



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

ISSN 1009-6868
CN 34-1228/TN

中兴通讯技术

ZTE TECHNOLOGY JOURNAL

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

2021年12月·第6期

专题：触觉通信技术



1 2 >

9 771009 686212



《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) 糜正琨(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 | 唐雄燕 | 中国联通研究院副院长 |
| 陈前斌 | 重庆邮电大学教授、副校长 | 陶小峰 | 北京邮电大学教授 |
| 段晓东 | 中国移动研究院副院长 | 王文博 | 北京邮电大学教授、副校长 |
| 葛建华 | 西安电子科技大学教授 | 王文东 | 北京邮电大学教授 |
| 管海兵 | 上海交通大学教授 | 王喜瑜 | 中兴通讯股份有限公司执行副总裁 |
| 郭庆 | 哈尔滨工业大学教授 | 王翔 | 中兴通讯股份有限公司高级副总裁 |
| 洪波 | 中兴发展股份有限公司总裁 | 王耀南 | 中国工程院院士 |
| 洪伟 | 东南大学教授 | 卫国 | 中国科学技术大学教授 |
| 黄宇红 | 中国移动研究院副院长 | 吴春明 | 浙江大学教授 |
| 纪越峰 | 北京邮电大学教授 | 邬贺铨 | 中国工程院院士 |
| 江涛 | 华中科技大学教授 | 向际鹰 | 中兴通讯股份有限公司首席科学家 |
| 蒋林涛 | 中国信息通信研究院科技委主任 | 肖甫 | 南京邮电大学教授 |
| 李尔平 | 浙江大学教授 | 解冲锋 | 中国电信研究院教授级高工 |
| 李红滨 | 北京大学教授 | 徐安士 | 北京大学教授 |
| 李厚强 | 中国科学技术大学教授 | 徐子阳 | 中兴通讯股份有限公司总裁 |
| 李建东 | 西安电子科技大学教授 | 续合元 | 中国信息通信研究院副总工 |
| 李乐民 | 中国工程院院士 | 薛向阳 | 复旦大学教授 |
| 李融林 | 华南理工大学教授 | 薛一波 | 清华大学教授 |
| 李少谦 | 电子科技大学教授 | 杨义先 | 北京邮电大学教授 |
| 李自学 | 中兴通讯股份有限公司董事长 | 叶茂 | 电子科技大学教授 |
| 林晓东 | 中兴通讯股份有限公司副总裁 | 易芝玲 | 中国移动研究院首席科学家 |
| 刘健 | 中兴通讯股份有限公司高级副总裁 | 张宏科 | 中国工程院院士 |
| 刘建伟 | 北京航空航天大学教授 | 张平 | 中国工程院院士 |
| 隆克平 | 北京科技大学教授 | 张钦宇 | 哈尔滨工业大学教授 |
| 陆建华 | 中国科学院院士 | 张卫 | 复旦大学教授 |
| 马建国 | 浙江大学教授 | 张云勇 | 中国联通集团产品中心总经理 |
| 毛军发 | 中国科学院院士 | 赵慧玲 | 工业和信息化部通信科技委信息通信网络专家组组长 |
| 孟洛明 | 北京邮电大学教授 | 郑纬民 | 中国工程院院士 |
| 任品毅 | 西安交通大学教授 | 钟章队 | 北京交通大学教授 |
| 石光明 | 西安电子科技大学教授、副校长 | 周亮 | 南京邮电大学教授 |
| 孙知信 | 南京邮电大学教授 | 朱近康 | 中国科学技术大学教授 |
| 谈振辉 | 北京交通大学教授、原校长 | 祝宁华 | 中国科学院院士 |
| 唐宏 | 中国电信IP领域首席专家 | | |

目次

中兴通讯技术 (ZHONGXING TONGXUN JISHU)
总第 161 期 第 27 卷 第 6 期 (卷终) 2021 年 12 月

专题：触觉通信技术

专题导读 01
周亮

触感技术的现状与发展 02
曾繁荣, 房颖, 赵铁松

触觉通信中的信号获取与恢复关键技术 08
刘恒发, 魏昕

面向 6G 的触觉互联网关键技术与挑战 15
尤子硕, 李强, 唐冰, 葛晓虎

面向触觉通信的预测窗口和计算卸载策略优化 21
吴岩, 吴丹, 王嵩

一种面向视触融合人机交互的柔性触觉传感阵列 29
王爽, 吴兵, 刘倩

专家论坛

触觉通信技术发展的若干思考 35
陈鸣镔, 周亮

企业视界

40 800 Gbit/s 光模块技术及应用
王会涛, 张平化, 苏展

技术广角

47 TSN 与 5G 融合部署的需求和网络架构演进
朱瑾瑜, 张恒升, 陈洁

53 基于典型无线场景库的 LTE-V2X 信道特性
苏昭阳, 刘留, 冯毅

58 5G 电信云网络的容器演进方案
史庭祥, 徐法禄, 章璐

65 基于对称密钥算法的 5G 网络防伪基站技术
毛玉欣, 游世林, 闫新成

综合信息

I 《中兴通讯技术》第 27 卷总目次

III 《中兴通讯技术》2022 年专题计划

2021 年第 1—6 期专题计划及策划人

1. 视频技术和用户体验评测

华中科技大学教授 江涛
中兴通讯股份有限公司副总裁 陆平

2. 6G 愿景及技术挑战

中国工程院院士 张平
北京邮电大学教授 张建华

3. 边缘计算与算力网络

工信部通信科技委信息通信网络
专家组组长 赵慧玲

4. 高铁智能通信技术与应用

北京交通大学教授 艾渤

5. 低轨卫星通信技术与应用

哈尔滨工业大学教授 郭庆

6. 触觉通信技术

南京邮电大学教授 周亮

MAIN CONTENTS

ZTE TECHNOLOGY JOURNAL Vol. 27 No. 6 Dec. 2021

Special Topic:

Haptic Communication Technology

Editorial **01**
ZHOU Liang

Status and Development of Haptic Technology **02**
ZENG Fanrong, FANG Ying, ZHAO Tiesong

Key Techniques of Signal Acquisition and Recovery in
Haptic Communication **08**
LIU Hengfa, WEI Xin

Key Technologies and Challenges of 6G-Oriented
Tactile Internet **15**
YOU Zishuo, LI Qiang, TANG Bing, GE Xiaohu

Optimization of Prediction Window and Computing
Offloading Strategy for Haptic Communication **21**
WU Yan, WU Dan, WANG Song

A Novel Flexible Haptic Sensor Array for
Visual Haptic Fusion Human-Machine Interaction **29**
WANG Shuang, WU Bing, LIU Qian

Expert Forum

35 Reflections on Development of the Haptic Communication
CHEN Mingkai, ZHOU Liang

Enterprise View

40 Technologies and Applications of 800 Gbit/s Optical
Module
WANG Huitao, ZHANG Pinghua, SU Zhan

Technology Perspective

47 Requirements and Network Architecture Evolution of
TSN and 5G Integrated Deployment
ZHU Jinyu, ZHANG Hengsheng, CHEN Jie

53 Channel Characteristics of LTE-V2X Based on
Typical Wireless Scenario Library
SU Zhaoyang, LIU Liu, FENG Yi

58 Evolution Solution for Container-Based 5G Telecom
Cloud Network
SHI Tingxiang, XU Falu, ZHANG Lu

65 Anti-Fake Base Station in 5G Network Based on
Symmetric Cryptography Algorithm
MAO Yuxin, YOU Shilin, YAN Xincheng

期刊基本参数: CN 34-1228/TN*1995*b*16*71*zh*P*¥20.00*6500*12*2021-12

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。



触觉通信技术 专题导读

专题策划人



周亮

南京邮电大学教授、博士生导师、通信与信息工程学院院长，教育部宽带无线通信与传感网技术重点实验室主任；先后获教育部“长江学者奖励计划”特聘教授、中共中央组织部“海外高层次青年专家”等荣誉称号，获国家自然科学基金委员会“优秀青年基金”资助；作为项目负责人和主要完成人，主持并参与了多项国家级重点科技攻关项目；发表论文多篇，拥有授权专利 25 项。

触觉是人类感知并与外部环境交互的一类重要方式，它通过来自皮肤表面的敏感神经来感受接触物的材质、运动、压力、温度等属性。区别于传统网络信息传输，触觉通信需要通过对触觉控制信息的传输和反馈来进行精细的动作控制，通过与实际或虚拟的物体进行远程实时交互来实现“技能传输”，通过视、听、触的多感知融合，实现“所见即所触”。近年来，触觉通信技术受到了全球学术界和工业界的极大关注。其中，国际电信联盟（ITU）和电气与电子工程师协会（IEEE）等国际标准化组织已着手触觉通信相关领域的标准制订。未来，触觉通信将在智能交通、远程医疗、工业物联网、军事训练、远程教育等人机交互场景中得到广泛应用。

触觉通信的关键挑战之一是在无线移动场景下实现高可靠低延迟的触觉信号传输。针对触觉通信中的关键技术和挑战，本期专题的多篇文章展开了讨论。《触感技术的现状与发展》探讨了触感信号的主要特点和触感交互面临的挑战，并对触感信息获取与再现、编码、传输等技术的研究进行了综合讨论。《触觉通信中的信号获取与恢复关键技术》聚焦于触觉通信中的发送端信号获取和接收端触觉恢复技术，对触觉感受、触觉信号的采集设备以及触觉信号的表征进行了描述，并针对触觉通信过程中可能存在的信号缺失、干扰等问题提出了同模态和跨模态的触觉信号恢复技术。《面向 6G 的触觉互联网关键技术与挑战》从

网络架构、信息传输、资源管理 3 个层面系统性地提出了触觉通信关键使能技术框架。《面向触觉通信的预测窗口和计算卸载策略优化》构建了联合预测和边缘计算的触觉通信框架，明确了时延和能耗之间的折中关系，并提出了发送者的预测效益概念，并以总预测效益最大化为目的，定义了多发送者的预测窗口和计算卸载策略联合优化问题。《一种面向视触融合人机交互的柔性触觉传感阵列》提出了一种面向视触融合人机交互的三维柔性触觉传感器以及相应的触觉采集软硬件装置，用于采集真实物理操作过程中的触觉力信息。

本期的作者来自于知名高校与科研机构。面向触觉通信技术与应用，作者们从网络架构、信号处理、性能评估等方面介绍了该领域最新的研究成果和经验。希望本期的内容能为读者提供有益的借鉴与启示。在此，对所有作者的大力支持表示由衷的感谢！

周亮

2021 年 11 月 20 日

触感技术的现状与发展

Status and Development of Haptic Technology



曾繁荣/ZENG Fanrong, 房颖/FANG Ying, 赵铁松/ZHAO Tiesong

(福州大学 福建省媒体信息智能处理与无线传输重点实验室, 中国 福州 350108)

(Fujian Key Lab for Intelligent Processing and Wireless Transmission of Media Information, Fuzhou University, Fuzhou 350108, China)

摘要:触觉作为互联网的第三种媒体流,将广泛应用于未来的触感应用中。在触觉交互任务中,要将用户的体验质量维持在最佳水平,仍是一个具有挑战性的任务。介绍了触觉信号的主要特点和触觉交互面临的挑战,并对触觉信息获取与再现、编码、传输等技术的研究进行了综合讨论。未来多模态的触觉设备、高效率的触觉编码与双边遥操作控制方法的相互融合,以及标准的网络传输协议的发展,都将推动触感技术在多媒体交互任务中的应用。

关键词: 触觉; 肤感信号; 触觉编码; 触觉通信

Abstract: As the third media stream of the Internet, haptics will be widely used in haptic applications. However, it is still a challenge to maintain the user's quality of experience at the optimal level in the haptic interaction. The main characteristics of haptic signals and the challenges of haptic interaction are introduced and the technologies of haptic signals acquisition, reproduction, coding, and transmission are analyzed. In the future, the mutual integration of multi-modal tactile devices, high-efficiency tactile coding, and bilateral tele-operation control methods, as well as standard network transmission protocols will all promote the growth of tactile technology in multimedia interactive tasks.

Keywords: kinesthetic; tactile signal; haptic encoding; haptic communication

DOI: 10.12142/ZTETJ.202106002

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211208.1032.008.html>

网络出版日期: 2021-12-08

收稿日期: 2021-11-08

随着数字信号处理及多媒体通信技术的飞速发展,人们的视听需求得到了极大的满足。人们开始追求多媒体交互体验的沉浸感,希望实现从视听互动到多感知交互的转变。其中,触觉作为仅次于听觉与视觉的第三种交互媒体,为人机互动增加了一个新的维度。触感技术根据反馈信息实现了触觉与肤感的再现,人类或机器可以远程且实时地感知或控制真实或虚拟物体。研究表明,将触感技术整合到现有的视听多媒体服务中,实现人与环境的双向交

流,可以增强多媒体交互性及用户的沉浸感^[1]。视、听、触的多感知融合,能够实现“所见即所触”,在智能交通、远程医疗、工业物联网等人机交互场景中得到广泛应用。

触觉交互系统是一个“人在回路”的混合闭环控制系统,以实现双边触觉远程操作。融合触觉的多媒体交互系统典型结构如图1所示^[2],由主控域、网络域和受控域组成。主控域通常由操作人员和人机触觉交互界面组成。人机触觉交互界面实现触觉信号(位置、速度等信号)的采

集、处理与传输,同时接收受控域的多模态(听觉、视觉、触觉)信号的反馈,从而实现远程交互环境的感知与交互命令的调整。在受控域中,远程控制的机器人或受控操作人员执行主控域的远程交互命令,同时将场景信息通过多模态信号反馈到主控域中。主控域与受控域之间通过网络域上的双向数据通信链路形成全局控制回路,从而实现多感知的交互。

触感技术使现实或虚拟环境的完全沉浸式多感官远程体验成为可能,将成为未来交互应用中的第三种

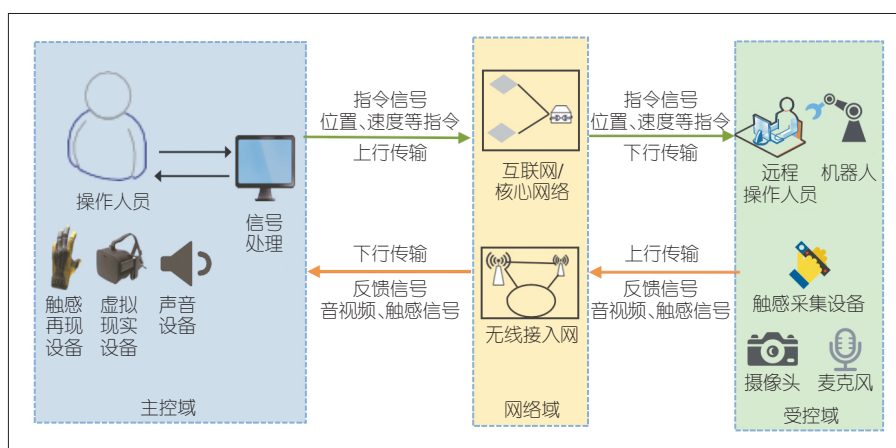
媒体流。然而,将用户的体验质量维持在最佳水平,仍是一个具有挑战性的任务。因此,目前学者对触觉信息获取及呈现、压缩编码、信息传输等方面进行了研究。电气与电子工程师协会(IEEE)等国际标准组织也着力于P1918.1触觉互联网相关标准的制定。本文讨论触觉信息获取及再现技术、触觉信号的压缩技术,同时结合触觉网络传输架构及触觉数据通信协议等,对触觉媒体传输关键技术进行了总结。

1 触觉信息的获取与再现

触觉包括位于关节和韧带内的动觉感知和位于皮肤真皮层及表皮层的肤感。动觉感知包括力、扭矩、位置、速度、角速度等信息。肤感包括静态压力、振动、摩擦、表面纹理、皮肤拉伸、三维形状、热觉和痛觉等^[3-4]。触觉反馈技术基于人体触觉感知机制,通过生成刺激信号,再现了触觉交互过程。动觉和肤觉由于记录的刺激数据不同,适用的触觉设备也有所差异。

1.1 动觉信息获取与再现

动觉信息(力、扭矩、位置、方向)的获取与再现目前主要是利用力反馈设备^[5]。力反馈设备包括传感器和由直流电动机控制的执行器。传感器提供关于设备在虚拟/现实世界中的位置和方向的信息。一旦设备与物体发生互动,执行器就会向用户显示所产生的力/扭矩。如图2所示,力反馈设备包括接地设备与非接地设备:Novint Falcon、Phantom的Omni和Premium等接地设备将其反作用力施加在一个巨大的静止物体上;Rutgers Master、Gyro Moment Display等非接地设备能够产生自平衡力,无须机械接地,便能将其反作用力施加在操作



▲图1 融合触觉媒体的交互系统



▲图2 交互信息获取与再现方式

者身体的某个部位。

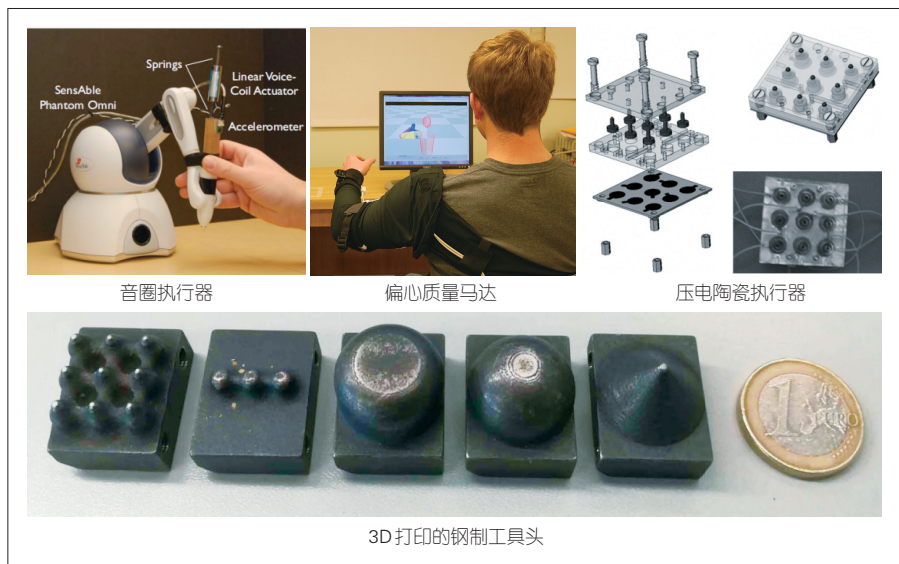
1.2 肤感信息获取与再现

传感器是产生肤感的设备,通常安装在用户终端,用于感测与周围环境交互产生的振动、温度等信息,并反馈传达给用户。目前传感器主要是对摩擦力、硬度、温度、粗糙度进行测量^[6]。相应的触觉再现设备主要是基于振动、超声波以及静电驱动方式^[6]的。振动触觉执行器可以使用音圈执行器、偏心质量马达、压电陶瓷执行器或3D打印的钢制工具头,如图3所示。

以上触感再现技术只关注到个别的触觉维度,而多模态触觉反馈设备能同时刺激人的肤感与动觉感知,产生多个维度的触觉刺激,包括力、

振动、扭矩、位置、热刺激等,使用户感知到物体的多种属性,支持更为精细的操作。目前,多模态触觉反馈设备可利用运动学机制、软执行器、针式矩阵和热显示器、静电效应、平台渲染物体表面、音圈与制动器的融合,或是内置紧凑的触觉显示器和一个振动模块来实现用户的多模态触觉反馈^[4]。例如,H.CULBERTSON和K.J.KUCHENBECKER利用Phantom Omni设备和振动机械执行器产生了特定触觉刺激(摩擦、硬度和粗糙度)^[6],利用螺线管柱塞和一个滚动的不锈钢球来呈现不同的摩擦力^[7]。

但是到目前为止,只有少数设备可以同时模拟3种基本模式(柔软度、纹理和热反馈),没有设备能同时模拟更多的模态。为了产生多模态的



▲图3 振动触觉常用执行器

触觉,新型触觉设备可以使用能同时支持多模态启动的新型功能材料,或将触觉设备执行器内嵌多个传感器。需要注意的是,当多种模式被整合时,不同制动器之间可能产生空间干扰;当用户同时感知多种模式时,某些模式组合的耦合效应与用户的感知阈值会出现变化。

2 触觉信号的编码

触觉信号采样率越高,信号幅值则越小。动觉信号的采样率为1 kHz甚至更高,位置和力的信号集中在 $-1.0 \sim +1.0$ 和 $-3.0 \sim +3.0$ 的范围内^[8],如图4所示^[9]。肤感信息包括硬度、导热性、摩擦力、微观粗糙度和宏观粗糙度,目前主要通过振动触觉信号来解决微观粗糙度和摩擦力问题。在高精度触觉交互任务中,振动触觉信号采样率在5 kHz以上。在数据库^[10]中,振动触觉信号采样率为2 800 Hz,其信号幅度值为0~1.5,如图5所示。

较高的采样率需要较高的数据包传输率,这会产生大量的数据开销,从而导致延迟的增加。随着触觉应用所需的自由度数量的增加,触觉数据将与之成比例地增加。尽管5G

网络的数据传输能力可以满足用户的触觉数据传输需求,但在实际的多媒体交互任务中,还需要同时传输音频和视频数据,因此必须对触觉数据进行压缩。动觉信号与肤感信号分别应用于不同的场景,对于通信延迟

