#### 3) 安全防护功能测试

- 安全防护<sup>[6]</sup>：通过设置安全策略、访问安全策略等，来保障数据稳定转发。

### 4 服务维度的测试方法分析

服务维度的测试同样也是SD-WAN独特的测试方法。与方案维度和设备维度不同，服务维度将底层资源抽象为上层服务，主要作用于用户层和应用层，需要从用户条款、约束条款、服务商免责条款以及服务赔偿条款方面，对SD-WAN服务进行服务管理规定测评和技术服务能力测评。

#### 1) 服务管理规定

- 用户条款：用户需要提供的相关用户信息应真实、准确、完整，变更服务时需要及时提供、更新相应身份资料及信息等。

- 约束条款：针对服务资质审查、服务类型、服务范围、服务产品管理等。

- 服务商免责条款：SD-WAN服务提供商应告知用户自身负责的条款，并将其以网站等方式对外正式公布，做到充分透明。SD-WAN服务提供商免责条款应包括免责的范围和解释、免责的情景等信息。

- 服务赔偿条款：SD-WAN服务提供商应承诺，在某项服务指标没有达到服务协议中的要求时，提供相应的赔偿，同时明确可获得赔偿的指标项、赔偿的具体方式、赔偿额度的计算方式、最低赔偿、最高赔偿等。SD-WAN服务提供商应明确服务方不承担责任的赔偿，例如由于用户自身操作不当导致的可用性不达标，可不承担赔偿责任。

#### 2) 技术服务能力

- 服务业务弹性：用户可按需选择基础网络连接拓扑、增值服务订购等业务能力。

- 服务可管理性：全生命周期的服务管理，包括服务的创建、删除、暂停、恢复、停止等操作。

- 服务安全性：通信安全、数据安全、用户安全。

- 服务质量保障：网关节点骨干网络的端到端可用率、不同客户场景的标准化服务、完整成熟的服务流程，保障用户业务数据正常传输的质量指标。

- 故障快速响应：对于系统故障，服务提供商应在一定时间内对故障进行处理，并将故障处理结果反馈给客户。

- 服务计量准确性：根据选择的服务类型，按照所列明的要求保证所出具的计量校准结果的准确性和溯源性。

### 5 测试实践与问题分析

#### 5.1 相关 SD-WAN 2.0 测试实践

当前，基于SD-WAN 2.0测试方法体系，中国SD-WAN厂商做了相关的测试实践。在一些功能上，各厂商实现方式不同，且测试重点不同，例如：服务提供商更侧重于服务质量及服务时效性，整体解决方案提供商更侧重于整体组网的健壮性和灵活性，设备提供商更侧重于设备的转发能力和与控制器的兼容性。本文中我们选取国际组织MEF和IETF开展的一些测试工作进行对比，如表1所示。从对比结果来看，相较于SD-WAN 2.0测试方法，这些测试工作都或多或少地存在测试维度片面问题，导致测试不完整性，不能完全反映SD-WAN的技术完备性。

▼表1 全球研究机构的SD-WAN 2.0测试对比

机构名称	测试维度	测试模块	测试指标
MEF	服务维度、方案维度	API载入测试、互操作性测试	产品上线和配置支持、测试数据管理、有效载荷
IETF	服务维度	SD-WAN云网络测试、SD-WAN服务融合测试	确定性网络测试、识别应用程序、SD-WAN隧道测试
中国信息通信研究院	服务维度、方案维度、设备维度	服务维度:融合SASE服务、SD-WAN基础服务、SD-WAN增值服务 方案维度:融合SASE方案、SD-WAN经典安全、融合零信任、融合垂直行业 设备维度:SD-WAN融合下一代防火墙、端点保护等	服务维度:服务管理要求和技术要求等 方案维度:组网功能、安全、可靠等 设备维度:设备功能、安全、可靠、兼容、性能等

API: 应用程序编程接口 IETF: 互联网工程任务组 MEF: 域域以太网论坛 SASE: 安全访问服务边缘 SD-WAN: 软件定义广域网络

## 5.2 测试工作中的问题探讨

### 1) 测试思路局限性

目前已经开展的测试工作仍然聚焦于，以SD-WAN设备为代表的转发单元能力和以SD-WAN控制器为代表的控制单元能力。SD-WAN 2.0融合多种技术的方案级测试工作，目前尚处于初级阶段。