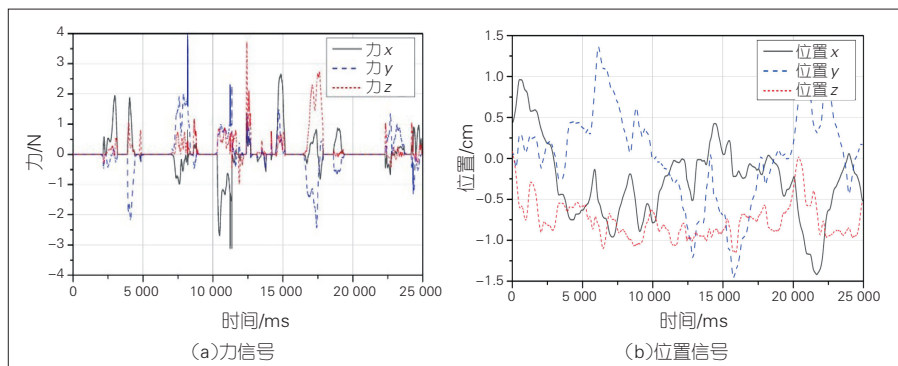
有着不同的需求,这两类信号的压缩有其根本区别。

2.1 动觉信号的编码

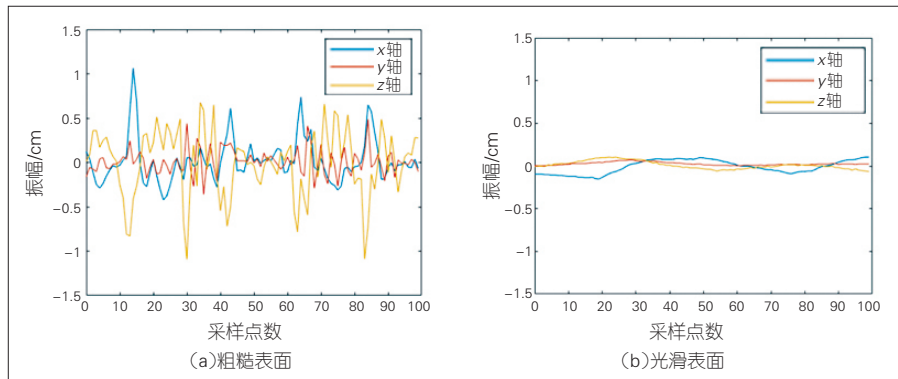
动觉信号主要应用于闭环的通信场景中,对网络延迟有着较高的要求,因此动觉编码器的主要目标是降低网络数据传输时的平均数据包率,同时保持较高的用户体验质量。目前,动觉信号编码主要包括实时编码及延迟编码^[11]:前者在完成信号采样的同时进行数据传输,在交互过程中不需要控制机制来保证交互稳定性;后者则允许在通信延迟超过5 ms的情况下,在交互过程中增加系统稳定控制机制。

(1) 实时编码

实时编解码器主要基于感知死区的动觉压缩方案^[11]。该类编解码器是基于韦伯的可察觉差异定律(JND)的,如公式(1)所示。只有当前信号 X 与之前传输的最后动觉信号



▲图4 动觉信号实例



▲图5 肤感信号实例

X_{n-1} 之间的相对差异超过韦伯分数 δ 时,信号 X 才需要被传输,否则当前信号将被丢弃。

$$\left| \frac{X - X_{n-1}}{X_{n-1}} \right| \geq \delta \quad (1)$$

δ 在编码方案中被称为死区参数。死区参数阈值决定了压缩率和感知的失真。 δ 值较小将导致压缩率不足; δ 值过大可能导致显著的失真。图6^[12]为一维动觉信号的感知死区压缩原理。带黑点的采样点代表输出动觉信号,灰色区域定义了感知死区。如果信号样本值位于定义的感知死区范围内,则说明信号间的相对差异值较小且无法被人所感知。在编码过程中,该部分信号将被舍弃。

基于感知死区及预测模型的感知编码可以进一步降低触觉数据包的速率。如图7^[3]所示,相同的预测器在主控域和受控域两边并行。在发送端,预测器在每个采样瞬间产生预测的触觉信号。如果预测误差小于相应的感知死区阈值,则不触发更新;否则,输入样本被传输到另一侧,同时传输的样本被用于更新预测模型。如果接收端收到了一个数据包,就直接作为动觉输出,收到的触觉信号被用于更新预测模型;否则,接收端预测器会产生一个预测的触觉信号作为当前的输出。基于感知死区的方法可能会引起用户感知的“抖动效应”,并且随着触觉设备自由度的增加,压缩率可能会有所下降。

(2) 延迟编码

延迟编码器用额外的延迟来换取更好的编码效率。延迟可能会影响主观体验和客观质量,甚至影响触觉设备的可用性,因此通常需要增加一个控制机制来保持系统的稳定性。差分脉冲编码调制(DPCM)以及自适应差分脉冲编码调制(ADPCM)技术

被引入该类编解码器中,压缩率可以达到11:1,同时将视觉和动觉失真保持在人类生理感知的阈值以下。除此之外,离散余弦变换(DCT)、快速DCT(FDCT)以及小波包变换(WPT)算法也应用于单一自由度的动觉系统中。其中,在DCT变换中加入气泡排序算法的选择DCT(sDCT)算法^[13],使得动觉信号的压缩率和信噪比进一步提高。文献[9]提出了一种高效、低延迟、感知无损的混合结构动觉编解码器。该解码器能够减少编码延迟,优化熵编码器,并通过预处理消除编码损失。该编码器由放大器、DCT、量化器、行程编码、熵编码器组成。利用文献[9]提出的算法可使位置信号和力信号的传输率比其他算法降低了50%以上,最大缓冲延迟仅为7 ms,有效保证了动觉交互的用户体验。

2.2 肤感信号的编码

肤感信息通常发生在开环通信场景中,对延迟的容忍度有所提高。目前肤感信号编解码器通常是指振动触觉数据的压缩,主要包括波形编码、参数编码以及混合感知编码。

(1) 波形编码

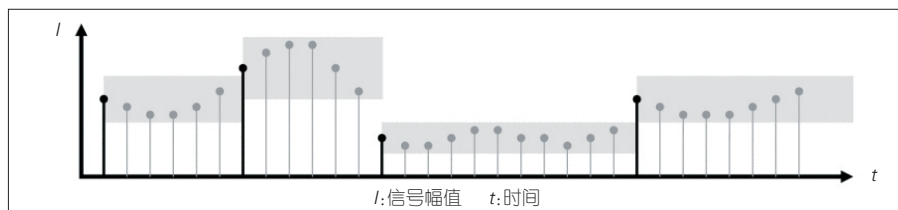
波形编码可以使用正交线性变换,例如DCT或离散小波变换(DWT),将振动触觉信号转换至其他数据域中,通过量化消除振幅低于阈值的信号,减少原始信号中的关联性再进行熵编码。

(2) 参数编码

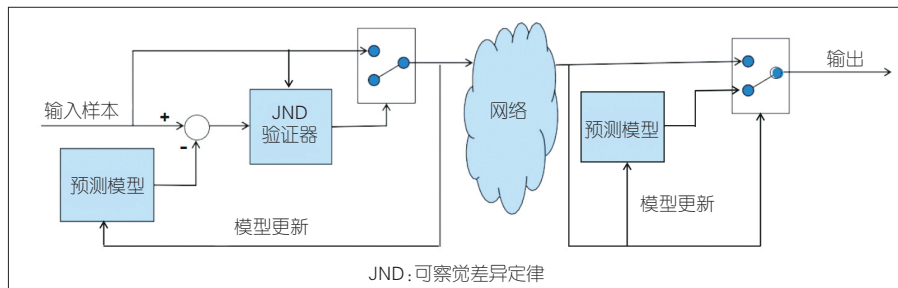
肤感特征可以被提取,以形成参数化表征。例如,物体表面的摩擦、粗糙度等被捕捉后,每种材料的特征向量被发送到远端触觉渲染框架,用于在虚拟环境中再现触觉感知。肤感信号特征是基于人为提取或基于深度学习自动提取的。基于线性预测编码(LPC)的振动触觉信号压缩方法采用分析合成技术,只存储或传输于分析阶段捕获的滤波器参数中^[14]。

(3) 混合感知编码

感知压缩的目标是通过尽可能多地保留可感知信号信息,来降低所需的数据传输速率。目前性能较好的感知小波量化振动编解码器(VCPWQ)^[16]利用心理触觉模型,将损伤阈值纳入感知绝对阈值函数;利用感知的绝对阈值和掩蔽现象,动态提取其接收到的输入信号的必要属性;同时使用DWT有效地去除输入信号的相



▲图6 感知死区原理



▲图7 基于感知死区及预测模型的感知编码机制

关性;采用嵌入式均匀量化器,使用多级树集合分裂(SPIHT)算法及算术编码,得到了较好的触觉编码性能。基于稀疏线性预测的感知振动触觉信号压缩算法(PVC-SLP)^[16]则采用了加速敏感度函数(ASF)以构建触觉敏感度模型。该压缩算法在线性预测方案中对残差和预测系数都引入了稀疏性约束,目前被选为IEEE P1918.1.1触觉编解码标准的一部分。

迄今为止,业界开发的触觉数据压缩算法只适用于动觉或肤感信号。但在现实中,动觉信号和肤感信号是相互影响的。因此,有必要设计一种兼容两种信号的通用编码方法^[1]。此外,大部分算法只解决单点触觉互动。随着触觉数据采集的改进,多点场景的重要性越来越高,因此需要开发针对多点触觉场景的压缩算法,并可以考虑利用通道间或空间的相关性来获得最大的压缩性能。最后,必须定义多模态编码方案,以支持触觉互联网信息不同模态,而不增加端到端的延迟。

3 触觉信号的传输

触觉交互实现了实时感知、操纵或控制真实和虚拟的物体或进程,其信号本质上是主动和双向的。互联网需要由原先的内容传输进一步演变为技能传输。相比于视频、声音、图像等其他多媒体应用,触觉应用对数据丢失和带宽的容忍度高,对延迟、可靠性的要求更高。为了保证更自然的交互操作,触觉互联网必须是高可靠低延时的,因此它的网络架构及传输协议都与传统互联网不同。

3.1 网络体系结构

触觉通信网络架构对于确保低时延、高可靠、鲁棒性,以及与其他网络的兼容至关重要。触觉互联网架

构分为独立于具体应用的通用结构和与具体应用相关联的专用架构。

通用架构包括IEEE P1918.1标准组提出的通用架构^[11,17]、移动边缘计算(MEC)增强蜂窝架构^[18]、基于以太网无源光网络(EPON)的网络架构^[19]、FiWi增强型蜂窝架构^[20](HetNet)以及新型互联网架构FlexNGIA^[21]。

IEEE P1918.1标准组提出的触觉互联网(TI)通用架构,旨在支持大多数TI的实际使用。该通用结构由网络域和边缘域组成,包括触觉设备、网关节点、网络控制器、支持引擎、触觉服务管理器、用户平面实体和控制平面实体。每个触觉边缘的触觉设备可以通过网络域(包括共享无线网络、共享有线网络、专用无线网络、专用有线网络)进行触觉信息的通信。除了传统的物理接口外,参考架构还确定了各种逻辑接口,这些接口可以通过包括5G在内的任何连接技术,映射到各种触觉互联网应用程序。

边缘计算增强蜂窝架构为一个多层云系统,该系统在蜂窝网络中提供迁移能力,以支持触觉互联网应用。该系统分别由在最低、中间和最高级别上的微云、小型云和核心云单元组成。该架构旨在将蜂窝网络的业务转移到云端,以减少数据往返的延迟和网络拥塞。

基于EPON^[18]的光云分布网络(OCLDN)架构提供高速、可编程可扩展的光接入网络,以支持低延迟服务。它允许光网络单元(ONU)同时通过多个波长通道传输流量,增加网络容量。

HetNet通用架构具有光纤回程共享和WiFi迁移功能。它利用EPON技术在核心和边缘网络之间形成回程链路,并利用千兆无线局域网(WLAN)技术构建FiWi接入网络。

新型互联网架构FlexNGIA则利

用未来网络基础设施中计算资源的可用性,来实现网络内计算。该架构还定义了一个商业模式:网络运营商不仅可以提供数据传输,还可以提供对性能、可靠性等有严格要求的服务功能链。该架构可使网络提供更好的拥塞控制和可靠性服务。

触觉互联网专用架构主要针对具体应用设计,包括实时远程控制任务、虚拟现实、工业应用、智慧城市等。其中,文献[22]提出了一种全新的体验质量(QoE)驱动的智慧城市触觉互联网专用架构。该架构包含5层:感知层、传输层、存储层、计算层和应用层。感知层的主要任务是从智慧城市中的分布式传感器收集大数据,然后对数据进行预处理;传输层旨在为智慧城市的数据传输提供超可靠的连接;存储层利用分布式数据存储和区块链技术,为数据存储和用户隐私信息保护方面提供保障;计算层使用渗透计算技术实现对触觉设备的控制;应用层通过挖掘触觉和现有的视听数据提升用户的QoE。

3.2 传输协议

近年来,针对触觉数据的特殊性,研究人员提出了新的触觉通信中数据传输协议、应用层迁移协议、切换协议以及触觉握手协议。触觉通信协议(FRTPS)^[23]基于实时传输协议(RTP)协议,使用连续网络探测流控制机制和模糊逻辑速率自适应机制,提高了触觉数据传输性能。敏捷云迁移(ACM)的应用层迁移协议^[24]由客户端、控制器和引擎组成,为触觉交互提供更高的可靠性。文献[25]提出的垂直切换协议,可实现IEEE 802.11p和长期演进(LTE)网络之间的无缝切换。IIYOSHI等^[26]着重设计了一种触觉握手机制,以协调(通常)异构的触觉设备。他们还开发了“触

觉互联网信息”的信息传递格式,用于在触觉握手期间交换元数据。

4 结束语

触觉交互系统是一个“人在回路”的混合闭环控制系统,它可以实现双边触觉远程操作,人类或机器可以实时获取、感知、操作或控制远端真实或虚拟的物体。本文分别对触觉信息的获取与再现、数据压缩编码技术及数据传输架构与触觉通信协议进行了总结与分析,并对相应技术未来的挑战进行了探讨。

参考文献

- [1] ZHOU L, WU D, CHEN J, et al. Cross-modal collaborative communications [J]. IEEE wireless communications, 2020, 27(2): 112–117. DOI: 10.1109/MWC.001.1900201
- [2] SHARMA S K, WOUNGANG I, ANPALAGAN A, et al. Toward tactile internet in beyond 5G era: recent advances, current issues, and future directions [J]. IEEE access, 2020, 8: 56948–56991. DOI: 10.1109/ACCESS.2020.2980369
- [3] ANTONAKOGLU K, XU X, STEINBACH E, et al. Toward haptic communications over the 5g tactile Internet [J]. IEEE communications surveys & tutorials, 2018, 20(4): 3034–3059. DOI: 10.1109/COMST.2018.2851452
- [4] WANG D, OHNISHI K, XU W. Multimodal haptic display for virtual reality: a survey [J]. IEEE transactions on industrial electronics, 2020, 67(1): 610–623. DOI: 10.1109/TIE.2019.2920602
- [5] STEINBACH E, STRESE M, EID M, et al. Haptic codecs for the tactile Internet [J]. Proceedings of the IEEE, 2019, 107(2): 447–470. DOI: 10.1109/JPROC.2018.2867835
- [6] CULBERTSON H, KUCHENBECKER K J. Importance of matching physical friction, hardness, and texture in creating realistic haptic virtual surfaces [J]. IEEE transactions on haptics, 2017, 10(1): 63–74. DOI: 10.1109/TOH.2016.2598751
- [7] CULBERTSON H, KUCHENBECKER K J. Un-grounded haptic augmented reality system for displaying roughness and friction [J]. IEEE/ASME transactions on mechatronics, 2017, 22(4): 1839–1849. DOI: 10.1109/TMECH.2017.2700467
- [8] IEEE P1918.1.1 Haptic Codecs for the Tactile Internet Task Group. Kinesthetic reference setup data traces [EB/OL]. (2020–03–30) [2021–11–07]. <https://cloud.lmt.ei.tum.de/s/4FmHUCsoUvwRle3>
- [9] ZENG C Y, ZHAO T S, LIU Q, et al. Perception-lossless codec of haptic data with low delay [C]// Proceedings of the 28th ACM International Conference on Multimedia. ACM, 2020: 3642–3650. DOI:10.1145/3394171.3413728
- [10] KIRSCH J, NOLL A, STRESE M, et al. A low-cost acquisition, display, and evaluation setup for tactile codec development [C]// 2018 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE). IEEE, 2018: 1–6. DOI: 10.1109/HAVE.2018.8547512
- [11] HOLLAND O, STEINBACH E, PRASAD R V, et al. The IEEE 1918.1 “tactile Internet” standards working group and its standards [J]. Proceedings of the IEEE, 2019, 107(2): 256–279. DOI: 10.1109/JPROC.2018.2885541
- [12] STEINBACH E, HIRCHE S, KAMMERL J, et al. Haptic data compression and communication [J]. IEEE signal processing magazine, 2011, 28(1): 87–96. DOI: 10.1109/MSP.2010.938753
- [13] BARAN E A, KUZU A, BOGOSYAN S, et al. Comparative analysis of a selected DCT-based compression scheme for haptic data transmission [J]. IEEE transactions on industrial informatics, 2016, 12(3): 1146–1155. DOI: 10.1109/TII.2016.2555982
- [14] CHAUDHARI R, SCHUWERK C, DANAEI M, et al. Perceptual and bitrate-scalable coding of haptic surface texture signals [J]. IEEE journal of selected topics in signal processing, 2015, 9(3): 462–473. DOI:10.1109/JSTSP.2014.2374574
- [15] NOLL A, NOCKENBERG L, GÜLECYÜZ B, et al. VC-PWQ: vibrotactile signal compression based on perceptual wavelet quantization [C]//2021 IEEE World Haptics Conference (WHC). IEEE, 2021: 427–432. DOI:10.1109/WHC49131.2021.9517217
- [16] HASSEN R, GUELECYUEZ B, STEINBACH E G. PVC-SLP: perceptual vibrotactile-signal compression based-on sparse linear prediction [J]. IEEE transactions on multimedia, 2020, 26(7): 99: 1. DOI: 10.1109/TMM.2020.3042674
- [17] AJAZ A, DAWY Z, PAPPAS N, et al. The IEEE P1918.1 reference architecture framework for the tactile Internet and a case study [C]//GLOBECOM 2020 – 2020 IEEE Global Communications Conference. IEEE, 2020: 1–6. DOI: 10.1109/GLOBECOM42002.2020.9322128
- [18] ATEYA A A, VYBORNOVA A, KIRICHEK R, et al. Multilevel cloud based tactile Internet system [C]//2017 19th International Conference on Advanced Communication Technology (ICACT). IEEE, 2017: 105–110. DOI: 10.23919/ICACT.2017.7890067
- [19] NEAIME J, DHAINI A R. Resource management in cloud and tactile-capable next-generation optical access networks [J]. Journal of optical communications and networking, 2018, 10(11): 902–914. DOI: 10.1364/JOCN.10.000902
- [20] BEYRANVAND H, LÉVESQUE M, MAIER M, et al. Toward 5G: FiWi enhanced LTE-A HetNets with reliable low-latency fiber backhaul sharing and WiFi offloading [J]. IEEE/ACM transactions on networking, 2017, 25(2): 690–707. DOI: 10.1109/TNET.2016.2599780
- [21] ZHANI M F, ELBAKOURY H. FlexNGIA: a flexible Internet architecture for the next-generation tactile Internet [J]. Journal of network and systems management, 2020, 28(4): 751–795. DOI: 10.1007/s10922-020-09525-0
- [22] WEI X, DUAN Q, ZHOU L. A QoE-driven tactile Internet architecture for smart city [J]. IEEE network, 2020, 34(1): 130–136. DOI: 10.1109/MNET.001.1900078
- [23] KONTOGIANNIS S, KOKKONIS G. Proposed fuzzy real-time HaPticS protocol carrying haptic data and multisensory streams [J]. International journal of computers communications & control, 2020, 15(4). DOI: 10.15837/ijccc.2020.4.3842
- [24] BRAUN P J, PANDI S, SCHMOLL R, et al. On the study and deployment of mobile edge cloud for tactile Internet using a 5G gaming application [C]//14th IEEE Annual Consumer Communications & Networking Conference (CCNC). IEEE, 2017: 154–159
- [25] DU W Q, LIU Q, GAO Z G, et al. Seamless vertical handoff protocol for LTE-802.11p hybrid vehicular communications over the tactile Internet [C]//2018 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE). IEEE, 2018: 1–5. DOI: 10.1109/HAVE.2018.8548055
- [26] IYOSHI K, TAUSEEF M, GEBREMEDHIN R, et al. Towards standardization of haptic handshake for tactile Internet: a WebRTC-based implementation [C]//2019 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE). IEEE, 2019: 1–6. DOI:10.1109/HAVE.2019.8921013

作者简介



曾繁荣,福州大学电子信息工程专业在读本科生;主要研究方向为触觉通信。



房颖,福州大学物理与信息工程学院讲师;主要研究方向为触感信息与触感网。



赵铁松,福州大学物理与信息工程学院教授、福建省媒体信息智能处理与无线传输重点实验室主任、国家级青年引才计划入选者、福建省“闽江学者”特聘教授、福建省引进海外高层次人才,现任第6届中国青年科技工作者协会理事,并担任计算机学会多媒体专委会委员、电子学会青年科学家俱乐部成员及网络多媒体专委会委员、图象图形学会多媒体专委会委员、IEEE高级会员、《IET Electronics Letters》编委等;入选福建省高层次人才创新创业百人计划;主要研究方向为图像处理与计算机视觉、智能视频编码与通信、触感信息与触感网;承担多项科研项目;发表论文80余篇。

触觉通信中的信号获取与恢复关键技术

Key Techniques of Signal Acquisition and Recovery in Haptic Communication



刘恒发 /LIU Hengfa, 魏昕 /WEI Xin

(南京邮电大学, 中国 南京 210003)

(Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

摘要: 面向未来以低时延、高可靠、双向互动为目标的触觉通信,从触觉感受以及触觉信号特性出发,描述了触觉信号的采集与表征方法;针对以远程工业控制、远程医疗、沉浸式游戏等为代表的典型触觉业务、多模态业务中存在的触觉信号受损或丢失问题,分别从同模态和跨模态两个角度提出了触觉信号的恢复架构。所提出的触觉通信原型系统,能够实现远程触觉信息的精准感知和反馈,增强用户的沉浸式体验。

关键词: 触觉通信; 信号采集; 信号恢复; 人工智能

Abstract: Aiming at future haptic communication with low time delay, high reliability, and bi-directional interactions, the acquisition and representation methods of haptic signals are given based on touch sensation and characteristics. For the damage and loss of haptic signals in typical haptic and multi-modal scenarios such as remote industrial manipulation, telesurgery, immersive games, and architectures of haptic signal recovery are proposed from perspectives of both the within-modal and cross-modal, respectively. Moreover, several haptic communication prototype systems are built, which can effectively realize remote haptic perception and feedback, and promote the user's immersive experience.