### 2) 测试标准滞后性

针对系统性和规范性的测试方法标准体系的缺失，已成为制约测试工作进一步发展的重要瓶颈。标准化组织相对滞后，与技术要求配套的测试方法体系标准成为当前亟须完善的重点。

### 3) 测试工具制约性

当前传统的商业仪表及开源项目难以满足系统级的方案测试。开源项目自身在可靠性和稳定性方面存在缺陷。开源社区在技术支持等服务质量方面存在重大短板。整体上看，缺少可靠的测试工具已成为制约测试工作进一步开展的另一个重要因素。

## 6 结束语

当前，SD-WAN 2.0相关的测试工作已处于相对成熟的发展阶段。未来，伴随着商业化进程的不断深入，SD-WAN 2.0相关的测试工作也将逐步得到业界的重视，吸引越来越多的技术专家参与其中，以完善测试方法，研发系统级评测工具，全面提升SD-WAN 2.0技术评测水平。

## 致谢

感谢中国信息通信研究院技术与标准研究所毕立波、韩淑君工程师对本论文做出的贡献！

## 参考文献

[1] IDC. 百花争艳,2021年中国SD-WAN市场增长再创新高 [EB/OL]. (2022-05-

- [10][2023-10-16]. <https://www.idc.com/getdoc.jsp?containerId=prCHC49101522>
- [2] 陆平, 欧阳新志, 高雯雯. 5G行业虚拟专网能力提升与实践 [J]. 中兴通讯技术, 2022, 28(2): 68–74. DOI: 10.12142/ZTETJ.202202011
- [3] 魏亮, 查选, 戴方芳. 面向云网融合的网络安全互操作 [J]. 中兴通讯技术, 2023, 29(1): 68–74. DOI: 10.12142/ZTETJ.202301003
- [4] 周吉喆, 杨思远, 王志勤. 面向业务感知的算网融合关键技术研究 [J]. 中兴通讯技术, 2022, 28(5): 2–6. DOI: 10.12142/ZTETJ.202205002
- [5] 赵慧玲, 史凡. SDN/NFV 的发展与挑战 [J]. 电信科学, 2014, 30(8): 13–18. DOI: 10.3969/j.issn.1000-0801.2014.08.002
- [6] 穆琰博, 马军锋, 徐晓麟. SDN/NFV 测试方法的研究 [J]. 中兴通讯技术, 2017, 23(2): 27–32. DOI: 10.3969/j.issn.1009-6868.2017.02.006
- [7] 柴瑶琳, 穆琰博, 马军锋. SD-WAN 关键技术 [J]. 中兴通讯技术, 2019, 25(2): 15–19. DOI: 10.12142/ZTETJ.201902003

## 作者简介



党小东，中国信息通信研究院技术与标准研究所互联网中心SD-WAN项目经理、中国通信标准化协会算网融合产业及标准推进委员会泛在计算组(CCSA TC621 WG10)副组长；主要研究领域为SD-WAN、SASE等；主持或参与10余项行业/团体标准的制订，参与撰写2篇行业白皮书，拥有3项软件著作。



柴瑶琳，中国信息通信研究院技术与标准研究所高级项目主管；主要从事SD-WAN、零信任、算网安全等相关技术研究工作；起草并参与20多项行业/团体标准的制订，发表论文10余篇，授权/申请技术专利6项，拥有软著7项，组织发布行业白皮书3个。



穆琰博，中国信息通信研究院技术与标准研究所互联网中心副主任，高级工程师，中国通信标准化协会算网融合产业及标准推进委员会秘书长；主要从事算网融合、未来网络、云计算等方面的研究工作；牵头多个技术体系的标准研制，拥有12项软件著作，发表论文10余篇。

# 《中兴通讯技术》第29卷总目次

卷 · 期 · 页

## 卷首特稿

算力网络研究与探索 ..... 张宏科, 权伟, 刘康 29-1-01

## 热点专题

### 面向云网安全的新型防护技术

- 专题导读 ..... 解冲锋, 杨义先 29-1-06  
面向云网融合的网络安全互操作 .....  
魏亮, 查选, 戴方芳 29-1-07  
基于超级SIM的5G端云安全体系架构与关键技术 .....  
李佩源, 刘建伟 29-1-13  
云网融合下的安全能力池关键技术与应用 .....  
余启明, 吴爽, 黄帅, 刘紫千 29-1-20  
未来网络内生安全通信技术 .....  
闫新成, 周娜, 蒋志红 29-1-26  
云平台DNS安全体系研究 ..... 宋林健, 马永, 梁卓 29-1-33  
构建可扩展的RPKI依赖方系统部署机制 ..... 马迪 29-1-40  
大型企业SASE解决方案及应用实践 .....  
王茜, 陈晨, 井俊丰, 季家震 29-1-45

## 语义通信

- 专题导读 ..... 陆建华, 陶晓明 29-2-01  
信源信道联合的新范式: 语义通信 .....  
吕守晔, 戴金晟, 张平 29-2-02  
语义信息论的回顾与展望 ..... 辛港涛, 樊平毅 29-2-09  
基于信息论的语义通信: 理论与挑战 .....  
施雨轩, 吴泳澎, 张文军 29-2-13  
面向语义通信的语义知识库研究综述 .....  
孙亚萍, 崔曙光, 张平 29-2-19  
基于CSI反馈的语义图像传输 .....  
姜培文, 韩瑜, 金石, 李潇 29-2-24  
基于深度联合信源信道编码的CSI反馈技术 .....  
许佳龙, 陈为, 艾渤 29-2-29  
语音信源的语义编码传输方法研究 .....  
牛凯, 姚圣时, 戴金晟 29-2-34  
智简语义通信的链路设计及关键技术研究 .....  
孙梦颖, 熊华超, 王怡宁, 韩书君, 许晓东 29-2-40

卷 · 期 · 页

- 语义通信性能评估体系及指标 .....  
郑远, 王凤玉, 许文俊 29-2-46  
基于深度学习的图像语义通信系统 .....  
张振国, 杨倩倩, 贺诗波 29-2-54  
基于语义重要度的不等错误保护数据传输机制 .....  
郭帅帅, 李树静 29-2-62

## 数字孪生技术

- 专题导读 ..... 陈前斌 29-3-01  
基于数字孪生网络的6G无线网络自治 .....  
刘光毅, 邓娟, 郑青碧 29-3-02  
基于数字孪生网络的6G智能网络运维 .....  
王威丽, 唐伦, 陈前斌 29-3-08  
基于长短期记忆网络的数字孪生移动通信网络环境生成技术 .....  
梁广明, 杨鲲, 刘强, 陈建军 29-3-15  
数字孪生边缘网络 ..... 张彦, 卢云龙 29-3-21  
无线网络孪生中的统计信道建模方法: 现状与前沿 .....  
张树韬, 薛烨, 史清江, 张纵辉 29-3-26  
面向6G的无线接入网络数字孪生技术 .....  
段向阳, 康红辉, 吕星哉, 芮华 29-3-32  
数字孪生赋能下的6G网络资源管控机制 .....  
程子豪, 刘向南, 高宏伟, 管婉青, 张海君 29-3-38  
基于数字孪生的算力网络自优化技术研究 .....  
许胜, 许方敏, 赵成林 29-3-46  
天地一体网络场景下的数字孪生关键技术 .....  
陈新宇, 张强, 陆光辉 29-3-51

## 算力网络和东数西算

- 专题导读 ..... 赵慧玲 29-4-01  
东数西算场景下的算力网关研发及实践 .....  
马思聪, 孙吉斌, 孙一豪 29-4-02  
算力网络四面三级算力度量技术体系 .....  
杜宗鹏, 李志强, 陆璐 29-4-08  
东数西算下面向业务的路由策略分析与探索 .....  
魏汝翔, 刘琦, 赵广, 曹畅, 唐雄燕 29-4-14  
面向算力网络的多路径时敏优先调度机制 .....  
夏华屹, 权伟, 张宏科 29-4-19  
算力网络资源协同调度探索与实践 .....