Keywords: haptic communication; signal acquisition; signal recovery; artificial intelligence

DOI: 10.12142/ZTETJ.202106003

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211208.0352.002.html>

网络出版日期: 2021-12-09

收稿日期: 2021-11-12

触觉是人类感知体验并与外部环境交互的一类重要方式,它通过来自皮肤表面的敏感神经传感器的触感,来感受接触物的材质、运动、压力、温度等属性。这些属性在很多情况下是视、听觉所不能完整、精确提供的。

通过实现各种感官信息的通信,人类可以更好地感知和探索世界。传统的多媒体通信系统主要通过通过对音视频数据的捕获、编码、传输、解码等,实现音视频通信,使用户获得丰富的

视听觉体验。另一方面,近年来面向人类触觉感受而设计的触觉通信系统以及触觉相关服务,受到了学术界和工业界的关注。与传统以音视频为主的多媒体通信不同的是,触觉通信主要关注本地终端与远程终端的高保真触觉信息的传输与接收^[1]。通过丰富的感官信息增强交互,触觉通信可以提升用户的沉浸式体验。因此,触觉通信在远程医疗^[2]、线上购物、在线教育^[3]、远程工业控制等领域有着广泛的应用。

本文主要聚焦于触觉通信中的发

送端和接收端,研究发送端的触觉信号获取技术以及接收端的触觉信号恢复技术,并简要介绍课题组所开发的触觉通信原型系统。

1 触觉信号获取

1.1 触觉感受

触觉通信的最终目标是实现人类触觉感受的传递与接收,因此我们首先需要明晰何为触觉感受。当前,学术界和工业界最为关注的是触摸和动觉两种触觉感受^[4]。其中,触摸感受

基金项目: 国家自然科学基金(62071254)

来源于人体的皮肤,通过表皮下4种不同类型的感受器,向中枢神经系统和大脑的体感区域提供关于皮肤表面发生的机械事件的信息,进而获取形状、粗糙度、温度、摩擦等属性。动觉感受则来自肌肉,通过肌梭向中枢神经系统提供有关肌肉长度和肌肉长度变化速度的信息^[5],使人类能够感知肢体运动的方向、幅度和速度以及相关的肢体位置变化。因此,对于触摸感受,可以通过压力、摩擦和振动等触觉信号表征;而对于动觉感受,则可以通过力反馈装置表征为位置、速度和力等触觉信号。

1.2 触觉与视听觉信息的本质区别

与视听觉信息相比,触觉具有一些显著的区别,如表1所示。

双向交互性:与视听觉信息单向输入或输出不同,触觉感知不仅来自于被动接触(如把东西放手上),还来自于主动探索环境(如用手主动触摸环境)。这样的双向特性使得触觉通信比传统以视听为主的多媒体通信更具交互性。此外,在触摸物体的过程中,还可能会改变物体的属性,如使之发生形变、温度变化等,这种不可逆的相互作用则不会发生在视觉和听觉中。

突发性:与音视频信号相比,触觉信号需要高的采样率来确保系统的稳定性和透明性^[6]。当交互设备与外界接触时,如执行高频敲动、推动物体、低频按压和抓取释放等交互时,由于触觉包速率急剧变化,易导致触觉流量突发^[7]。此外,在实际触觉通信系统中,根据上述特点并考虑到节约传输资源,触觉信号往往采用间断性传

输,而音视频信号则需要连续传输。因此,在突发流量期间,两个传输数据包之间间隔很小,容易导致传输不可靠,如延迟或数据包丢失,最终影响用户体验。基于上述分析,触觉的突发性是影响触觉通信质量的关键因素之一。

敏感性:触觉的感知敏感性与其空间分辨率和时间分辨率密切相关。空间分辨率是指可以检测到的触觉刺激之间的空间分隔距离。例如,在指尖上,我们可以分辨出大约1 mm的间隔,这个数值位于视觉和听觉空间分辨率之间。时间分辨率是指传递到触觉感受器的两个脉冲被感知为连续不同的刺激而非同一刺激所需的时间差。例如,在触摸过程中,人类可以分辨出5 ms的时间差异,这比视觉的时间差(25 ms)好,但是比听觉的时间差(0.01 ms)差。此外,触觉分布感受器遍布全身,感知特定触觉刺激的能力在皮肤表面上变化,相同的触觉刺激将在不同的位置发生不同的感知体验。因此,敏感性也影响着触觉通信系统的设计与实现。

1.3 触觉信号的采集

触觉信号的采集是实现触觉通信的前提。与通过摄像机、照相机等设备以非接触式方式所采集的音视频信号不同,触觉信号主要通过安装在机械装置或机器人上的触觉传感器与物体接触时采集,如触摸时物体的位置、速度、力、扭矩和形变等触觉属性数据。

(1) 摩擦力和压力:当相关传感器与外界物体表面材质接触时,产生的触觉信号最常见的是接触力,例如切向摩擦力和法向压力,通常用接触

时X、Y和Z3个轴的力的幅度来表示。

(2) 加速度:在物体表面上敲击刚性工具或拍打物体表面会导致工具产生一定的加速度,这些振动可以用三轴或单轴加速度传感器测量,相应的一维信号代表材料表面触觉特性。同时在某些情况下,可以使用声学传感器(麦克风)来检测振动,用于触觉传感。

(3) 动觉信息:动觉信息主要通过记录位置变化(运动轨迹)、力、扭矩、速度等来呈现,可以通过计算机、力传感器、速度传感器等记录、采集。

与上述触觉属性相对应的触觉信号采集设备如图1所示。

此外,当探索不同的物体属性时,人类会使用不同的交互方式,这种运动称为人类的探索性程序^[10]。为了让机器也能获得感知触觉信息的能力,在触觉采集时,可以让机器进行与人类类似的探索性活动。即通过使用配备有压力、温度和加速度等传感器的机械手,采用轻拍、按压、保持、环绕和滑动等采集方式来匹配人类触摸物体时的原型动作集,获取期望的触觉属性。

(1) 轻拍:机械手迅速围绕物体闭合,待与物体有接触时便打开、释放物体。这样可初步探知材料的温度信息,模仿人类快速触摸以感知温度的行为。

(2) 按压:人类对压力的探索性活动,可以辨别物体的硬度^[10]。夹具以恒定的速度关闭,待按压深度达到特定阈值后再以相同的恒定速度打开,直到任一手指都感觉不到接触。

(3) 静态保持:人类对全局形状、体积、温度和热导率等探索性活动^[10]。当夹具与物体接触并保证物体不发生滑动后,机器人轻轻握住物体10 s,同时让加热的手指与物体达到热平衡。

(4) 环绕:适合手的物体大小和

▼表1 触觉与视听信息区别

	双向特性	突发性	空间分辨率	时间分辨率
视觉	×	×	高	低
听觉	×	×	低	高
触觉	√	√	中	中



▲ 图1 不同类型的采集设备

形状,可以根据皮肤压痕和手指的姿势来感知,因为手可以包围物体。但当物体较大时,必须用手依次探索以确定形状。依据触觉探索的这种性质,通过环绕多次接触物体,才能获得全局形状。

(5) 横向运动:通过让机械手在轻握住物体的同时向下滑动,来模拟人类的横向运动(手指在表面来回移动)。其中,包括速度为 1 cm/s 的缓慢移动和速度为 2.5 cm/s 的快速移动,快速移动的接触强度弱于缓慢移动^[9]。

1.4 触觉信号的表征

对于采集到的触觉信号,我们需要以一定的形式对其进行表征。触觉信号的表征是对真实表面的感觉进行描述和建模,之后才能进行编码、传输等操作。触觉信号的表征形式按信号类型可以分为基于一维时间序列的表征和基于二维图像数据的表征,如图2所示。

(1) 基于一维时间序列的表征。物体材质表面的触觉信息(包括材质的纹理轮廓、摩擦力、振动信号等)都可以基于上述采集设备与目标物体

交互时产生的接触力和振动、速度、加速度和三轴方向的位置等,以一维信号(如波形)形式表达。基于真实的触觉交互数据便于构建表面纹理细节,但存在触觉数据规模较小、触觉信号处理方法尚未成熟等问题。

(2) 基于二维图像数据的表征。触觉数据需要经过专门的采集设备来收集,难度较大,且现存的触觉数据集规模较小。与之相比,图像数据集包含丰富的视觉和触觉特征,且数量庞大。基于图像特征提取的方法可以从图片的灰度、深度等信息中获取纹理表面的特征,从而实现几何信息到触觉信息的映射。另外,当采集设备与物体进行接触时,物体表面会发生一定程度的形变。利用附在传感器表面的薄膜记录下这种形变,并通过图

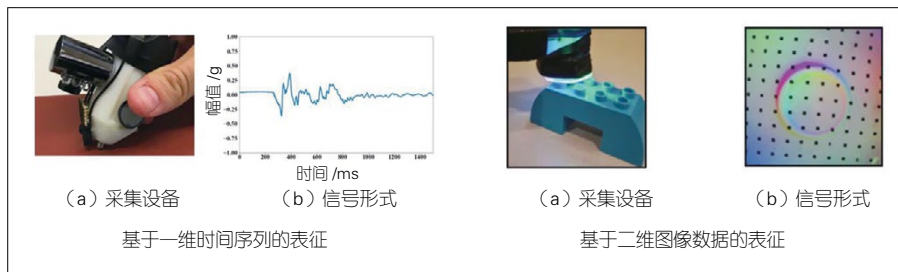
像的形式呈现可以表征触觉信息。此外,图像频域变换所提取出的图像频谱特征也能反映物体表面的纹理分布。基于二维图像数据表征的方法可以充分利用现在的大规模图像数据集,但其缺点在于无法表征微观特性。

2 触觉信号的恢复

虽然当前全球学者提出的高效触觉编码方案^[11]、自适应传输速率控制方案^[7]等可以在很大程度上保障触觉传输的可靠性。但是通过实际测试发现,当触觉信号通过无线网络传输时仍会不可避免地遭遇各种因素的干扰(如可用传输资源的动态变化、不可预测的通信延迟等)。这会导致触觉信号在接收端出现不同程度的损耗甚至缺失。此外,触觉信号具有突发性和不可预测性,从而容易使接收端收到时序错乱的触觉信号^[12]。因此,在上述情况下,需要采取相关技术来实现触觉信号的恢复,从而保障终端用户的触觉感知质量。具体而言,我们认为可以采用两类方法来实现触觉信号的恢复。

2.1 同模态触觉信号恢复

在以远程工业控制为代表的触觉业务中,接收端触觉信号的时序特性以及上下文数据间的相关性可以作为受损信号重建的依据。本文中,我们使用深度神经网络来学习触觉信号的时序特性以及相关性,建立一个缺失触觉信号到完整触觉信号的映射,从



▲ 图2 触觉信号表征形式

而实现缺失触觉信号的重建。近年来,生成对抗网络依靠其学习数据之间复杂分布的能力,在图像修复及超分辨率重建、电力系统缺失数据重建^[13]等应用场景中都取得了巨大成功。其核心思想在于利用不完整数据和给定的上下文约束去生成符合客观规律的缺失部分。基于此,我们使用生成对抗网络来完成缺失触觉信号的恢复。此外,由于传统的生成对抗网络使用随机向量来生成期望的信号,无法充分利用已有的受损信号,因此我们考虑利用条件生成对抗网络(GAN)以充分利用已有的部分受损信号,并将其作为条件输入到生成网络和鉴别网络中去,从而提高恢复信号的准确性。所提出的同模态的触觉信号恢复框架如图3所示。

首先,将受损的触觉信号和隐变量 z 输入生成器中,此时受损的触觉信号提供一个条件约束,以提高所生成的触觉信号的上下文一致性。然后,把生成的触觉信号输入判别器,以鉴别生成信号的真实性的。最后,使用真实性约束和上下文约束来优化生成器参数和输入的隐变量 z 。需要说明的是,真实性约束的目的是使生成器生成的触觉信号尽可能接近真实的触觉信号,

以使鉴别器不能区分两者,从而保证了通过生成网络而恢复出的触觉信号质量。上下文约束的目的是从所生成的触觉信号中去搜寻与受损触觉信号中残留片段最相似的样本来优化隐变量,从而保证生成的触觉信号片段与残留触觉信号片段具有上下文一致性。此外,当触觉信号中存在标签信息时(如触摸不同类别的材质),还可以在图3框架中增加一个语义标签,以达到进一步提升触觉恢复质量的目的。

2.2 跨模态触觉信号恢复

在以远程针灸、沉浸式游戏等为代表的多模态业务中,接收端除了收到受损触觉信号以外,还可以接收到视频信号。不同于触觉业务中只能利用同模态的触觉信号恢复,多模态业务可以利用与受损触觉信号具有相同语义的音视频信号来辅助完成触觉信号恢复,即跨模态触觉信号恢复。

在跨模态触觉信号恢复中,最大的挑战在于不同模态的信号(音频、视频、触觉信号)是高度异构的,且维度差异较大,因此我们需要探究其内在相关性,从而减少模态间的语义偏差。此外,在支撑多模态业务的实际通信场景中,可用作训练集的音视

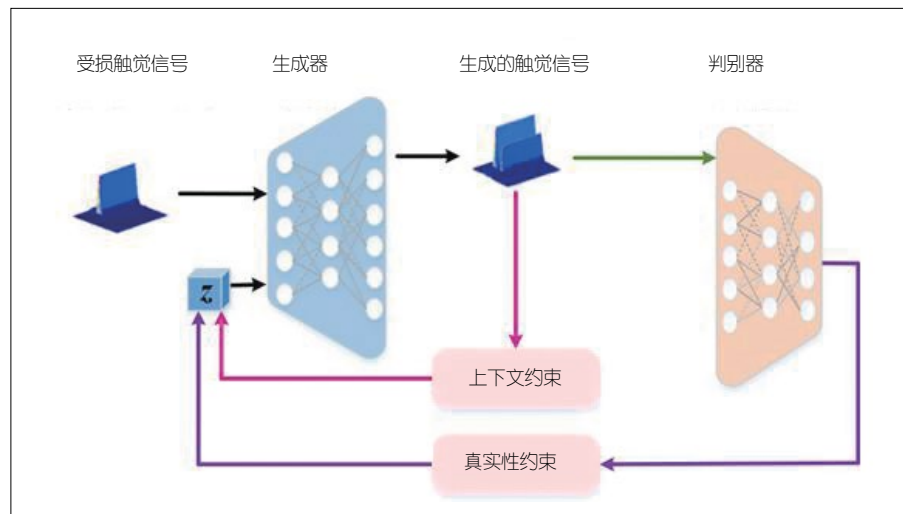
频信号数量通常十分匮乏。这些都是实现触觉恢复所亟待解决的问题。

对于不同模态信号间的异构问题,我们认为对于同一对象,不同模态信号所表达的语义具有潜在的关联性和一致性。因此,不同模态的深层语义可以互相提供丰富的辅助信息。通过建立公共语义空间,能够探索不同模态信号间的语义一致性,从而实现模态特征的对齐、融合与互补。针对实际通信系统中用于训练的音视频信号数量匮乏问题,我们可以从其他海量的公共音视频数据集中提取有用知识并实现知识的迁移来解决。

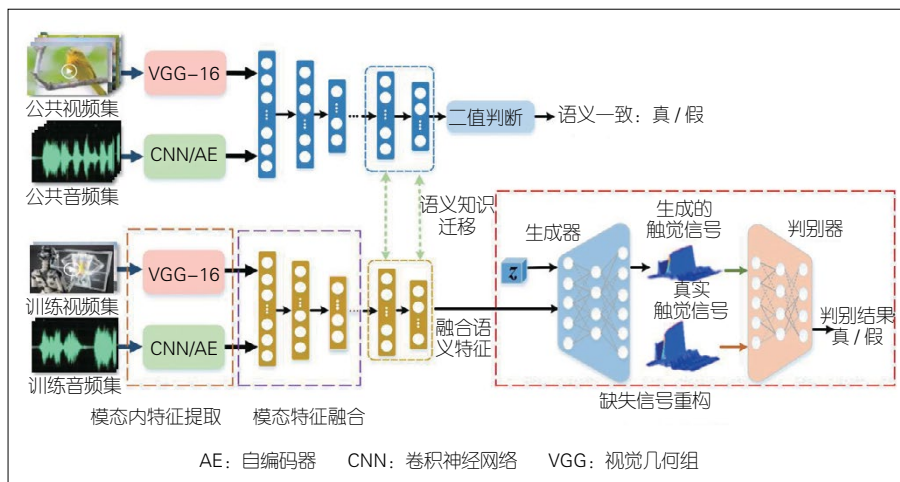
鉴于上述分析,我们提出了跨模态触觉信号恢复框架^[14],具体如图4所示。首先,基于自监督学习的思想,从海量音视频数据中汲取并迁移语义知识;接着,通过搭建融合网络,实现对所提取的语义信息的融合建模,从而弥补模态间的语义鸿沟;最后,利用融合的语义信息,并通过生成对抗网络,实现缺失触觉信号的重建。

(1) 基于跨模态知识迁移的语义特征提取

我们将接收到的音频、视频数据作为训练集合,并从中提取与模态无关的抽象语义信息。对于视频数据,可以采用在大规模公共数据集(如ImageNet)上预训练的视觉几何组(VGG 16)或RESNET网络进行视觉特征的提取;对于音频数据,基于语谱图,可以将其输入至卷积神经网络或编码器中以提取特征。值得注意的是,在利用大规模公共数据集时,由于音视频数据没有显式的语义标签,因此我们可以基于自监督学习技术来设计一个分类任务,用于判断视频和音频是否属于同一场景或内容。通过该方式学习可以得到语义知识。接着,将语义知识传递给基于当前接收到的音视触数据所建立的触觉恢复模型。



▲ 图3 同模态触觉信号恢复框架



▲ 图4 跨模态触觉信号恢复框架

基于跨数据集的知识迁移有效地解决了实际触觉通信系统中用于训练的数据稀少的问题，并且显著增强触觉恢复模型中的语义特征提取效果。

(2) 基于共享语义子空间的模态特征融合

经过上述特征提取步骤，我们可以分别获得视觉、听觉模态的语义特征，但单个模态通常无法包含实现精确触觉信号重构所需的全部有效信息。因此，我们需要进行模态特征的融合，以实现有用信息的相互补偿。首先，我们需要实现视觉特征和听觉特征的对齐，这可以基于标签信息来实现；

接着，通过搭建公共语义空间，将对齐的视觉、听觉特征输入至由多个全连接层组成的神经网络，从而实现模态特征的融合。

(3) 基于对抗学习的缺失触觉信号恢复

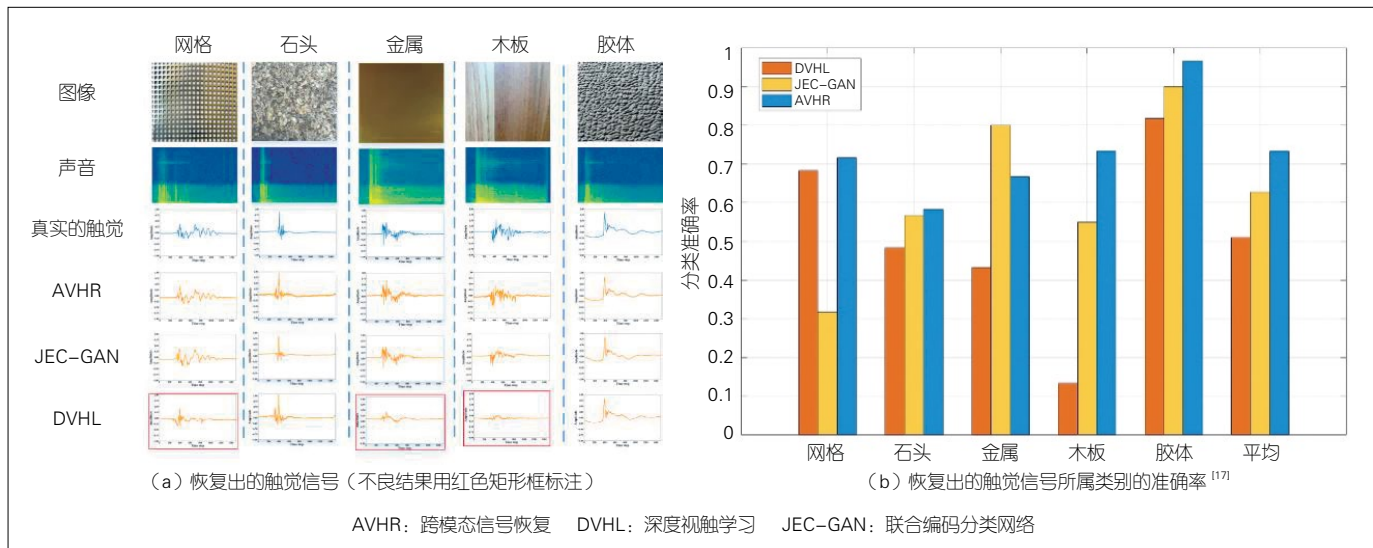
基于模态融合后的语义，我们建立触觉模态生成模型。与同模态触觉信号恢复框架类似，我们使用生成对抗网络来实现缺失触觉信号的生成。具体而言，利用生成对抗网络在学习数据分布和特征表示上的能力，我们将从语义空间获得的融合特征和随机向量输入至生成网络；随后，把生成