彭开来, 王旭, 唐琴琴	29-4-26
面向算力网络的云边端协同调度技术	
周旭, 李琢	29-4-32
一种面向服务的算网路由架构方案	
黄光平, 谭斌, 吉晓威	29-4-38
通用在网计算系统架构及协议设计	
姚柯翰, 陆璐, 徐世萍	29-4-43
大规模语言模型的跨云联合训练关键技术	
潘圃丞, 侯永帅, 杨卿, 余跃, 相洋	29-4-49

## 6G 网络技术

专题导读	王文东	29-5-01
6G智能内生网络架构及关键技术分析		
李文璟, 喻鹏, 张平	29-5-02	
面向6G的星地融合网络架构		
徐晖, 陈山枝, 艾明	29-5-09	
6G网络架构展望	刘玉芹, 邢燕霞, 陈鹏	29-5-16
6G网络架构和关键技术展望	王友祥, 唐雄燕	29-5-21
6G网络架构研究进展及建议	谢峰	29-5-28
面向6G无线环境可预测的感知重构、语义表征及应用		
张建华, 田艺璇, 孙语瞳, 于力	29-5-38	
面向下一代移动通信的前传网络关键技术		
赵俊皓, 张俊文, 迟楠	29-5-45	
基于机器学习的智能路由解释方法		
孟子立, 徐明伟	29-5-56	
网络智能传输研究进展	廖乙鑫, 王子逸, 崔勇	29-5-61
NetGPT: 超越个性化生成服务的内生智能网络架构		
陈宇轩, 李荣鹏, 张宏纲	29-5-68	

## 面向双碳的新一代无线通信网络

专题导读	葛晓虎, 李建东	29-6-01
基于多点信道海图的低能耗网络覆盖结构优化		
叶远青, 张四海, 朱近康	29-6-03	
面向5G-A的无线网络节能关键技术		
郭诚, 陈梦竹	29-6-11	
面向节能减排的跨制式融合感知通信系统		
周均翼, 周琳, 张舜卿	29-6-16	
基于多时隙业务联合整形的低能耗资源调度方法		
李建东, 牛淳隆, 赵晨曦, 刘俊宇	29-6-23	
低开销智能反射面辅助无线通信研究综述		
陈天贝, 李娜, 陶小峰	29-6-29	
基于视距概率模型的6G室内无线通信系统性能分析		
毛曦晨, 王承祥, 杨松江, 黄杰, 黄晨	29-6-39	

## 专家论坛

关于发展中国安全浏览器的建议		
魏小强, 张义荣, 黄亚洲	29-1-51	
2022年度十大天线技术进展	张跃平	29-2-67
智能计算中心发展态势研究	焦奕硕, 邱绍岩	29-3-59
面向新型智算中心的全调度以太网技术		
段晓东, 程伟强, 王瑞雪, 王雯萱	29-4-57	
6G新型信息通信网络架构设计		
段晓东, 孙滔, 刘超, 施南翔	29-5-76	
6G绿色网络发展趋势和关键技术		
李福昌, 郭希蕊	29-6-54	
零信任关键技术与产业发展研究		
张云畅, 柴瑶琳, 穆域博	29-6-60	

## 企业视界

新型家庭全光网技术	王新余, 孔雪, 贺峰	29-1-56
通感一体化关键技术与应用	李萍, 郭晓江	29-2-72
5G XR关键技术挑战和方案研究		
王新台, 袁知贵	29-3-64	
数据管理系统发展趋势与挑战		
韩银俊, 牛家浩, 屠要峰	29-4-64	
相干光收发器件未来技术演进	沈百林, 王会涛	29-5-82
集成电路产业技术发展趋势探讨		
刘新阳, 晁沛荫, 李婷宇	29-6-66	

## 技术广角

基于两跳IRS辅助的下行无线能量和上行信息传输WPCN性能优化	冯璇, 吕斌, 杨震	29-1-63
面向卫星通信系统的寻呼方法	毛玉欣, 闫新成	29-1-72
神经辐射场加速技术综述	郑清芳	29-2-79
光接入网络遥测技术的分析与实践		
刘刚, 张德智, 陈爱民	29-3-73	
基于5G连接的集中式PLC新型工业组网架构		
邓伟, 于天意, 侯庆东	29-4-72	
基于数字子载波和概率整形的相干光通信系统设计及应用		
陆源, 牛文林, 王永奔, 胡子荷	29-4-78	
软件定义广域网络(SD-WAN)2.0测试方法		
党小东, 柴瑶琳, 穆域博	29-6-71	

# 喜报

## 《中兴通讯技术》入选《中文核心期刊要目总览》



### 《中文核心期刊要目总览》2023年版入编通知

《中兴通讯技术》主编先生/女士：

我们谨此郑重通知：依据文献计量学的原理和方法，经研究人员对相关文献的检索、统计和分析，以及学科专家评审，贵刊《中兴通讯技术》入编《中文核心期刊要目总览》2023年版（即第10版）之电子技术、通信技术类的核心期刊。

《中文核心期刊要目总览》2023年版从2021年10月开始研究，研究工作由北京大学图书馆主持，共32个单位的148位专家和工作人员参加了本项研究工作，全国各地9473位学科专家参加了核心期刊表的评审工作。经过定量筛选和专家定性评审，从我国正在出版的中文期刊中评选出1987种核心期刊。

评选核心期刊的工作是运用科学方法对各种刊物在一定时期内所刊载论文的学术水平和学术影响力进行综合评价的一种科研活动。该研究成果只是一种参考工具书，主要是为图书情报界、出版界等需要对期刊进行评价的用户提供参考，例如为各图书情报部门的中文期刊采购和读者导读服务提供参考帮助等，不应作为评价标准。谨此说明。

顺颂  
撰安

编号：2023-J30056

《中文核心期刊要目总览》  
2023年版编委会  
图书馆



# 中兴通讯技术杂志社

## 促进产学研合作青年专家委员会

**主任** 陈为(北京交通大学)

**副主任** 秦晓琦(北京邮电大学) 卢丹(中兴通讯股份有限公司)

**委员** (按姓名拼音排序)

曹进	西安电子科技大学	秦志金	清华大学
陈力	中国科学技术大学	史颖欢	南京大学
陈琪美	武汉大学	王景璟	北京航空航天大学
陈舒怡	哈尔滨工业大学	王兴刚	华中科技大学
陈为	北京交通大学	王勇强	天津大学
官科	北京交通大学	温森文	华南理工大学
韩凯峰	中国信息通信研究院	吴泳澎	上海交通大学
何姿	南京理工大学	夏文超	南京邮电大学
胡杰	电子科技大学	徐梦炜	北京邮电大学
黄晨	紫金山实验室	徐天衡	中国科学院上海高等研究院
李昂	西安交通大学	杨川川	北京大学
刘春森	复旦大学	尹海帆	华中科技大学
刘凡	南方科技大学	于季弘	北京理工大学
刘俊宇	西安电子科技大学	张娇	北京邮电大学
卢丹	中兴通讯股份有限公司	张宇超	北京邮电大学
陆游游	清华大学	章嘉懿	北京交通大学
宁兆龙	重庆邮电大学	赵昱达	浙江大学
祁亮	上海交通大学	周伊	西南交通大学
秦晓琦	北京邮电大学	朱秉诚	东南大学

### 刊物相关信息



投稿须知



投稿平台



过刊下载



论文索引与  
引用指南

# 中兴通讯技术

(ZHONGXING TONGXUN JISHU)

## 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

## 产业顾问（按姓名拼音排序）：

段向阳、高 音、胡留军、华新海、刘新阳、  
陆 平、史伟强、屠要峰、王会涛、熊先奎、  
赵亚军、赵志勇、朱晓光

双月刊 1995 年创刊 总第 173 期

2023 年 12 月 第 29 卷 第 6 期（卷终）

总编辑：王喜瑜

主编：蒋贤骏

执行主编：黄新明

编辑部主任：卢丹

责任编辑：徐烨

编辑：杨广西、朱莉、任溪溪

设计排版：徐莹

发行：王萍萍

编务：王坤

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

《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

电子邮箱：magazine@zte.com.cn

电话：(0551) 65533356

发行方式：自办发行

印刷：合肥添彩包装有限公司

出版日期：2023 年 12 月 26 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