的触觉信号与真实的触觉信号输入至鉴别器；通过生成器和鉴别器的对抗训练，不断地迭代优化降低损失直至最优。

为了验证上述跨模态信号恢复（AVHR）方案，我们将AVHR在触觉纹理（LMT）数据集^[18]上进行实验，并与另外两种触觉信号恢复方法进行了比较。我们采用了联合编码分类网络（JEC-GAN）^[15]和深度视触学习（DVHL）^[16]的对比方法。其中，JEC-GAN将每个模态的实例编码到共享的固有空间，然后利用成对约束使视触样本在固有空间接近，最后将视觉信息作为输入，重构相应的触觉信号；DVHL方法扩展了具有潜在变量的传统自编码器并学习其深度相关性。图5分别展示了不同方法重建出的触觉信号以及所属类别的准确率，结果表明本文所提的方法能够重建出更接近真实的高质量触觉信号。

2.3 触觉信号的呈现

在成功接收到触觉信号或完成触觉恢复后，我们需要对触觉信号进行呈现，即利用特定的硬件装置来处理接收到的信号，然后模拟产生与物体



▲ 图5 不同恢复方案的性能比较

接触时的触感,使用户感受到物体的弹性、刚度、摩擦等物理属性和形状纹理等几何属性,从而提升用户的交互式体验。

典型的触觉呈现硬件装置分为触摸屏、触觉笔、可穿戴系统、空气触觉系统、动觉呈现设备等,如图6所示。

(1) 触摸屏:基于触摸屏的表面触觉信息,通过改变触觉面板与指尖之间的摩擦力来呈现触觉。

(2) 触觉笔:常用于手写学习系统或触觉游戏中,可分为基于静电力振动的触觉笔和基于振动电机和压力传感器的触觉笔。

(3) 可穿戴设备:可以依附于身体的各个部位而不受工作空间的限制,例如,触觉手套、触觉机械臂、触觉背心等。通过向特定皮肤提供触觉反馈,可以大大提高交互性。例如,Immersion公司的触觉反馈手套 Cyber Touch,在手掌与手指部位设置多个触觉振动器,通过改变各个振动器的振动强度产生不同的触感压力,让用户感受到物体的形状。

(4) 空气触觉系统:通过气压冲击、超声波等方式传递能量,使用户无须直接接触便能产生触觉感知。例如,触觉反馈装置 HaptiRead 能通过精确模式的超声波脉冲,让视力受损的人感受到再现于半空中的盲文。

3 触觉通信原型系统

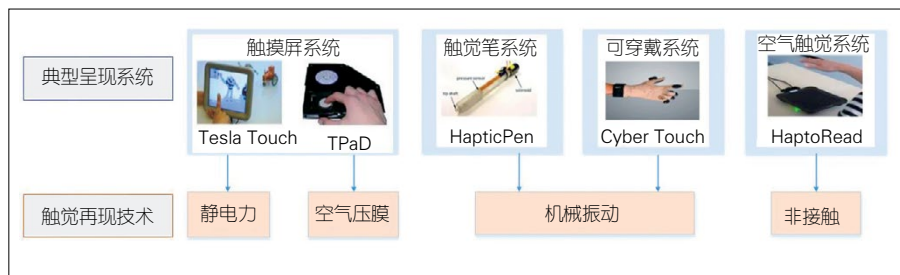
基于现有触觉通信技术,我们搭建了基于触觉感知的医疗诊断模拟系统以及基于机械手臂的远程触摸控制原型系统。

图7为我们开发的基于触觉感知的医疗诊断模拟系统,用于辅助外科实习医生进入临床环境之前的训练和教学。如图7左图所示,我们首先对人体皮肤和伤口进行3D建模,并给模型附着对应的真实触觉感应,然后与

触觉力反馈设备 Geomagic Touch X 进行联合编程调试,最后实现了皮肤缝合诊断模拟。在图7右图中,通过利用真实口腔的3D模型和收集到的相关触觉感知信息,能够在接收端感知到口腔不同部位的真实触觉感受,比如牙龈的光滑柔软感和牙齿的硬物感。未来,我们的目标是将系统应用于远程康复、针灸中。这不仅包含触觉信息,也包括器官表面及内部的图像纹理等视听觉信息,以实现音视触信息的协同传输与综合处理,增强用户在操作端的真实感和沉浸体验。

图8是我们开发的远程表面材质触摸控制原型系统。该系统能够应用于远程工业控制中需要感知物体材质表面触觉的场景和任务。该原型系统采用触觉力反馈设备 Geomagic Touch

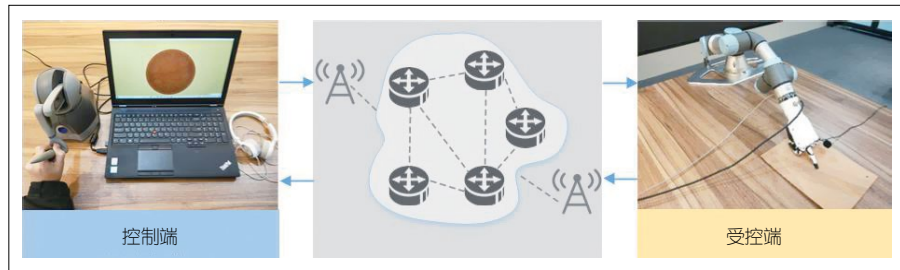
作为控制端,优傲机械臂 (UR3) 作为受控端,并在手指处装有 Teckscan 薄膜触觉压力传感器,来感知不同的物体属性。通过机械手在材料表面触摸和滑动,Teckscan 薄膜触觉压力传感器收集机械手指尖按压材料得到的压力信号,并将该压力数据作为触觉信号。其中,压力传感器的采样频率约 200 Hz。同时,摄像机和麦克风能采集到对应的视频和音频信号。控制端可以发送控制命令,以控制机械手臂、接收触觉反馈。远程机械手臂端可以执行相应的指令,并提供触觉信号以及对应的音视频信号。目前,该系统已经成功实现了木材、纸板、丝绸、泡沫等9类材质表面触觉的采集、传输、识别,识别准确率高达90%。此外,当发生触觉信号的受损、缺失时,



▲ 图6 不同的触觉再现技术



▲ 图7 基于触觉感知的诊断模拟系统



▲ 图8 远程表面材质触摸控制原型系统

依据上述跨模态触觉信号恢复算法,我们能够利用接收的音、视频信号进行相应的触觉信号恢复。

4 结束语

本文中,聚焦于触觉通信中的发送端信号获取和接收端触觉恢复技术,我们首先对触觉感受、触觉信号的采集设备以及触觉信号的表征进行了描述,接着针对触觉通信过程中可能存在的信号缺失、干扰等问题,面向触觉业务和多模态业务场景,分别提出了同模态和跨模态的触觉信号恢复技术,最后简要介绍了目前所开发的一些触觉通信原型系统。对于未来触觉通信的研究,除了探索更有效的触觉编码方案、更可靠的传输机制、更灵活的资源管理方案^[19]等外,面对用户对视听触感沉浸式综合体验的更高需求,建立面向音频、视频、触觉信号协同传输与处理的跨模态通信策略^[20]将成为学术界和工业界研究和关注的热点。

参考文献

- [1] QIAO Y, ZHENG Q, LIN Y, et al. Haptic communication: toward 5G tactile Internet [C]//2020 Cross Strait Radio Science & Wireless Technology Conference (CSRSWTC). IEEE, 2020: 1–3. DOI: 10.1109/CSRSWTC50769.2020.9372659
- [2] GUPTA R, TANWAR S, TYAGI S, et al. Tactile-Internet-based telesurgery system for healthcare 4.0: an architecture, research challenges, and future directions [J]. IEEE network, 2019, 33(6): 22–29. DOI: 10.1109/mnet.001.1900063
- [3] WEI X, DUAN Q, ZHOU L. A QoE-driven tactile Internet architecture for smart city [J]. IEEE network, 2020, 34(1): 130–136. DOI: 10.1109/MNET.001.1900078
- [4] CULBERTSON H, SCHORR S B, OKAMURA A M. Haptics: the present and future of artificial touch sensation [J]. Annual review of control, robotics, and autonomous systems, 2018, 1(1): 385–409. DOI: 10.1146/annurev-control-060117-105043
- [5] JONES L A, LEDERMAN S J. Human hand function [M]. USA: Oxford University Press, 2006
- [6] LAWRENCE D A. Stability and transparency in bilateral teleoperation [J]. IEEE transactions on robotics and automation, 1993, 9(5): 624–637. DOI: 10.1109/70.258054
- [7] GUI M, XU X, STEINBACH E. Adaptive packet rate control for the mitigation of bursty haptic traffic in teleoperation systems [C]//2020 IEEE Haptics Symposium (HAPTICS). IEEE, 2020: 134–139. DOI: 10.1109/HAPTICS45997.2020.ras.HAP20.18.6a51e1e1
- [8] STRESE M, SCHUWERK C, IEPURE A, et al. Multimodal feature-based surface material classification [J]. IEEE transactions on haptics, 2017, 10(2): 226–239. DOI: 10.1109/TOH.2016.2625787
- [9] CHU V, MCMAHON I, RIANO L, et al. Robotic learning of haptic adjectives through physical interaction [J]. Robotics and autonomous systems, 2015, 63: 279–292. DOI: 10.1016/j.robot.2014.09.021
- [10] LEDERMAN S J, KLATZKY R L. Extracting object properties through haptic exploration [J]. Acta psychologica, 1993, 84(1): 29–40. DOI: 10.1016/0001-6918(93)90070-8
- [11] STEINBACH E, STRESE M, EID M, et al. Haptic codecs for the tactile Internet [J]. Proceedings of the IEEE, 2019, 107(2): 447–470. DOI: 10.1109/JPROC.2018.2867835
- [12] GOKHALE V, NAIR J, CHAUDHURI S. Congestion control for network-aware telehaptic communication [J]. ACM transactions on multimedia computing, communications, and applications, 2017, 13(2): 1–26. DOI: 10.1145/3052821
- [13] 王守相, 陈海文, 潘志新, 等. 采用改进生成式对抗网络的电力系统量测缺失数据重建方法 [J]. 中国电机工程学报, 2019, 39(1): 56–64
- [14] WEI X, ZHOU L. AI-enabled cross-modal communications [J]. IEEE wireless communications, 2021, 28(4): 182–189. DOI: 10.1109/MWC.001.2000448
- [15] PURRI M, DANA K. Teaching cameras to feel: estimating tactile physical properties of surfaces from images [EB/OL]. (2020-04-29)[2021-11-07]. <https://arxiv.org/abs/2004.14487>
- [16] TAKAHASHI K, TAN J. Deep visuo-tactile learning: estimation of tactile properties from images [C]//2019 International Conference on Robotics and Automation (ICRA). IEEE: 8951–8957, 2019. DOI: 10.1109/ICRA.2019.8794285
- [17] WEI X, SHI Y, ZHOU L. Haptic signal reconstruction for cross-modal communications [EB/OL]. [2021-11-10]. <https://ieeexplore.ieee.org/document/9577216>. DOI: 10.1109/TMM.2021.3119860, 2021
- [18] STRESE M, SCHUWEK C, IEPURE A, et al. Multimodal feature-based surface material classification [J]. IEEE transactions on haptics, 2017, 10(2): 226–239. DOI: 10.1109/TOH.2016.2625787
- [19] YUAN Z, WEI X, CHEN J, et al. Ultra-reliability connectivity with redundant D2D transmission scheme for tactile Internet [C]//2019 IEEE Globecom Workshops (GC Wkshps). IEEE, 2019: 1–6. DOI: 10.1109/GCWkshps45667.2019.9024454
- [20] ZHOU L, WU D, CHEN J, et al. Cross-modal collaborative communications [J]. IEEE wireless communications, 2020, 27(2): 112–117. DOI: 10.1109/MWC.001.1900201

作者简介



刘恒发, 南京邮电大学在读博士研究生; 主要研究方向为多媒体大数据分析与管理。



魏昕, 南京邮电大学教授; 主要研究领域为多媒体通信、多媒体大数据分析与管理; 先后主持和参加各类科研项目20余项; 发表论文70余篇, 出版英文学术专著1部。

面向 6G 的触觉互联网关键技术挑战

Key Technologies and Challenges of 6G-Oriented Tactile Internet

尤子硕 /YOU Zishuo, 李强 /LI Qiang, 唐冰 /TANG Bing, 葛晓虎 /GE Xiaohu

(华中科技大学, 中国 武汉 430074)

(Huazhong University of Science and Technology, Wuhan 430074, China)



摘要: 为了顺应 6G 网络发展趋势, 围绕触觉互联网 (TI) 超低时延、超高可靠的关键性能需求, 从网络架构、信息传输、资源管理 3 个层面系统性地提出了 TI 关键使能技术框架。在网络架构方面, 提出基于软件定义网络 (SDN) / 网络功能虚拟化 (NFV) 和人工智能 (AI) 赋能的新型网络架构; 在信息传输方面, 提出采用多模态通信增强沉浸体验, 并利用语义通信促进智能体间的达意通信; 在资源管理方面, 提出通过通信 - 计算 - 存储多维资源融合调度来提高网络服务承载能力。最后, 探讨了 TI 在发展和应用中存在的问题、挑战以及可能的解决思路。

关键词: 触觉互联网; 网络架构; 多模态通信; 语义通信; 通信 - 计算 - 存储融合调度

Abstract: Following the development trend of 6G networks and focusing on the key performance requirements of the tactile Internet (TI) for ultra-low latency and ultra-high reliability, TI's key enabling technology framework is systematically proposed from three levels including network architecture, information transmission, and resource management. In terms of network architecture, a new network architecture based on software defined-network (SDN)/network function virtualization (NFV) and artificial intelligence (AI) is proposed. In terms of information transmission, it is proposed to use multi-modal communication to enhance the immersive experience, and to use semantic communication to promote idea-passing communication between agents. In terms of resource management, it is proposed to improve the carrying capacity of network services through the integration of communication-computing-caching multi-dimensional resource scheduling. Finally, the problems, challenges, and possible solutions in TI's development and application are discussed.

Keywords: tactile Internet; network architecture; multimodal communication; semantic communication; communication-computing-caching fusion scheduling

DOI: 10.12142/ZTETJ.202106004

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211208.0943.006.html>

网络出版日期: 2021-12-09

收稿日期: 2021-11-10

随着智能终端设备的普及与移动高清视频、虚拟现实 (VR) / 增强现实 (AR) 等业务的不断涌现, 全球移动业务数据流量呈爆炸式增长, 预

计 10 年后将达到 5G 网络的极限。与此同时, 5G 网络在数据传输速率、覆盖范围、接入密度等方面面临瓶颈, 无法充分满足车联网、工业互联网、远程医疗、远程教育等新兴业务的发展需求。为了进一步打破现实世界和虚拟世界的界限, 缩小现实世界与数字世界的鸿沟, 全球主要国家和经济

体已启动 6G 的研究探索。6G 将极大满足人与人、人与物、人与环境之间的沟通交互需求, 提供不受时间、空间限制的完全沉浸化交互体验^[1]。这将有助于构建一个平行于现实世界的开放数字虚拟世界, 促进人类社会的数字化转型^[2]。

6G 时代新技术的应用, 使得实时

基金项目: 国家自然科学基金 (61971461); 湖北省重点研发项目 (2020BAA002); 湖北省重点研发项目 (2021BAA015)

传输触觉控制信息的触觉互联网 (TI) 成为可能。区别于传统网络信息传输, TI 通过对触觉控制信息的传输和反馈来进行精细的动作控制, 通过与实际或虚拟的物体进行远程实时交互来实现“技能传输”。英国咨询公司 IDTechEx 的报告显示: 触觉反馈技术市场正呈指数级增长, 预计到 2025 年市场规模将接近 50 亿美元。此外, TI 技术将使 VR/AR 更具有沉浸感, 感官互联预期将成为未来主流的通信方式。这将为移动通信、制造业、医疗健康等新兴行业的发展带来更多、更大的市场机遇。

目前, TI 作为 B5G/6G 的关键应用, 已经引起学术界和工业界的广泛关注。国际电信联盟 (ITU) 和电气与电子工程师协会 (IEEE) 等国际标准化组织已经开始 TI 相关领域的标准制定^[3-4]。文献 [5] 指出借助 TI 可以实时控制真实或虚拟对象, 从而构建新一代信息物理系统。文献 [6] 比较了 5G、物联网、TI 3 种网络, 指出 TI 区别于传统网络的关键在于利用触觉设备实现人-机 (H2M) 通信。文献 [7] 提出了 TI 的一些关键性能要求, 包括: 极低时延极高可靠性、安全性、多感官的整合。文献 [8] 提出了一些 TI 潜在应用场景, 包括医疗健康领域的外骨骼、远程教育中的交互式虚拟现实体验、自动驾驶中的远程汽车控制, 以及智能制造中机器人参与大规模装配生产线等。文献 [9] 指出实现 TI 面临的挑战, 包括设计能够提供触觉和触觉信号反馈并且能与蜂窝网络自由通信的触觉设备, 开发实现触觉数据压缩的触觉编码器, 融合具有不同采样/传输速率和延迟需求的视觉、听觉、触觉多模态信息等。文献 [10] 提出, TI 需要在共享的基础设施上实现实时控制与网络边缘高效计算能力的结合。文献 [11] 提出一种基于多层云

的蜂窝系统以支持 TI 应用, 该系统包含微云单元、负责管理区域内的微云单元并形成边缘计算架构的小型云、负责管理网络内所有小型云的核心网络云以及作为大型数据中心的远程公共云。文献 [12] 提出一种面向软件定义网络 (SDN)/网络功能虚拟化 (NFV) 的多层云 TI 结构, 有助于降低时延, 提高系统性能。文献 [13] 提出一种基于用户体验质量 (QoE) 的感知模型, 可用于工业物联网中触觉应用的动态资源分配。

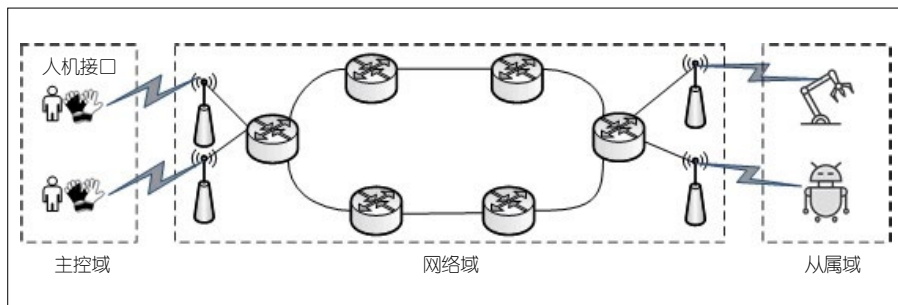
如上所述, 现有研究工作大多从应用场景、网络架构、编/译码器等单一角度研究 TI。然而, TI 作为一个系统工程, 需要从不同层面开展关键使能技术的融合创新。基于网络架构、信息传输以及资源管理这 3 个不同层面, 本文提出了 TI 关键使能技术框架。为了满足触觉控制和反馈信息的实时、可靠传输, 需要设计基于 SDN/NFV 和人工智能 (AI) 赋能的新型网络架构, 通过提供灵活的网络功能和基于本地的服务保障时延、可靠性等关键性能要求; 为了促进多感官信息融合, 需要利用多模态通信技术实现多感官信息的传输、接收和重建, 并借助语义通信技术融合提取多模态信息的语义特征, 减少数据冗余, 提高通信有效性; 为了灵活匹配网络资源, 需要设计通信、计算、存储等多维网络资源融合调度机制, 在提高用户沉浸式体验的同时提高网络资源利用率。

1 TI 网络范式

下一代移动网络 (NGMN) 联盟将 TI 定义为: “通过互联网远程提供对真实或虚拟对象以及物理触觉体验的实时控制的能力”^[14]。为了在沉浸式远程教育、高精度远程医疗诊断、工业控制和自动化等新兴触觉应用领域提供对真实或虚拟对象实时的控制交互能力, TI 应具备以下的网络范式。

如图 1 所示, TI 网络范式主要包括 3 个部分: 主控域、网络域、从属域^[15-16]。主控域主要包含操作者和操作系统接口, 操作系统接口或人机接口通常是一个触觉设备, 可以通过各种编码方案将操作者的输入转换为触觉输入信息并将其发送到网络域。从属域由远程操作设备构成, 受主控域中各种命令信号直接控制, 可以和远程环境中的各种对象进行交互, 同时将触觉反馈信息返回给主控域。网络域为主控域和从属域之间的双向控制通信提供媒介, 可以将操作者动态地耦合到远程环境。理想情况下, TI 网络可以让操作者拥有远程环境中沉浸式的体验感受, 从而实现物理世界的数字化重现。

视觉、听觉、触觉等不同感官信号的传输需要保持实时性、可靠性、一致性和协调性^[9]。为了更好地支持这些关键性能需求, 需要为 TI 设计灵活、智能的网络架构以提供超低时延、超高可靠和稳定的网络连接, 为 TI 设计高效的多模态信息传输技术以促进



▲ 图 1 触觉互联网网络范式

多感官信息的融合,为 TI 设计多维网络资源融合调度策略以在充分保证业务性能要求的基础上提高网络资源利用率。以下我们将从网络架构、信息传输、资源管理这 3 个方面探讨 6G 背景下新型 TI 的关键使能技术。

2 TI 关键使能技术

2.1 基于 SDN/NFV 的新型 TI 架构

传统的网络采用逻辑控制与数据传输紧密耦合的方式,网络节点同时进行控制决策和数据传输,无法满足 TI 超低时延、超高可靠的关键性能需求。实现触觉通信的关键在于设计一个灵活、高效、能提供一致性 QoE 保障的网络架构,实现网络的统一管理与资源的动态调度,使网络资源可以基于物理基础设施在不同的服务和应用之间有效切分、共享,从而降低数据传输时延,提高可靠性。

如图 2 所示,本文提出一种基于 SDN/NFV 和 AI 赋能的 TI 3 层架构。感知层通过感知设备收集环境和用户数据,然后将其上传至网络层,同时接收来自网络层的控制信号。网络层通过 SDN 和 NFV 实现逻辑控制与数据传输分离、虚拟资源与硬件资源分离,促进了对虚拟网络资源的统一管理和动态配置,提高了 TI 架构的灵活性。控制平面通过在网络的不同层级(如云计算中心和边缘计算平台等)构建网络控制器,如 SDN 控制器、域控制器、边缘控制器等,可以构建集中和分布协同的多层级网络(如云边融合网络),有效弥补网络空隙,提高网络弹性。一方面,集中式架构可以提供稳定的网络连接,从而满足确定性组网需求和服务质量保证;另一方面,分布式架构可以实现以用户为中心的控制和管理,灵活适应新业务新场景的需求。数据平面根据控制平面下达的指令,

利用移动基础设施提供连接和计算服务,实现高效的信息传输与业务交付。应用层基于分布式与集中式融合的云边协同网络,为 TI 用户提供差异化、多维度、确定性服务。根据不同的应用需求,应用层可以通过嵌入不同的 AI 引擎实现 TI 网络的智能部署、智能运维、智能优化、智能管理,从容应对各种实时的变化,实现网络的自学习、自运行、自维护。在此基础上,通过融合内生智能技术,并充分利用网络通信、感知、计算、存储等能力,应用层能够更好地承载大规模智能分布式协同服务,实现从传统的云 AI 向网络 AI 转变。

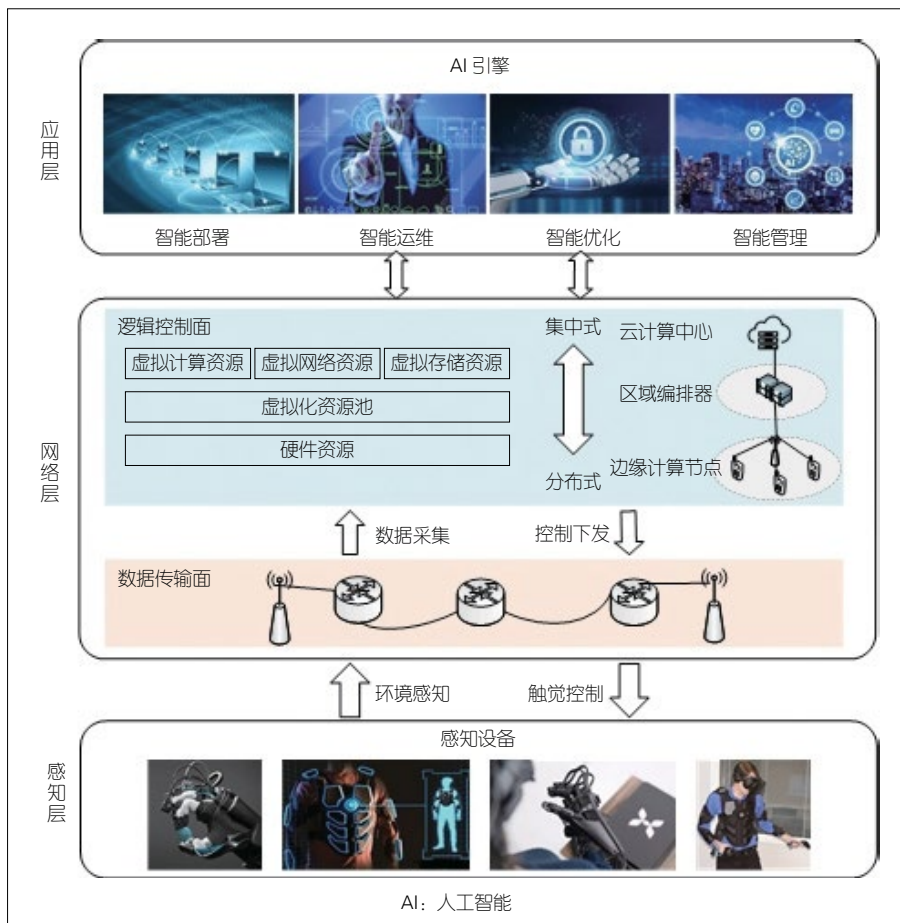
如上所述, TI 3 层网络架构能够有效结合云、网、边、端的算力资源和能力,对不同业务进行智能分发处

理。例如,将需要大量计算资源的 TI 场景渲染放在云计算中心集中处理,将多模态信号的编解码放在边缘的分布式节点进行处理,实现集散共存、分布自治的新型网络架构^[1]。

2.2 多模态语义通信

多模态通信是一种通过文字、图片、音频、视频等多种信息载体进行信息交互的方式。在面向 6G 的 TI 中,除了音、视频等传统的信息模态,还需要进行触觉信息的传输。因此,为了提供沉浸式体验,必须结合 TI 应用特点,设计相应的多模态通信系统,实现视觉、听觉、触觉等多感官信息的高效传输与融合^[17]。

多模态通信过程包含多模态信号编码、传输以及恢复重建 3 部分^[18],



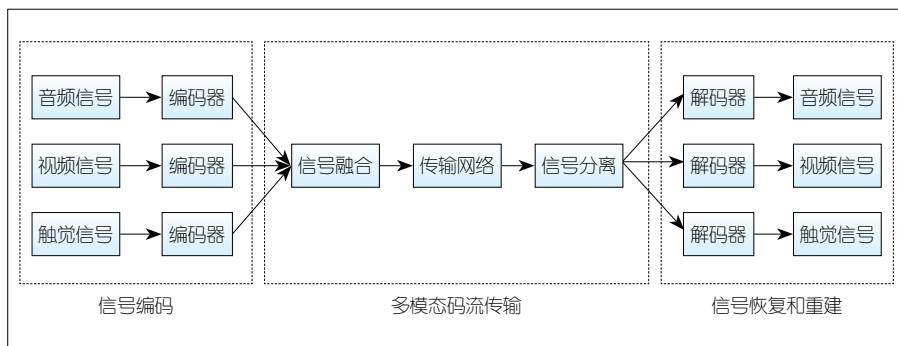
▲ 图 2 基于软件定义网络 / 网络功能虚拟化和人工智能赋能的触觉互联网架构

如图 3 所示。同一对象的多模态信号之间通常具有一定的相关性,在传输前需要进行跨模态编码^[17]。在跨模态编码中,如何对多模态信号特征进行提取压缩,从而降低冗余、提高传输效率是亟待解决的问题。

传统的香农信息论主要关注信息的形式,而用户更为关注的是信息的内容和价值,由此引出了语义信息的概念^[19]。语义通信泛指不同的智能体之间进行的以“达意”为目的通信,这里的“智能体”可以是人类或机器^[20]。鉴于 TI 应用中存在广泛的智能体之间的会意交互应用,语义通信可被用于传输相应的触觉控制指令,这将有助于实现超低时延、超高可靠的关键性能要求。提取、融合、传输多模态联合信息的语义特征为提高 TI 传输效率、满足用户沉浸式体验需求提供了一种新思路。

语义通信的关键在于设计实现语义特征提取与恢复的语义编译码器。因为 TI 应用中通常需要传输多模态信号,所以对多模态信息的表征学习是 TI 中实现语义通信的关键技术之一。目前常用的多模态信息表征学习方法有联合表示、编解码器架构、协同表示等^[21]。联合表示旨在将单模态的表示映射到一个共享语义子空间中,以融合多模态特征^[22];编解码器架构早期用于自然语言处理中的机器翻译^[23]等任务,现广泛应用在将一种模态映射到另一模态的跨模态任务中;协同表示将多模态的信息分别映射到各自的表示空间中学习,但映射后的向量之间满足一定的相关约束性^[24]。由于不同模态中包含的信息是不平等的,学习分离表征有助于保持独有的有用模态特定特征。采用高效的多模态信息表征方法能够更好地解决数据冗余问题。

在对多模态信号进行语义特征提



▲ 图 3 多模态通信过程

取和编码的基础上,还需要满足一定的传输时延和带宽要求。不同模态信号的通信带宽、可靠性等需求不同,因此在多模态码流传输时需要设计合适的传输策略^[18]。借助边缘计算和云边融合架构,可以将网络核心的计算处理功能下沉到网络边缘;通过在网络边缘对多模态信号进行处理和传输,能够有效缩短与用户的距离,进一步提高多模态通信的效率。

多模态码流经过传输和解码后会有一定程度的损耗,因此对接收到的多模态信息进行高效恢复重建也是影响 TI 用户体验的关键技术。现存方法多基于深度学习模型,例如 WEI X.^[25]等使用迁移学习和生成对抗网络的方法从不同的模态中提取、迁移和融合信息,恢复损坏的信号,提升接收器的渲染效果。

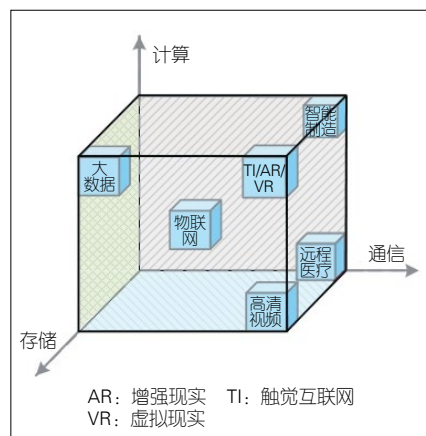
2.3 通信、计算、存储资源联合调度

传统网络以连接为中心,旨在通过匹配相应的通信资源(如带宽等)实现端到端的可靠连接。新型业态的不断涌现,如自然语言处理、VR/AR、人脸识别等应用稳步发展,不仅要求高通信数据速率,还要求通信系统提供足够强大的存储和计算资源。通信不再是网络提供的唯一核心业务,通信能力也不再等同于网络的服务能力。为了提供全面沉浸化的 TI 服务,往往需要融合调用通信、计算、存储

等网络多维资源,如图 4 所示。

因此,为了满足用户“身临其境”的沉浸式体验需求,同时避免由于网络单一类型资源匮乏导致的 TI 业务服务质量的不确定性,需要为 TI 设计网络多维资源的融合调度策略。具体来说,为了实现高速率、低时延的传输触觉控制和反馈信息,需要为 TI 匹配相应的通信资源;为了进行多模态信号的编译码、语义特征的提取与融合重建,需要为 TI 匹配相应的计算资源;为了提高计算和通信效率、存储语义库以及多模态信息的编译码映射信息,需要为 TI 匹配相应的存储资源。值得注意的是,多维资源之间存在折衷互补特性。针对由于某种类型资源匮乏导致业务需求无法满足的情况,可以通过通信、计算、存储资源的深度交互、融合进行补足和增强。

例如,通过提取语义特征进行多



▲ 图 4 通信、计算、存储资源融合

模态信息的压缩编码,可以有效降低数据冗余度,提高通信效率,实现“以计算换通信”;反过来,通过将数据分析和计算任务卸载到边缘计算节点或云端进行协作处理,可以有效提高计算效率,实现“以通信换计算”。通过将某些热点流行内容,如音频、视频等,预先存储在网络边缘,能够有效降低网络传输时延,提升用户体验,实现“以存储换通信”。此外,通过缓存某些中间计算结果和相关数据,在进行类似业务处理时能够快速响应、缩短处理时延,实现“以存储换计算”。

6G 时代, TI 作为典型应用,将通过融合通信、计算、存储等网络多维资源,构建“算网一体”“存网一体”的网络新格局。在实时感知网络状态和资源占用情况的基础上,通过网络多维资源的动态分配和调整,可以将业务流量匹配至最优节点。这将有助于缓解甚至消除资源竞争状况,赋能一致化、确定性用户体验,保证用户的沉浸化服务质量需求。

3 存在的问题

3.1 多模态信息的编码传输与融合重建

在多模态通信中,视听编码技术已经较为成熟,而触觉信号由于种类繁多、感知机理复杂,目前尚没有成熟的编码方案。文献[17]提出了一种基于韦伯的显著差异定律的感知盲区方案来解决触觉信号编码问题,并提出了基于波形信号的压缩、数据驱动下的特征提取两种编码方案来解决触觉信号编码问题。由于现实应用场景下两种信号相互作用、影响,需要结合触觉和触觉信息来开发普适的标准触觉编解码器。视听流的稳定传输需要基于高带宽、低抖动,而触觉信号的传输则有着低时延、高可靠性的要

求。现有传输方案大多仅关注触觉流^[9]或视听流^[26],严重影响用户沉浸式体验。在实际的多模态通信中,传输环境是动态的,网络资源是有限的,因此设计自主的跨模态协同传输方案是 TI 保证用户沉浸式体验的关键所在^[25]。此外,对接收跨模态信号进行恢复,一般只能得到和发送端分辨率相同的信号,无法实现全息远程通信的沉浸式体验和虚拟现实下高精度渲染场景。通过深度学习的方法,可以实现多模态信号从低分辨率到超高分辨率的重建,从而增强用户在 TI 下身临其境的体验^[18]。如 C. LEDIG^[27]等从照片感知角度出发,通过生成对抗网络来进行超高分辨率重建,提高了图像超高清重建真实感。

3.2 多模态信息的语义特征提取与压缩

由于 TI 中视觉、听觉、触觉多感官信息存在一定的相关性,因此需采用多模态信息联合编码压缩冗余、提高通信有效性。与此同时,多模态信息也具有差异互补性,多种感官信息的充分融合给用户带来“身临其境”的体验。为了在提高通信效率的基础上增强沉浸式的体验,多模态语义信息提取的准确性与有效性显得尤为重要。多模态语义信息的最佳表示是什么?什么是无损语义压缩?如何量化评估语义保真度和体验沉浸度?这些问题的解决对实现多模态语义通信、提高 TI 用户体验感至关重要。J. BAO^[28]等分析了无损语义压缩,在特定的假设下建立了无损语义数据压缩的理论界限,然而在实际多模态语义通信过程中这些假设都难以成立。至于如何量化评估语义保真度还存在很多困难。针对文本传输的评估量化,现有方法多使用机器翻译任务中的 BLEU^[29]指标。该指标虽然能在一定程度上反映传输准确率,但是无法评估

传输前后的语义一致性。值得借鉴的做法是文献[30]提出的新的度量:利用预训练好的语言模型和传输文本的语义信息来描述它们的语义相似度。

3.3 安全性

传统集中式网络架构缺乏安全内生设计,用户隐私数据泄露问题难以解决。同时,网络安全与 TI 业务各成体系,安全投入高。基于分布式技术设计 AI 赋能的 TI 网络,实现去中心化的安全可信机制,能够有效提高 TI 的安全自治能力。分布式技术在数据产生的地方进行数据处理,无须将数据全部发往云计算中心,这能够在一定程度上缓解隐私数据泄露的问题。网络 AI 可以实现主动免疫、弹性自治、安全泛在、智能协同,为 TI 助力网络安全内生。此外,语义通信需要交换双方感知的知识信息,并协同更新双方的知识库,这可能会带来本地隐私数据泄露的风险。如何开发一种通信参与者之间的高效协调机制,而不造成任何私人数据泄露,仍是一个亟待解决的问题^[31]。由于区块链技术具有高冗余存储、去中心化、公开透明和高安全性等特点,利用区块链技术来存储 TI 中本地隐私数据,可以在一定程度上避免因中心化机构遭受攻击造成的数据泄露^[32]。

4 结束语

为了顺应 6G 时代全面沉浸化的发展需求,本文从研究背景与现状、网络范式、关键使能技术、存在的问题等方面对 TI 进行了系统介绍,并从网络架构、信息传输、资源管理 3 个方面对 TI 关键使能技术进行了探讨。虽然 TI 在多模态信息编码传输与融合重建、语义特征提取与压缩、安全性等方面仍然存在很多需要解决的问题与挑战,但随着新技术的发展,相信

在不久的将来 TI 作为 6G 关键技术将发挥越来越重要的作用。

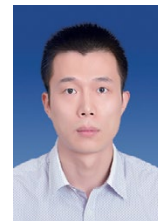
参考文献

- [1] IMT-2030(6G) 推进组. 6G 网络架构愿景与关键技术展望白皮书 [EB/OL]. (2021-09-16) [2021-10-25]. <http://www.caict.ac.cn/>
- [2] 刘子涵. 元宇宙: 人类数字化生存的高级形态 [J]. 新闻阅读, 2021(9): 78-79
- [3] TU-T FG NET2030. New services and capabilities for network 2030: description, technical gap and performance target analysis [EB/OL]. [2021-10-25]. https://www.itu.int/en/ITU-T/focusgroups/net2030/Documents/Deliverable_NET2030.pdf
- [4] HOLLAND O, STEINBACH E, PRASAD R V, et al. The IEEE 1918.1 "tactile Internet" standards working group and its standards [J]. Proceedings of the IEEE, 2019, 107(2): 256-279. DOI: 10.1109/JPROC.2018.2885541
- [5] 胡泳. 所到之处皆媒介: 5G 对媒体产业的影响分析 [J]. 新闻记者, 2020(7): 59-69. DOI: 10.16057/j.cnki.31-1171/g2.2020.07.006
- [6] MAIER M, CHOWDHURY M, RIMAL B P, et al. The tactile Internet: vision, recent progress, and open challenges [J]. IEEE communications magazine, 2016, 54(5): 138-145. DOI: 10.1109/MCOM.2016.7470948
- [7] 王凌豪, 王淼, 张亚文, 等. 未来网络应用场景与网络能力需求 [J]. 电信科学, 2019, 35(10): 2-12. DOI: 10.11959/j.issn.1000-0801.2019223
- [8] FETTWEIS G P. The tactile Internet: applications and challenges [J]. IEEE vehicular technology magazine, 2014, 9(1): 64-70. DOI: 10.1109/MVT.2013.2295069
- [9] AIJAZ A, DOHLER M, AGHVAMI A H, et al. Realizing the tactile Internet: haptic communications over next generation 5G cellular networks [J]. IEEE wireless communications, 2017, 24(2): 82-89. DOI: 10.1109/MWC.2016.1500157RP
- [10] 赵亚军, 郁光辉, 徐汉青. 6G 移动通信网络: 愿景、挑战与关键技术 [J]. 中国科学: 信息科学, 2019, 49(8): 963-987. DOI: 10.1360/N112019-00033
- [11] ATEYA A A, VYBORNOVA A, KIRICHEK R, et al. Multilevel cloud based tactile Internet system [C]//Proceedings of 2017 19th International Conference on Advanced Communication Technology (ICACT). IEEE, 2017: 105-110. DOI: 10.23919/ICACT.2017.7890067
- [12] 魏萌, 吕廷勤. 基于软件定义网络的触觉互联网端到端系统 [J]. 计算机应用研究, 2020, 37(12): 3760-3763. DOI: 10.19734/j.issn.1001-3695.2019.09.0549
- [13] AAZAM M, HARRAS K A, ZEADALLY S. Fog computing for 5G tactile industrial Internet of Things: QoE-aware resource allocation model [J]. IEEE transactions on industrial informatics, 2019, 15(5): 3085-3092. DOI: 10.1109/TII.2019.2902574
- [14] NGMN. 5G white paper [EB/OL]. (2016-04-26) [2021-10-25]. <http://www.ngmn.org/5g-white-paper.html>
- [15] SIMSEK M, AIJAZ A, DOHLER M, et al. 5G-enabled tactile Internet [J]. IEEE journal on selected areas in communications, 2016, 34(3): 460-473. DOI: 10.1109/JSAC.2016.2525398
- [16] GE X H, ZHOU R, LI Q. 5G NFV-based tactile Internet for mission-critical IoT services [J]. IEEE Internet of Things journal, 2020, 7(7): 6150-6163. DOI: 10.1109/JIOT.2019.2958063
- [17] 周亮, 袁哲. 跨模态编码 [J]. 南京邮电大学学报 (自然科学版), 2020, 40(5): 95-100. DOI: 10.14132/j.cnki.1673-5439.2020.05.010
- [18] 高赞, 魏昕, 周亮. 跨模态通信理论及关键技术初探 [J]. 中国传媒大学学报 (自然科学版), 2021, 28(1): 55-63. DOI: 10.16196/j.cnki.issn.1673-4793.2021.01.009
- [19] 钟义信. 信息生态学与语义信息论 [J]. 图书情报知识, 2017(6): 4-11. DOI: 10.13366/j.dik.2017.06.004
- [20] 石光明, 李莹玉, 谢雪梅. 语义通讯 - 智能时代的产物 [J]. 模式识别与人工智能, 2018, 31(1): 91-99. DOI: 10.16451/j.cnki.issn1003-6059.201801008
- [21] GUO W, WANG J, WANG S. Deep multimodal representation learning: a survey [J]. IEEE access, 2019, 7(99): 63373-63394. DOI: 10.1109/ACCESS.2019.2916887
- [22] BALTRUSAITIS T, AHUJA C, MORENCY L P. Multimodal machine learning: a survey and taxonomy [J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 41(2): 423-443. DOI: 10.1109/TPAMI.2018.2798607
- [23] CHO K, MERRIENBOER V B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation [EB/OL]. (2014-09-03) [2021-10-25]. <https://arxiv.org/abs/1406.1078>
- [24] HE Y, XIANG S, KANG C, et al. Cross-modal retrieval via deep and bidirectional representation learning [J]. IEEE transactions on multimedia, 2016, 18(7): 1363-1377. DOI: 10.1109/TMM.2016.2558463
- [25] WEI X, ZHOU L. AI-enabled cross-modal communications [J]. IEEE wireless communications, 2021, 28(4): 182-189. DOI: 10.1109/MWC.001.2000448
- [26] LIU L, HU H, LUO Y, et al. When wireless video streaming meets AI: a deep learning approach [J]. IEEE wireless communications, 2020, 27(2): 127-133. DOI: 10.1109/MWC.001.1900220
- [27] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017: 105-114. DOI: 10.1109/CVPR.2017.19
- [28] BAO J, BASU P, DEAN M, et al. Towards a theory of semantic communication [C]//2011 IEEE Network Science Workshop. IEEE, 2011: 110-117. DOI: 10.1109/NSW.2011.6004632
- [29] PAPINENI K, ROUKOS S, WARD T, et al. BLEU: a method for automatic evaluation of machine translation [C]//Proceedings of the 40th Annual Meeting on Association for Computational Linguistics - ACL '02. Association for Computational Linguistics, 2001: 311-318. DOI: 10.3115/1073083.1073135
- [30] XIE H Q, QIN Z J, LI G Y, et al. Deep learning enabled semantic communication systems [J]. IEEE transactions on signal processing, 2021, 69: 2663-2675. DOI: 10.1109/TSP.2021.3071210
- [31] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to semantic-aware networking: model, architecture, and open problems [J]. IEEE communications magazine, 2021, 59(8): 44-50. DOI: 10.1109/MCOM.001.2001239
- [32] 袁勇, 王飞跃. 区块链技术发展现状与展望 [J]. 自动化学报, 2016, 42(4): 481-494. DOI: 10.16383/j.aas.2016.c160158

作者简介



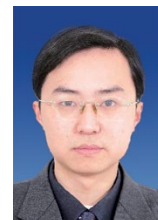
尤子硕, 华中科技大学电子信息与通信学院在读硕士研究生; 研究方向为触觉通信、边缘计算/雾计算、边缘缓存。



李强, 华中科技大学教授、IEEE GLOBECOM 2019 分会联合主席; 主要研究方向为无线协作通信、物联网、认知无线电、触觉通信、边缘计算/雾计算、边缘缓存等; 承担了国家级和省部级科研项目 10 余项; 发表论文 70 余篇, 获授权专利 20 余项。



唐冰, 华中科技大学电子信息与通信学院在读硕士研究生; 研究方向为视频内容理解、语义通信。



葛晓虎, 华中科技大学教授、博士生导师, 国家绿色通信与网络国际联合研究中心主任, 国际信息处理联合会 (IFIP) 中国代表和联合国咨商信息与通信技术专委会委员, 国家科技部和国家自然科学基金委评审专家, 湖北省科技厅评审专家, IET Fellow, 中国通信学会会士; 主要研究方向为绿色通信与网络、超密集网络、大规模 MIMO、毫米波通信技术等; 主持完成了 20 余项国家级科研课题和多项国际高水平科技合作项目; 发表论文 200 余篇, 获授权专利 50 余项。



面向触觉通信的预测窗口和计算卸载策略优化

Optimization of Prediction Window and Computing Offloading Strategy for Haptic Communication

吴岩/WU Yan, 吴丹/WU Dan, 王嵩/WANG Song

(陆军工程大学, 中国 南京 210000)
(Army Engineering University of PLA, Nanjing 210000, China)

摘要: 构建了联合预测和边缘计算的触觉通信框架, 明确了时延和能耗之间的折中关系, 并定义了发送者的预测效益。以总预测效益最大化为目的, 构建了多发送者的预测窗口和计算卸载策略联合优化问题。为便于求解, 将该优化问题转化为多发送者预测与计算卸载博弈, 并提出了一种基于最优响应的预测与计算卸载联合优化算法以收敛纳什均衡解。仿真结果表明, 所提算法可以获得近乎最优的预测窗口和计算卸载策略。

关键词: 触觉通信; 超低时延; 预测; 边缘计算; 博弈

Abstract: A joint prediction and edge computing framework for haptic communications is constructed. The tradeoff between delay and energy consumption is clarified, and the prediction benefit of the transmitter is defined. In order to maximize the sum prediction benefit, a joint optimization problem of multi-transmitter prediction window and computing offloading strategy is constructed. Then, the optimization problem is transformed into a multi-transmitter prediction and computing offloading game, and a joint optimization algorithm of prediction and computing offloading based on the best response is proposed to converge to the Nash equilibrium. Simulation results show that the proposed algorithm can achieve near-optimal performance.

Keywords: haptic communication; ultra-low latency; prediction; edge computing; game

DOI: 10.12142/ZTETJ.202106005

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211119.1936.004.html>

网络出版日期: 2021-11-22

收稿日期: 2021-11-06

无线通信、信号处理和多媒体通信技术的进步推动了触觉互联网的发展, 使触觉互联网可以在传统视听信号的基础上实现触觉信息的实时传输。触觉通信的发展为用户带来了更具沉浸感的服务体验, 也开启了远程操作、自动驾驶等多媒体应用的新时代^[1]。

触觉通信一般包含主域、网络域和控制域3个部分。触觉表现需要依靠肤觉的即时性和力觉的反馈性, 对时延有着极高的要求, 否则会使得通信双方对环境认知出现偏差。为了满足这一要求, 文献[2]从触觉设备信号采集、数据压缩、网络传输协议和编码、资源管理等方面展开了研究。但由于处理时延、传输时延等固有时延的存在, 传统方法在降低触觉通信时延方面并不理想。

考虑到触觉信号在时间上的连

续性和关联性, 通过引入预测思想, 发送者可以根据预测窗口大小对未来动作进行预测并提前发送, 以有效降低端到端时延。现有预测方法大致分为基于模型和基于数据的两种。前者较为简单, 但过于依赖模型的选择和构建, 因此适用场景有限; 后者较为复杂, 但可以通过机器学习等算法, 通过挖掘历史数据中的潜在规律来预测未来趋势。特别地, 由于神经网络具有时间序列预测的优势, 因此基于神经网络的预测方法在网络流

基金项目: 国家自然科学基金(61671474); 江苏省杰出青年基金(BK20180028); 江苏省优秀青年基金(BK20170089)

量、车流量预测方面得到了广泛应用^[3-5]。然而,这种方法的性能优势建立在对大量历史数据进行训练的基础上,加大了发送者自身计算资源的损耗。特别是当发送者计算资源受限时,仅在本地进行计算卸载很可能会增加预测未来动作所需的时延。为此,尝试在预测的同时加入边缘计算,可以使得发送者有机会将预测任务卸载到边缘服务器上,以减少由预测带来的额外时延。

目前,已有一些研究将边缘计算与神经网络相结合,例如:文献[6]将边缘计算融入深度神经网络架构,以应对计算敏感型的处理任务;文献[7]利用边缘计算将深度学习过程从云服务器迁移到边缘节点,以减少工业互联网网络中的数据传输需求并缓解网络拥塞。

如此一来,在联合预测和边缘计算的触觉通信中,发送者根据本地计算资源和预测所需计算资源,既可以将预测任务进行本地卸载,也可以进行服务器卸载。值得注意的是,虽然边缘服务器比本地终端有着更为丰富的计算资源,但过多的发送者同时选择将预测任务卸载到服务器,会使得每个发送者分得的计算资源无法满足相应需求,反而会增加预测所需的时延。因此,在众多发送者之间合理规划计算卸载策略对于每个发送者而言至关重要。此外,值得注意的是,虽然引入预测和边缘计算有利于满足触觉通信的低时延要求,但是预测和计算卸载也会造成发送者能耗的增加。这对于能量受限的发送者而言是不可忽略的。实际上,发送者的能耗不仅与计算卸载策略有关,还与预测窗口的大小有关,即预测窗口越大,预测所需的计算资源越多,由此带来的能耗则越大。因此,为了保证发送者的预测所耗时延和能耗性

能,对预测窗口和计算卸载策略进行联合优化势在必行。

鉴于此,本文通过引入预测和边缘计算的思想,提出了预测窗口和计算卸载策略联合优化方案,以满足触觉通信超低时延和低能耗的双重要求。主要工作总结如下:

(1)构建了联合预测和边缘计算的触觉通信框架,其核心思想是根据优化的预测窗口大小,采用基于长短期记忆神经网络预测模型来预测发送者未来时刻的动作,同时引入边缘计算,激励发送者在本地和边缘服务器之间自主选择计算卸载策略,以避免本地卸载导致的预测所耗时延过高的问题。

(2)定量分析了预测窗口和计算卸载策略对预测所耗时延和发送者能耗的影响,明晰了时延和能耗之间的折衷关系,进而提出预测效益的概念,以定义发送者对于低时延和低能耗的双重要求。

(3)以总预测效益最大化为目标,建立了预测窗口和计算卸载策略联合优化问题。为便于求解,将该优化问题转化为多发送者预测与计算卸载博弈,并证明了纳什均衡的存在性,而后提出一种基于最优响应的预测与计算卸载联合优化算法,以分布

式、低复杂度的方式获得近乎最优的预测窗口和计算卸载策略。

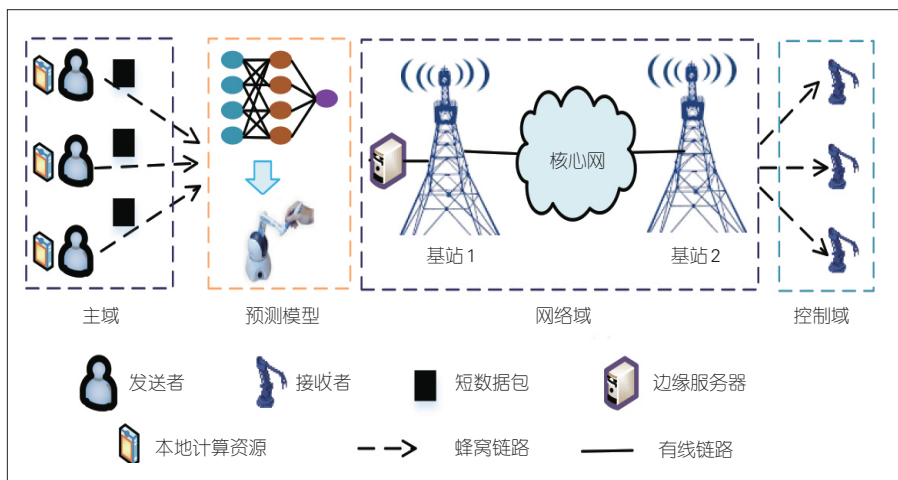
1 联合预测和边缘计算的触觉通信框架

图1展示了联合预测和边缘计算的触觉通信框架,包含主域、预测模型、网络域和控制域。位于主域的发送者通过网络域向位于控制域的接收者发送触觉信号,以对其进行实时操作。其中,触觉信号采用短数据包的形式进行传输。各部分的主要结构和作用概括描述如下。

- 主域:由发送者和人机接口组成,其中发送者为具备一定计算资源的终端设备,人机接口负责将发送者的动作通过编码转化为触觉信号。所有发送者构成集合 $\mathcal{N} = \{1, \dots, n, \dots, N\}$, 其中 N 为发送者的数目,发送者 n 自身具备的计算资源表示为 V_n 。

- 预测模型:由于触觉信号在时间上的连续性和关联性,发送者根据优化获得的预测窗口大小,采用基于长短期记忆(LSTM)神经网络预测模型对未来时刻动作进行预测并提前发送,以满足触觉通信超低时延的要求。

- 网络域:由基站和核心网组成,



▲图1 联合预测和边缘计算的触觉通信框架

负责为触觉通信提供传输媒介。基站1配备有一个边缘服务器,可提供的计算资源总量为 V_B 。

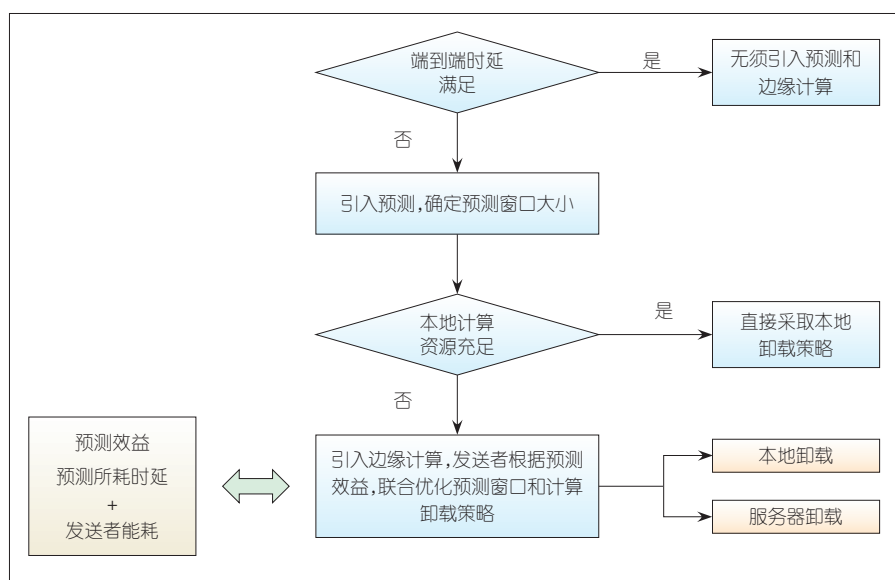
• 控制域:由触觉信号接收者组成,可以和远程环境进行交互。

图2展示了联合预测和边缘计算的触觉通信流程。首先,发送者根据接收者的反馈判断端到端时延是否满足触觉通信超低时延要求。若满足,则无须引入预测和边缘计算,发送者可直接与接收者进行触觉通信;若不满足,则引入基于LSTM神经网络的触觉动作预测模型,并根据预测所耗时延和发送者的能耗确定预测窗口大小。随后,考虑到预测需要以计算资源为支撑,计算资源越多,预测所耗时延就越低。若此时发送者自身具备的计算资源能够满足预测所耗时延要求,则直接采取本地卸载策略;否则引入边缘计算,发送者根据预测效益,联合优化预测窗口和计算卸载策略。

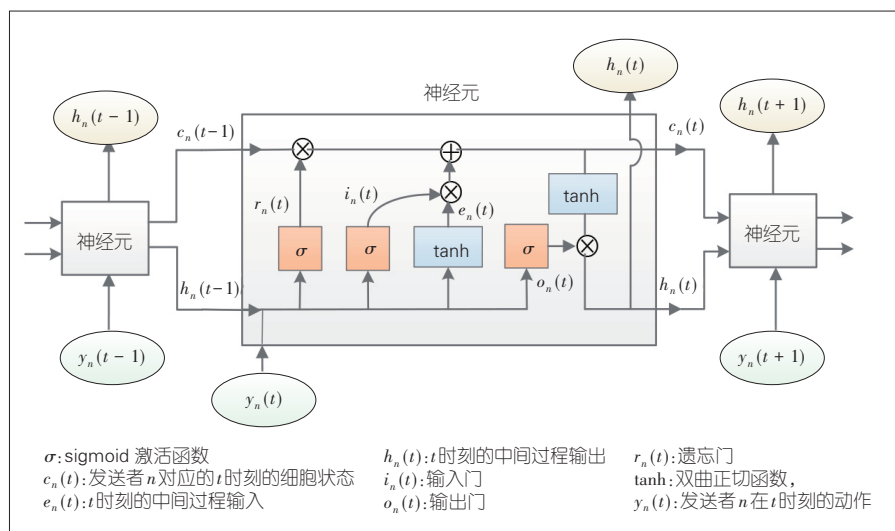
1.1 基于LSTM神经网络的触觉动作预测模型

针对触觉动作所具有的时间序列特性,本文采用基于LSTM的神经网络预测方法。LSTM神经网络不仅可以很好地解决时间序列的预测问题,而且可以通过在神经元中引入门结构,很好地协调历史信息和当前信息,从而有助于解决时间序列中的长期依赖问题。

图3展示了LSTM神经元结构,包含输入门 $i_n(t)$ 、输出门 $o_n(t)$ 和遗忘门 $r_n(t)$ 。其中,输入门用来处理当前序列位置的输入,输出门用来控制信息的输出,遗忘门用来控制是否遗忘细胞历史状态。相应的更新规则如下:



▲图2 联合预测和边缘计算的触觉通信流程



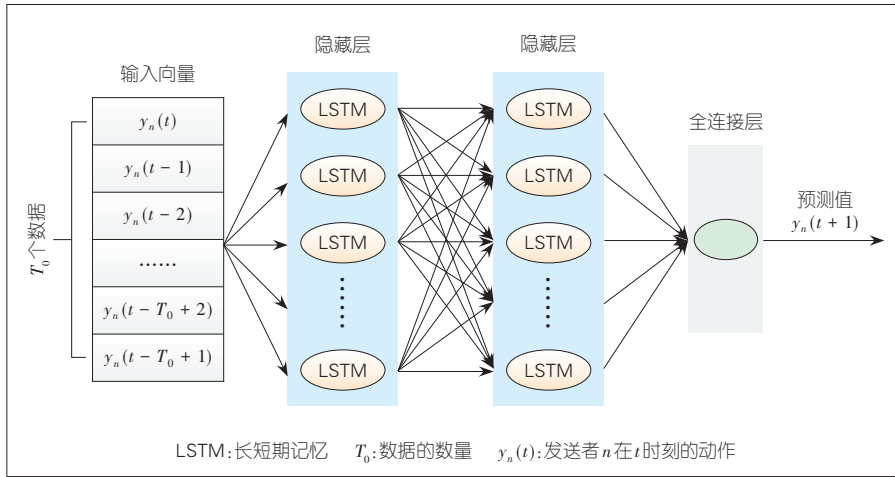
▲图3 长短记忆神经元结构

$$\begin{aligned}
 r_n(t) &= \sigma(Z_r \cdot h_n(t-1) + G_r \cdot y_n(t) + I_r) \\
 i_n(t) &= \sigma(Z_i \cdot h_n(t-1) + G_i \cdot y_n(t) + I_i) \\
 o_n(t) &= \sigma(Z_o \cdot h_n(t-1) + G_o \cdot y_n(t) + I_o) \\
 e_n(t) &= \tanh(Z_e \cdot h_n(t-1) + G_e \cdot y_n(t) + I_e) \\
 c_n(t) &= c_n(t-1) \otimes r_n(t) + i_n(t) \otimes e_n(t) \\
 h_n(t) &= o_n(t) \otimes \tanh(c_n(t))
 \end{aligned} \quad (1)$$

其中, $y_n(t-1)$ 、 $y_n(t)$ 和 $y_n(t+1)$ 分别表示发送者 n 在时刻 $t-1$ 、 t 和 $t+1$ 的动作, $c_n(t)$ 表示发送者 n 对应的 t 时刻的细胞状态, $e_n(t)$ 表示 t 时刻的中间过程输入, $h_n(t-1)$ 、 $h_n(t)$ 和 $h_n(t+1)$

1)分别表示时刻 $t-1$ 、 t 和 $t+1$ 的中间过程输出, σ 为sigmoid激活函数, \tanh 表示双曲正切函数, \otimes 表示对应元素相乘。此外, Z_r 、 Z_i 、 Z_o 、 Z_e 、 G_r 、 G_i 、 G_o 、 G_e 、 I_r 、 I_i 、 I_o 、 I_e 均为线性关系的系数和偏倚。

结合图3所示的LSTM神经元结构,图4展示了基于LSTM神经网络的触觉动作预测模型。令 t 为当前时刻, T_0 为预测未来动作所需要的历史动作数据量, w_n 为预测窗口。触觉动作的预测采用迭代预测方式,即当预



▲图4 基于LSTM神经网络的触觉动作预测模型

测未来 $t + w_n$ 时刻动作时,先用 $t - T_0 + 1$ 到 t 时刻的连续 T_0 个数据作为输入向量,并经过两层 LSTM 隐藏层和全连接层,此时输出 $t + 1$ 时刻的预测动作;而后再用 $t - T_0 + 2$ 到 $t + 1$ 时刻的数据作为输入向量来预测 $t + 2$ 时刻的触觉动作,依次迭代进行,直至输出 $t + w_n$ 时刻预测动作。由此可见,LSTM 神经网络预测所需要的计算资源与预测窗口大小密切相关,即预测窗口越大,迭代预测的次数越多,所需要的计算资源也越多。

进一步地,本文采用均方根误差(RMSE)作为上述预测模型的预测误差指标。当预测窗口为 w_n 时,发送者 n 的预测差错概率为:

$$\varepsilon_n^{\text{predict}} = \sqrt{\frac{1}{w_n} \sum_{i=1}^{w_n} (y_n(t+i) - \hat{y}_n(t+i))^2}, \quad (2)$$

其中, $y_n(t+i)$ 为发送者 n 在 $t+i$ 时刻实际采取的动作, $\hat{y}_n(t+i)$ 为发送者 n 在 $t+i$ 时刻预测的动作。

1.2 发送者预测效益定义

虽然引入预测和边缘计算有利于满足触觉通信超低时延要求,但是预测和计算卸载也会造成发送者能

耗的增加。这对于能量受限的发送者而言是不可忽略的。因此,本文通过明晰时延和能耗之间的折衷关系,定义预测效益这一性能指标。该指标的物理意义在于表征触觉通信低时延和低能耗的双重要求。

当采用本地卸载策略时,即 $x_n = 0$,发送者 n 的预测所耗时延取决于预测窗口和自身的计算资源,即:

$$t_n^{\text{predict}} = \frac{\mu w_n}{V_n}, \quad (3)$$

其中, w_n 为发送者 n 的预测窗口大小, μ 为预测一个传输间隔需要的计算资源, V_n 为发送者 n 自身具备的计算资源。迭代预测方法可以使预测所需的计算资源与预测窗口呈正相关关系。因此,这里采用线性计算方法,即预测窗口为 w_n 时需要的计算资源为 μw_n 。

根据文献[8],本地卸载策略下的能耗为:

$$E_n = \kappa \mu w_n V_n^2, \quad (4)$$

其中, κ 为芯片结构的能量系数。

结合公式(3)和(4),本地卸载策略下发送者 n 的预测效益为:

$$\delta_n = t_0 w_n - t_n^{\text{predict}} - \tau E_n, \quad (5)$$

其中, t_0 为数据包的传输间隔, $t_0 w_n - t_n^{\text{predict}}$ 表示本地卸载策略下触觉通信端到端时延的减少量, τ 为能耗的权重系数。在本地卸载策略下,考虑到发送者能量的有限性,为了避免发送者一味地增大预测窗口而造成过大的能耗,令 $\tau = \tau_0 \sqrt{w_n}$, 其中 τ_0 为常数。相应地,能耗的权重系数将随着预测窗口的增加而不断增加,以实现时延和能耗的折衷。

当采用服务器卸载策略时,即 $x_n = 1$,发送者 n 的预测所耗时延取决于预测窗口 w_n 和边缘服务器分配的计算资源,即:

$$t_n^{\text{predict}} = \frac{\mu w_n}{V_B^n}, \quad (6)$$

其中, V_B^n 为边缘服务器分配给发送者 n 的计算资源。为了满足计算资源分配的公平性原则,边缘服务器为采用服务器卸载策略的发送者等额分配计算资源,即:

$$V_B^n = \frac{V_B}{\sum_{i \in \mathcal{N}} x_i}, \quad (7)$$

其中, V_B 为服务器可提供的计算资源总量, $\sum_{i \in \mathcal{N}} x_i$ 表示所有采用服务器卸载策略的发送者数目。

由于边缘服务器为有源节点,其能耗相对于能量有限的发送者而言可以忽略不计,即 $E_n = 0$ 。因此,在服务器卸载策略下,发送者 n 的预测效益为:

$$\delta_n = t_0 w_n - t_n^{\text{predict}}, \quad (8)$$

其中, $t_0 w_n - t_n^{\text{predict}}$ 表示在服务器卸载策略下触觉通信端到端时延的减少量。需要注意的是,我们之所以没有考虑发送者上行传输的能耗,是因为在触觉通信中发送者本来就需要先向基站1发送触觉信号,而后才能向

核心网传送。这也从侧面体现了在触觉通信中引入边缘计算的优势,即降低能耗并减少回传时延。结合公式(5)和(8),发送者 n 的预测效益可以表示为:

$$\delta_n = \begin{cases} t_0 w_n - \frac{\mu w_n}{V_n} - \tau E_n & x_n = 0 \\ t_0 w_n - \frac{\mu w_n}{V_B^n} & x_n = 1 \end{cases} \quad (9)$$

1.3 触觉通信时延和可靠性约束

结合图1所示的触觉通信框架,发送者 n 的端到端时延可以表示为:

$$t_n = t_n^{\text{up}} + t_n^{\text{predict}} + t_n^{\text{core}} + t_n^{\text{down}} - t_0 w_n, \quad (10)$$

其中, t_n^{up} 为发送者 n 到基站1的传输时延, t_n^{predict} 为预测所耗时延, t_n^{core} 为基站1到基站2的传输时延, t_n^{down} 为基站2到相应接收者的传输时延。

令 t_{\max} 为触觉通信端到端时延约束,则有:

$$t_n \leq t_{\max} \quad (11)$$

发送者 n 的总传输差错概率可以表示为:

$$\varepsilon_n = 1 - (1 - \varepsilon_n^{\text{up}})(1 - \varepsilon_n^{\text{predict}})(1 - \varepsilon_n^{\text{core}})(1 - \varepsilon_n^{\text{down}}), \quad (12)$$

其中, $\varepsilon_n^{\text{up}}$ 为发送者 n 到基站1的传输差错概率, $\varepsilon_n^{\text{predict}}$ 为预测差错概率, $\varepsilon_n^{\text{core}}$ 为基站1到基站2的传输差错概率, $\varepsilon_n^{\text{down}}$ 为基站2到相应接收者的传输差错概率。

由于在高可靠传输下的差错概率一般为 $10^{-4} \sim 10^{-5}$,因此公式(12)可以近似为 $\varepsilon_n = \varepsilon_n^{\text{up}} + \varepsilon_n^{\text{predict}} + \varepsilon_n^{\text{core}} + \varepsilon_n^{\text{down}}$ 。令 ε_{\max} 为触觉通信传输差错概率约束,则有:

$$\varepsilon_n \leq \varepsilon_{\max} \quad (13)$$

2 预测窗口和计算卸载策略联合优化

本节中,我们首先以总预测效益最大化为目标,建立预测窗口和计算卸载策略联合优化问题;随后,将该优化问题转化为多发送者预测与计算卸载博弈,通过定义博弈的势能函数证明纳什均衡的存在性;最后,提出一种基于最优响应的预测与计算卸载联合优化算法,并分析其收敛性和复杂度。

2.1 问题形成

由以上的分析可知,无论是卸载策略还是预测窗口,都对发送者的预测效益产生非常关键的影响。令 $a_n = \{x_n, w_n\}$ 为发送者 n 的联合卸载策略和预测窗口策略。本文旨在优化所有发送者的策略以最大化总预测效益。优化问题可以表示为:

$$\begin{aligned} Q: \max_{a_n} \sum_{n \in \mathcal{N}} \delta_n \\ \text{s.t. C1: } t_n \leq t_{\max} \\ \text{C2: } \varepsilon_n \leq \varepsilon_{\max} \\ \text{C3: } x_n \in \{0, 1\}, \end{aligned} \quad (14)$$

其中,约束条件C1和C2分别表示触觉通信的时延和可靠性约束,C3为卸载策略约束。对于C3, $x_n = 0$ 表示本地卸载策略, $x_n = 1$ 则表示服务器卸载策略。由公式(10)可知,预测窗口仅会影响 $t_0 w_n$ 和 t_n^{predict} 两项时延。根据公式(9),不管采取哪种卸载策略,预测窗口越大,端到端时延的减少量就越多,因此约束条件C1可以视为预测窗口 w_n 的下界约束。同理,由公式(12)可知,预测窗口仅会影响预测差错误差 $\varepsilon_n^{\text{predict}}$,并且预测窗口越大,预测差错误差就越大,因此约束条件C2可以视为预测窗口 w_n 的上界约束。

2.2 博弈模型

公式(14)中的优化问题为混合

整数非线性规划问题。集中式的解决方案将导致较高的复杂度。此外,发送者的卸载策略将影响其他发送者可分配的服务器计算资源,进而影响其他发送者的预测效益。考虑到不同发送者策略之间的交互关系,我们采用博弈论^[9-10]对优化问题进行求解。相应地,优化问题被建立为多发送者预测与计算卸载博弈 $\mathcal{G} = [\mathcal{N}, \mathbb{A}, \{u_n\}_{n \in \mathcal{N}}]$,其中 $\mathcal{N} = \{1, \dots, N\}$ 为发送者的集合, $\mathbb{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_N$ 为所有发送者的策略空间, $\mathcal{A}_n = a_n \otimes w_n$ 为发送者 n 的策略空间, u_n 为发送者 n 的效用函数。

本文中发送者 n 的效用函数被定义为:

$$u_n(a_n, a_{-n}) = \delta_n(a_n, a_{-n}) + \sum_{i \neq n} (\delta_i(a_i, a_{-i}) - \delta_i(a_i, a_{-i|n})), \quad (15)$$

其中, a_{-n} 为除发送者 n 外其他发送者的策略集, $\delta_i(a_i, a_{-i|n})$ 为不考虑发送者 n 策略时发送者 i 的预测效益。因此, $\delta_i(a_i, a_{-i}) - \delta_i(a_i, a_{-i|n})$ 表示在发送者 n 采取策略前后,发送者 i 预测效益的变化量。公式(15)的第1项表示发送者 n 的预测效益,第2项表示发送者 n 的策略对其他发送者的影响。

2.3 纳什均衡的存在性

定义1(纳什均衡):当且仅当没有发送者可以单方面改变策略使得自身效用函数得到提升时, $a^* = (a_1^*, \dots, a_n^*, \dots, a_N^*)$ 被认为是所提多发送者预测与计算卸载博弈 \mathcal{G} 的纳什均衡,如公式(16)所示。

$$\begin{aligned} u_n(a_n^*, a_{-n}^*) &\geq \\ u_n(a_n, a_{-n}^*), \forall n \in \mathcal{N}, a_n \in \mathcal{A}_n. \end{aligned} \quad (16)$$

定义2(精确势能博弈):若存在势能函数 $\phi(a_n, a_{-n})$ 使得公式(17)成

立,则所提多发送者预测与计算卸载博弈 \mathcal{G} 为精确势能博弈。

$$\begin{aligned} u_n(a_n^+, a_{-n}) - u_n(a_n, a_{-n}) &= \\ \phi(a_n^+, a_{-n}) - \phi(a_n, a_{-n}) &= \\ \forall n \in \mathcal{N}, a_n^+ \in \mathcal{A}_n. \end{aligned} \quad (17)$$

公式(17)表示任意发送者策略变化引起的效用函数变化量等于势能函数变化量。

定理1:所提多发送者预测与计算卸载博弈 \mathcal{G} 为精确势能博弈,且最优的联合预测窗口和计算卸载策略为博弈 \mathcal{G} 的纯策略纳什均衡。

证明:定义势能函数为所有发送者的总预测效益,如公式(18)所示。

$$\phi(a_n, a_{-n}) = \sum_{n \in \mathcal{N}} \delta_n(a_n, a_{-n}). \quad (18)$$

当发送者 n 的策略从 a_n 改变为 a_n^+ 时,发送者 n 效用函数的变化量可以表示为:

$$\begin{aligned} u_n(a_n^+, a_{-n}) - u_n(a_n, a_{-n}) &= \\ \delta_n(a_n^+, a_{-n}) - \delta_n(a_n, a_{-n}) &+ \\ \sum_{i \neq n} (\delta_i(a_i, a_{-i}^+) - \delta_i(a_i, a_{-i})) - & \\ \sum_{i \neq n} (\delta_i(a_i, a_{-i}^+) - \delta_i(a_i, a_{-i})) &= 0. \end{aligned} \quad (19)$$

因为 $\delta_i(a_i, a_{-i})$ 为不考虑发送者 n 策略时发送者 i 的预测效益,所以即使发送者 n 的策略从 a_n 改变为 a_n^+ , $\delta_i(a_i, a_{-i}) = \delta_i(a_i, a_{-i})$ 依然成立。公式(19)可以进一步表示为:

$$\begin{aligned} u_n(a_n^+, a_{-n}) - u_n(a_n, a_{-n}) &= \\ \delta_n(a_n^+, a_{-n}) + \sum_{i \neq n} \delta_i(a_i, a_{-i}^+) - & \\ \{ \delta_n(a_n, a_{-n}) + \sum_{i \neq n} \delta_i(a_i, a_{-i}) \} &= \\ \phi(a_n^+, a_{-n}) - \phi(a_n, a_{-n}) &= 0. \end{aligned} \quad (20)$$

可以看出,所提多发送者预测与计算卸载博弈 \mathcal{G} 为精确势能博弈。对于精确势能博弈,势能函数的最大值为博弈 \mathcal{G} 的纯策略纳什均衡解^[11],

因此定理1得证。

2.4 基于最优响应的预测与计算卸载联合优化算法

如算法1所示,我们提出了一种基于最优响应^[12]的预测与计算卸载联合优化算法来获得所提博弈的纳什均衡解。具体地,在每次迭代中,只有一个发送者 n 被随机选择来更新其策略。在发送者 n 计算自身效用函数时,需要知道其他发送者的策略,这意味着策略的更新需要发送者之间的信息交互。一般而言,信息交互可以通过公共信道广播来实现。在每次算法迭代时,只有被随机选出的发送者更新策略,其他发送者将保持他们的策略不变。因此,在迭代过程中由信息交互导致的信令开销非常有限。

算法1 基于最优响应的预测与计算卸载联合优化算法

- 1: 初始化 $t = 0$, 根据公式(14)中约束条件 C1 和 C2 确定预测窗口的下界 w_n^{\min} 和上界 w_n^{\max} 。初始化每个发送者的策略 $a_n = \{x_n, w_n\}$, 其中 $x_n \in \{0, 1\}$, $w_n^{\min} \leq w_n \leq w_n^{\max}$ 。每个发送者在公共信道上广播其初始策略。
- 2: 随机选择一个发送者, 将其记为 n 。发送者 n 计算策略空间 \mathcal{A}_n 中每一个策略对应的效用函数值, 即 $u_n(a_n, a_{-n}), \forall a_n \in \mathcal{A}_n$ 。其他发送者的策略保持不变。
- 3: 在 $t + 1$ 时刻, 发送者 n 选择策略空间中效用函数值最大的策略。

$$a_n(t + 1) = \arg \max_{a_n \in \mathcal{A}_n} u_n(a_n, a_{-n}(t)). \quad (21)$$
 发送者 n 在公共信道上广播其更新后的策略。如果存在多个策略使得效用函数达到最大值, 则从多个策略中随机选择一个策略。
- 4: 如果运行到达了最大迭代次数, 则

算法终止, 否则返回步骤2。

定理2: 算法1可以收敛到所提多发送者预测与计算卸载博弈的纳什均衡解。

证明: 由定理1可知, 所提多发送者预测与计算卸载博弈 \mathcal{G} 为精确势能博弈, 因此发送者策略变化引起的效用函数变化量等于势能函数变化量。由于在算法每次迭代中, 发送者的策略更新总是可以使其效用函数值得到增加, 相应地, 势能函数值也在不断增加。由于发送者的策略空间为有限空间, 因此, 算法1可以在有限的迭代次数中收敛到势能函数的局部最优解或者全局最优解, 即所提多发送者预测与计算卸载博弈的纳什均衡解。

算法1的复杂度主要由发送者的效用函数计算过程决定。具体地, 在每次迭代中, 发送者需要计算策略空间 \mathcal{A}_n 中每一个策略对应的效用函数值。由于 $x_n \in \{0, 1\}$, $w_n^{\min} \leq w_n \leq w_n^{\max}$, 因此每次迭代的复杂度为 $2(w_n^{\max} - w_n^{\min} + 1)$ 。令 T_{\max} 为最大的迭代次数, 则算法1的复杂度为 $\mathcal{O}(T_{\max}(w_n^{\max} - w_n^{\min} + 1))$ 。

3 数值仿真结果

3.1 基于LSTM神经网络的触觉动作预测模型

本节首先对基于LSTM神经网络的触觉动作预测模型进行仿真分析。该模型采用两层LSTM隐藏层结构, 每层神经元个数分别设置为80和100, 预测所需要的历史数据量 T_0 为1 000。图5给出了训练次数分别为10、20和30的预测误差曲线。随着预测窗口的增加, 每条曲线的预测误差都不断增加。这是因为预测窗口越大, 由LSTM所引发的预测误差传

播对后续的触觉动作预测影响就越大。此外,训练次数的增加并不一定会降低预测差错概率,这是因为训练次数过多可能导致过拟合的现象。因此,在后续仿真中,我们设置训练次数为20。

3.2 基于最优响应的预测与计算卸载联合优化算法

这里,我们对算法1的性能进行仿真分析。仿真参数设置如下:发送者数目 N 为12,预测一个传输间隔需要的计算资源 μ 为 10^6 转,边缘服务器计算资源 V_B 为 20×10^9 转/s,每个发送者的计算资源 V_n 服从随机分布 $[2,3] \times 10^9$ 转/s,传输间隔 t_0 为1 ms,能效权重系数 τ_0 为 0.12×10^{-9} 。由公式(14)中的约束条件C1和C2可知,触觉通信中的预测窗口下界 w_n^{\min} 和上界 w_n^{\max} 需要结合主域、核心网以及控制域的传输时延和传输差错概率共同求得。在仿真中,为便于算法1的仿真分析,我们设 w_n^{\min} 和 w_n^{\max} 分别为5和20,即 $w_n \in [5, 20]$ 。

图6分析了算法1的收敛性。随着迭代次数的增加,系统总预测效益不断增加,并最终达到收敛。需要注意的是,所提算法只能收敛到局部最优解,并不能保证收敛到全局最优方案。算法1的优势在于复杂度较低,

且可以以较快的收敛速度达到接近最优方案的性能。

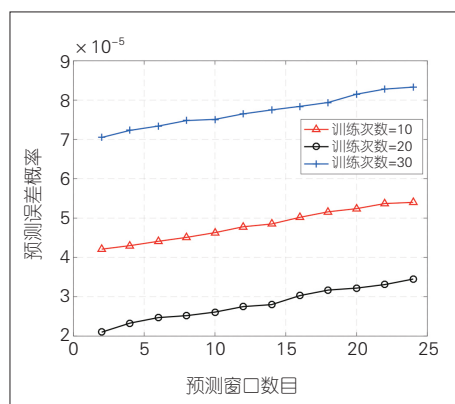
图7分析了系统总预测效益与发送者数目的关系。其中,“服务器卸载”是指所有发送者都采用服务器卸载策略,由公式(9)可知,此时每个发送者都采用最大预测窗口以使总预测效益最大;“本地卸载(最大窗口)”是指所有发送者都采用本地卸载策略,且预测窗口都采用最大值;“本地卸载(最优窗口)”是指所有发送者都采用本地卸载策略,且预测窗口都采用最优值;“随机卸载”是指每个发送者的卸载策略和预测窗口都随机分配。

首先,随着发送者数目的增加,所提算法的系统总预测效益不断增加,且优于其他方案的性能。这验证了所提算法的可行性和优越性。其次,当发送者数目较小时,服务器卸载方案性能优于本地卸载方案性能。这是因为:一方面,服务器的能耗可以忽略不计;另一方面,当发送者数目较小时,每个发送者从服务器获得的计算资源较多,从而预测所耗时延也就越低。然而,当发送者数目不断增加时,服务器卸载方案性能先增加而后逐渐降低,且性能劣于本地卸载方案。这是因为随着发送者数目的增加,每个发送者从服务器获得的计

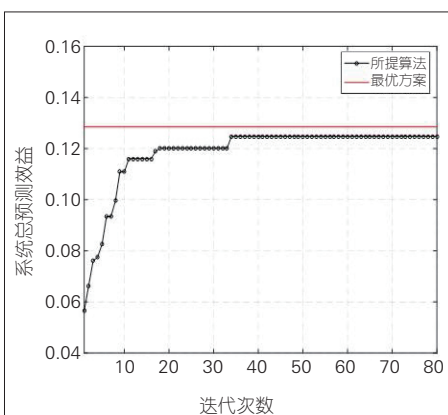
算资源逐渐减少,预测所需的时延也就越多,进而导致总体预测效益逐渐降低。最后,不管发送者数目是多少,本地卸载(最优窗口)方案的性能都优于本地卸载(最大窗口)方案的性能。这验证了优化预测窗口可以有效提升发送者的预测效益。

图8分析了系统总预测效益与能效权重系数的关系。可以看出,除服务器卸载方案外,其他方案的系统总预测效益都会随着能效权重系数的增加而不断降低。这是因为能效权重系数越大,预测效益就越低。而对于服务器卸载方案,其能效可以忽略不计,所以总预测效益不受能效权重系数的影响。此外,当能效权重系数较小时,本地卸载(最优窗口)方案与本地卸载(最大窗口)方案的性能相同,这说明此时最大预测窗口就是最优预测窗口。而随着能效权重系数的增加,本地卸载(最优窗口)方案逐渐优于本地卸载(最大窗口)方案的性能,且性能差距逐渐增大。

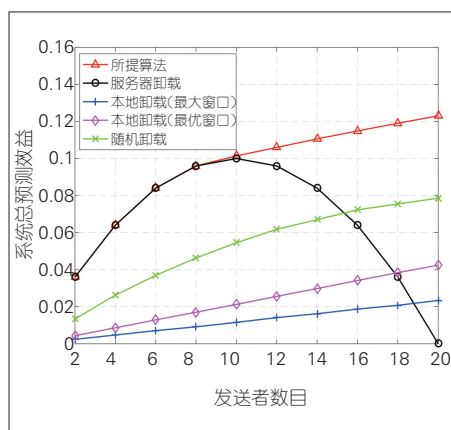
图9分析了系统总预测效益与服务器计算资源的关系。除本地卸载策略下,其他方案的总预测效益都随着服务器计算资源的增加而不断增加。这是因为当服务器计算资源不断增加时,发送者采用服务器卸载策略可以分配的计算资源就越多,预测



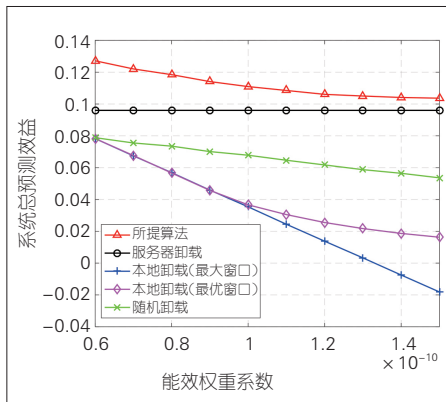
▲图5 基于长短期记忆神经网络的预测误差与预测窗口关系



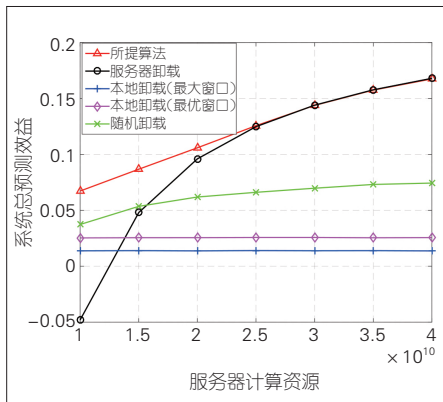
▲图6 算法1收敛性分析



▲图7 系统总预测效益与发送者数目的关系



▲图8 系统总预测效益与能效权重系数的关系



▲图9 系统总预测效益与服务器计算资源的关系

所耗时延会不断降低,预测效益就会不断增加。对于本地卸载方案,其总预测效益与服务器计算资源没有关系。当服务器计算资源较低时,本地卸载方案的性能优于服务器卸载方案。当服务器计算资源逐渐增加时,服务器卸载方案的性能逐渐超过本地卸载方案,且逐渐趋于所提算法。

4 结束语

本文构建了联合预测和边缘计算的触觉通信框架以满足触觉通信的超低时延要求,其核心思想是根据优化的预测窗口大小,采用基于LSTM神经网络预测模型来预测发送者未来时刻的动作,并通过引入边缘计算激励发送者在本地和边缘服务器之间自主选择计算卸载策略。此外,本文明晰了时延和能耗之间的折衷关系,并由此定义了发送者的预测效益;为便于求解,通过势能博弈对优化问题转化,并提出一种基于最优响应的预测与计算卸载联合优化算法,以分布式、低复杂度的方式获得近乎最优的预测窗口和计算卸载策略。

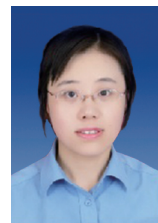
参考文献

- [1] ALJAZ A, DOHLER M, AGHVAMI A H, et al. Realizing the tactile Internet: haptic communications over next generation 5G cellular networks [J]. IEEE wireless communications, 2017, 24(2): 82–89. DOI: 10.1109/MWC.2016.1500157RP
- [2] ANTONAKOGLU K, XU X, STEINBACH E, et al. Toward haptic communications over the 5G tactile Internet [J]. IEEE communications surveys & tutorials, 2018, 20(4): 3034–3059. DOI: 10.1109/COMST.2018.2851452
- [3] BI J, ZHANG X, YUAN H T, et al. A hybrid prediction method for realistic network traffic with temporal convolutional network and LSTM [J]. IEEE transactions on automation science and engineering, 7537, 99: 1–11. DOI: 10.1109/TASE.2021.3077537
- [4] MACKENZIE J, RODDICK J F, ZITO R. An evaluation of HTM and LSTM for short-term arterial traffic flow prediction [J]. IEEE transactions on intelligent transportation systems, 2019, 20(5): 1847–1857. DOI: 10.1109/ITITS.2018.2843349
- [5] KANNAN S, YENGERA G, MUTTER D, et al. Future-state predicting LSTM for early surgery type recognition [J]. IEEE transactions on medical imaging, 2020, 39(3): 556–566. DOI: 10.1109/TMI.2019.2931158
- [6] LI E, ZENG L K, ZHOU Z, et al. Edge AI: on-demand accelerating deep neural network inference via edge computing [J]. IEEE transactions on wireless communications, 2020, 19(1): 447–457. DOI: 10.1109/TWC.2019.2946140
- [7] LIANG F, YU W, LIU X, et al. Toward edge-based deep learning in industrial Internet of Things [J]. IEEE Internet of Things journal, 2020, 7(5): 4329–4341. DOI: 10.1109/JIOT.2019.2963635
- [8] 吴学文, 廖婧贤. 云边协同系统中基于博弈论的资源分配与任务卸载方案 [J]. 系统仿真学报, 2021. DOI: 10.16182/j.issn1004731x.joss.21-0077
- [9] BACCI G, LASAULCE S, SAAD W, et al. Game theory for networks: a tutorial on game-theoretic tools for emerging signal processing applications [J]. IEEE signal processing magazine, 2016, 33(1): 94–119. DOI: 10.1109/MSP.2015.2451994
- [10] FANG T, YUAN F, AO L, et al. Joint task offloading, D2D pairing and resource allocation in device-enhanced MEC: a potential game approach [J]. IEEE Internet of Things journal, 7754, PP(99): 1. DOI: 10.1109/JIOT.2021.3097754
- [11] DUY Q L, CHEW Y H, SOONG B H. Potential game theory applications in radio resource allocation [M]. Berlin: Springer International Publishing, 2016
- [12] XU Y H, WANG C G, CHEN J H, et al. Load-aware dynamic spectrum access for small-cell networks: a graphical game approach [J]. IEEE transactions on vehicular technology, 2016, 65(10): 8794–8800. DOI: 10.1109/TVT.2015.2508998

作者简介

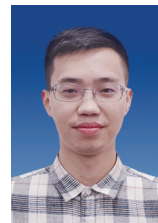


吴岩,陆军工程大学在读博士研究生;主要研究领域为跨模态通信、D2D内容共享;发表论文4篇。



励;发表论文20余篇。

吴丹,陆军工程大学副教授、国家优秀青年基金获得者;主要研究领域为跨模态通信、D2D内容共享、协同通信;先后主持国家自然科学基金项目10余项;获中国通信学会科学技术奖一等奖、教育部自然科学奖二等奖等多项科研奖励。



王嵩,陆军工程大学在读硕士研究生;主要研究领域为触觉通信、资源管理;发表论文1篇。



一种面向视触融合人机交互的 柔性触觉传感阵列

A Novel Flexible Haptic Sensor Array for Visual Haptic Fusion Human-Machine Interaction

王爽/WANG Shuang, 吴兵/WU Bing, 刘倩/LIU Qian

(大连理工大学, 中国 大连 116033)
(Dalian University of Technology, Dalian 116033, China)

摘要:提出了一种面向视触融合人机交互的三维柔性触觉传感器以及相应的触觉采集软硬件装置,用于采集真实物理操作过程中的触觉力信息。该触觉传感器由前、左、右3面组成,可佩戴在手指尖部位并捕捉抓取过程中的触觉信息。通过实验测量,该传感器的测量范围是0~20 N,检测精度为0.1 N,平均响应时间为23.8 ms,能够提供实时、高精度的触觉信息采集。

关键词:柔性触觉传感器;触觉互联网;虚拟现实

Abstract: A flexible 3D haptic sensor and the corresponding signal acquisition software and hardware to collect tactile information during real physical operation are proposed. The tactile sensor consists of the front, left, and right sides, which can be worn at the fingertips of the human hand and capture the tactile information during grasping. Through experimental results, the sensor can range from 0 to 20 N with a detection accuracy of 0.1 N and an average response time of 23.8 ms, which enables to provide real-time and high-precision haptic acquisition.

Keywords: flexible haptic sensor; tactile Internet; virtual reality

DOI: 10.12142/ZTETJ.202106006

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211208.0939.004.html>

网络出版日期: 2021-12-09

收稿日期: 2021-11-10

1 触觉通信的发展

早在1948年,美国Arogonne国家实验室开发的用于核反应堆操作的机械式力矩反馈主-从系统开创了遥操作研究的先河。随后,美国Utah大学开发了具有高保真力反馈的主-从遥操作系统,用以完成多种海底精细作业任务^[1]。日本港湾空港技术研究所也研制了水下作业工程机械力反馈遥操作系统,使水下作业

的效率基本达到了陆地水平^[2]。2010年,日本庆应义塾大学成功研制了触觉钳子手术机器人,并通过小鼠肝脏组织剥离实验证实:该触觉钳子能够实现在医院的楼宇范围内高性能的触觉传递,使医生在远离“患者”的位置也能清晰地感觉到钳子与活体组织间微弱的触感^[3]。近年来,随着中国空间站计划、“嫦娥”探月计划、“蛟龙”深海潜水器等项目的启动实施,有触觉传递功能的机械臂也正逐渐获得更密切的关注。可以看出,触觉通信技术目前在医疗外科、水下勘探、军事排雷等各个领域的应用不断

增加,这使得人类用户可以在安全距离执行复杂的甚至是非常危险的任务^[4]。

随着5G通信、云计算和虚拟现实(VR)等技术的发展,低延迟高可靠的触觉互联网通信成为可能。触觉互联网能够提供实时的触觉控制和远端物理触觉体验,在视触觉融合的人机交互等领域有广泛的应用前景。特别是在VR领域,触觉可以作为目前主流基于视觉VR技术的补充,给用户带来更加强烈的沉浸感。这主要体现在与皮肤接触时使用振动或静电冲击来模拟触感。通过使

基金项目:国家自然科学基金(62071083);中央高校基本科研业务费基金(DUT21GJ208)

用触觉执行器(专门开发的手套、鞋子、操纵杆等),用户可以从手中或身体的其他部位获得计算机程序的反馈。史蒂文·斯皮尔伯格的《Ready Player One》展示了未来VR世界,让用户注意到高质量沉浸式体验不仅需要视觉,还需要触觉反馈。现在越来越多的公司开发了面向触觉的设备,如Gloveone VR手套、微软PIVOT^[5]等。然而,人类的触觉感知和通过触觉导向设备进行的触觉演示之间仍然存在一定差距。在多模态人机交互应用中,这种差距一般通过触觉渲染^[6]来弥补,而触觉渲染算法的基础是真实物理操作过程中的触觉信息。因此,能够提供原始触觉数据基础的触觉传感器成为实现高保真视触融合人机交互应用的必然需求。

本文提出了一种新型的柔性分布式触觉传感器,该传感器采用压阻式原理,具有多接触面多(底、左、右3面)、重量轻、体积小等特点。它可以穿戴于人体的指尖、机械手或机械臂的末端夹持器上。同时,针对该分布式柔性触觉传感器,本文提出相应的信号检测软硬件系统。由于独特的3面设计,使触觉传感器更适用于采集人手或灵巧末端夹持机构(例如五指灵巧机械手)抓取时多姿态、多接触面的触觉信息,进而能够更准确地为VR触觉渲染和机械手仿人抓取提供数据支撑。

2 触觉传感器介绍

Z. KAPPASSOV将触觉传感器定义为一种能够获取被接触物体属性的设备^[7]。触觉传感器按照工作原理可分为压阻式、压电式、光学式、磁力式等。其中,压阻式触觉传感器通过测量敏感材料的电阻值变化来检测施加力的大小和位置等信息。压阻

式触觉传感器的原理简单、成本低,因此得到了广泛应用。压阻式触觉传感器根据敏感材料的不同,又分为微机电系统(MEMS)应变计类、导电聚合物类、导电橡胶类和导电溶液类。MEMS应变计利用的是导体或半导体在外力作用下产生机械变形时电阻会发生变化的特性。典型器件是美国伊利诺伊大学香槟分校利用聚酰亚胺(PI)作为基底设计的仿生皮肤。导电聚合物类触觉传感器是以导电聚合物(例如离子聚合物Flemion导电膜)作为敏感材料的。导电橡胶类传感器是在有机弹性材料(如硅橡胶)中添加炭黑、碳纳米管等导电颗粒制成柔性导电复合材料。典型器件是美国马里兰大学研发的触觉传感器,使用混有碳纳米管的聚二甲基硅氧烷(PDMS)作为敏感材料,以PDMS作为基底封装。导电溶液类触觉传感器的典型代表是美国南加州大学研制的BioTac触觉仿生手指,可以直接安装在机械手上^[8]。压阻式触觉传感器各种类别的优缺点总结如表1所示。

S. TESHIGAWARA等设计了一种基于压阻效应的触觉传感器,并利用物体滑动时的频率变化进行滑动检测^[9]。A. SCHMITZ等设计了一款电容式压力分布传感器,并部署在iCub仿人机器人上^[10]。P. A. SCHMIDT等设计了一款电容式传感器,能够感知到5 mN的力值变化,但

传感器在抓取过程中容易损坏^[11]。B. CHOI等设计了一种具有微型指尖触觉传感器的拟人机械手SKKU-Hand II,其机械手指尖安装了由基于偏氟乙烯(PVDF)和压变电阻组成的触觉传感器,能够感知机械手抓取过程中的力变化^[12]。BioTac是SynTouch公司开发的一款能够感知压力、温度和振动信息的触觉传感器,3种触觉信息通过3组独立的传感器实现,然后集成到一起。M. K. JOHNSON等设计一款能够显示物体表面几何信息的触觉传感器GelSight,主体由凝胶和相机构成^[13],传感器表面是一块涂有反光材料的透明弹性体膜。当物体接触到传感器时,弹性体发生变形。通过相机捕捉弹性体的变形信息,即可重建出接触物体的表面信息。S. SUNDARAM等设计了一款由压阻膜组成的可伸缩触觉手套,每个触觉手套上含有548个触觉传感器。通过触觉手套可以识别单个物体、估计其重量,还可以探索抓取物体过程中的典型触觉模式^[14]。

3 触觉传感阵列设计与触觉信号检测套件

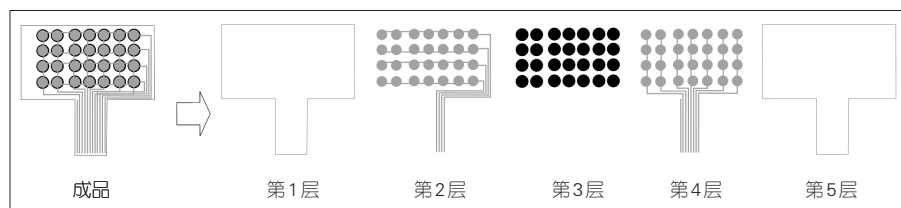
3.1 传感器设计与制作

传感器采用双面布线的方式,一共分为5层,整体结构如图1所示:第1层为柔性非导电材料层,使用聚对苯二甲酸乙二醇酯(PET)薄膜;第2

▼表1 不同压阻式传感器优缺点

种类	柔性基底+敏感材料举例	优点	缺点
MEMS应变计	<ul style="list-style-type: none"> PI+金属应变计 PI+NiCr PDMS+应变片 	<ul style="list-style-type: none"> 灵敏度高 空间分辨率高 	<ul style="list-style-type: none"> 单元易碎 成本较高 封装后尺寸大
导电聚合物	PDMS+Flemion	<ul style="list-style-type: none"> 弯曲性好 成本低 	灵敏度低
导电橡胶	<ul style="list-style-type: none"> PDMS+混有CNT的PDMS PDMS+INASTROMER导电橡胶 	<ul style="list-style-type: none"> 弯曲性好 可检测静态动力 	存在迟滞现象
导电溶液	柔性橡胶+导电液(BioTac)	<ul style="list-style-type: none"> 集成度高 稳定性好 	<ul style="list-style-type: none"> 信号串扰 密封性要求高

CNT:碳纳米管 MEMS:微机电系统 PDMS:聚二甲基硅氧烷 PI:聚酰亚胺



▲图1 传感器布线与结构图

层为电极层,采用喷墨打印的方式将导电银浆在第1层上印刷4行导线和28个圆形电极;中间层(第3层)是力敏材料层,可将石墨烯力敏油墨印刷于圆形电极处;第4层为对称电极层,也采用喷墨打印的方式将导电银浆在PET层上印刷7列导线和28个圆形电极;第5层也是柔性非导电材料PET聚酯薄膜,第4层的导电银浆也印刷在其上。第2层和第4层可以互换。因此,传感器有4行7列,共有28个传感单元、11条输出导线。其中,11条导线分别代表4行和7列,用来定位每个感应点。

为了更好地贴合机械手,传感器表面设有引导线,可按引导线进行弯折。另外,考虑弯折处应力产生的测量误差,我们需要在弯折引导线附近对基材进行镂空设计。这时信号采集的电极线向远离弯折线的方向

避让。

触觉传感器的制作过程共分为3步:第1步是在150°下预缩PET聚酯薄膜,并将它作为双面基底,厚度为50 μm;第2步是在上层基底喷墨印刷7列导电银浆布置电路,厚度为5 μm,下层基底喷墨印刷4行导电银浆布置电路,厚度为5 μm;第3步是在两层银导线中间喷墨印刷石墨烯力敏油墨,厚度为20 μm。整个传感器一共有4行7列,共28个单元,左边第2列和第3列以及右边第2列和第3列中间有引导线(可弯折形成3面结构)。这4列传感单元圆心距为8 mm,其他列传感单元圆心距为7 mm。传感器成品整体长58 mm,列宽32 mm,厚度为130 μm。

3.2 触觉信号采集

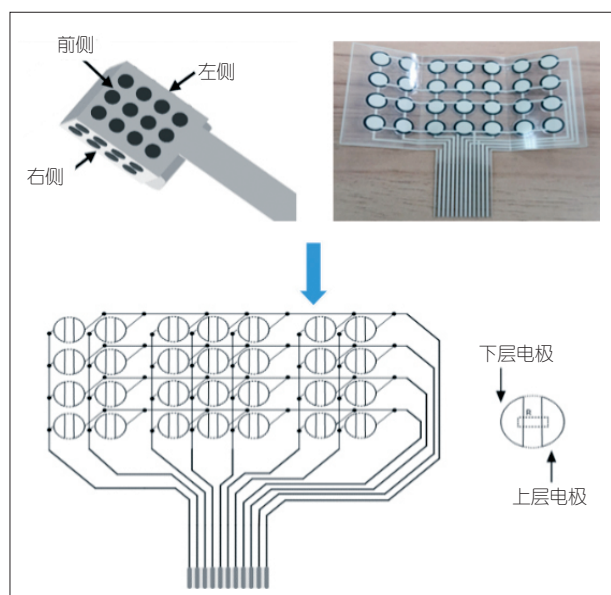
传感阵列的等效电阻如图2所

示,触觉信号的采集通过循环扫描电路实现,并利用零电位法(红色小电阻)来消除环路干扰(原理如图3所示)。当硬件实现时,7条列导线用单片机循环选通,并通过模数转换器(ADC)采集4条行导线的输出电压。假设位于第*i*行、第*j*列的传感单元受压力,则力敏电阻值 r_{ij} 发生变化。由式(1)可知,输出电压 V_{out} 的值将发生变化;受到的压力越大,输出电压值越大。

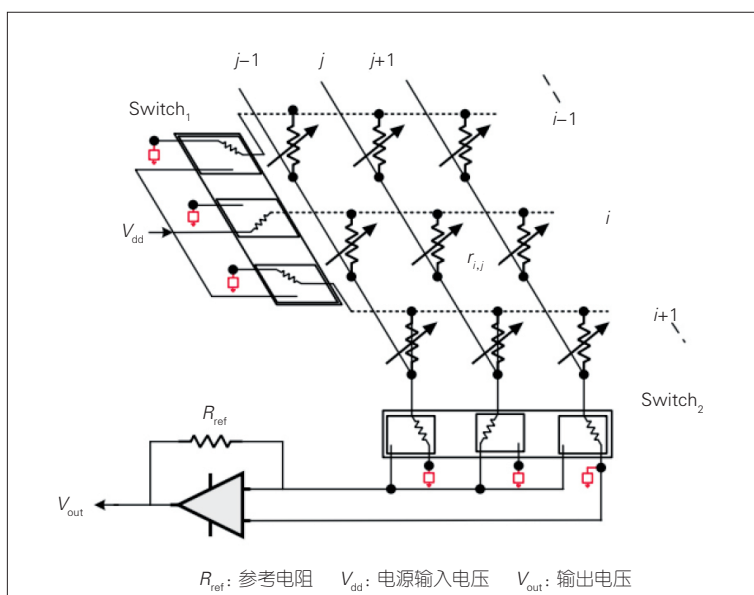
$$V_{out} = V_{dd} * \left(\frac{-r_{ij}}{R_{ref}} \right), \quad (1)$$

其中, V_{dd} 表示电源输入电压, R_{ref} 表示参考电阻。因此,通过循环扫描开关阵列1(Switch₁)的第*i*行和开关阵列2(Switch₂)的第*j*列,测量输出电压值,即可获得触觉传感器28个传感单元的受压力大小。

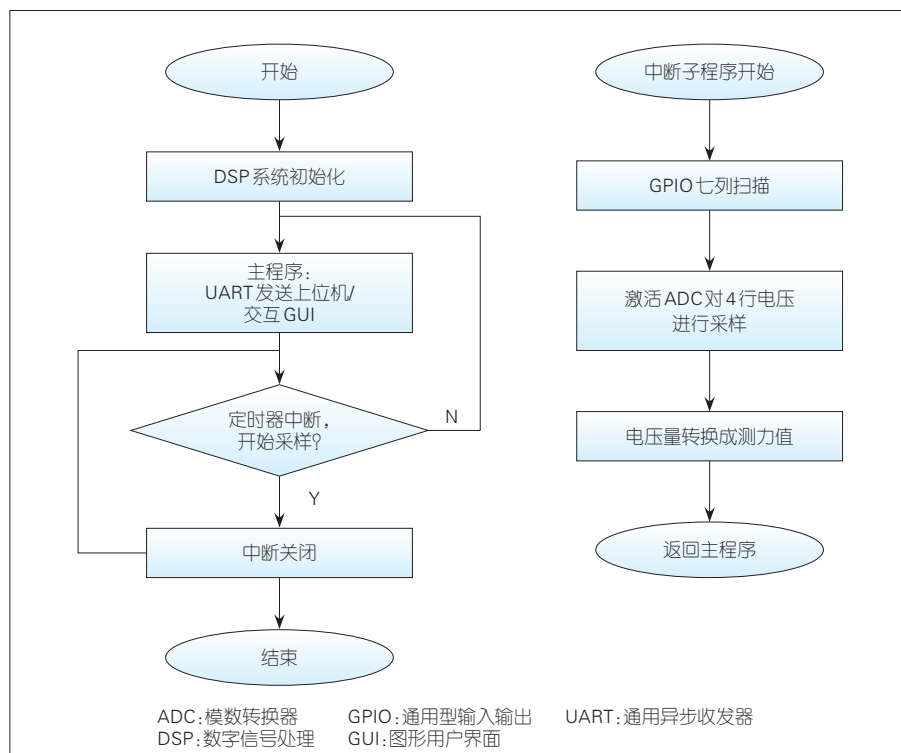
触觉信号的具体采集流程如图4所示。首先我们初始化7个通用型输入输出(GPIO)引脚口,用于选通并拉高电压;然后定义4路ADC采集通道,用于将模拟信号转换为单片机可以处理的数字信号;最后初始化通用



▲图2 等效电阻示意图



▲图3 循环扫描电路图



▲图4 主控制程序流程图

异步收发器(UART)模块,用于将单片机采集回的数据上传给上位机。初始化完成后,待上位机发送指令,触觉信号采集软硬件设备开始采集触觉信号。首先,GPIO选通某一列,然后激活ADC,从而对四行电压进行采样。如此循环7次,得到28个电压采样值,再通过图5所示的测力计进行标定。

4 有限元仿真

当传感器受力时,中间的力敏材料层产生微小位移,阻值发生改变。通过测量电阻值的变化量可求出每个传感单元上方的一维力大小,通过获取多个单元的电阻变化就可以计算出整个传感器的三维合力大小。

传感器为集成在机器人指端的触觉传感器。通过将传感器贴附在

机器人指端,我们可以获取指端接触物体时的触觉信息。传感器三维模型与尺寸如图6所示,每个传感单元可等效为直径3 mm、厚30 μm的圆柱。传感器主要由PET聚酯薄膜、导电银浆和石墨烯力敏油墨组成,贴附在指端基体上。为简化仿真实验,我们将导电银浆和石墨烯力敏油墨合并成一层。

我们使用有限元分析软件对传感器进行仿真实验,分析每个传感单元受法向载荷和切向载荷作用时对应的应力分布情况。图7展示了不同情况下的指端与接触面的相对表现和受力分析。

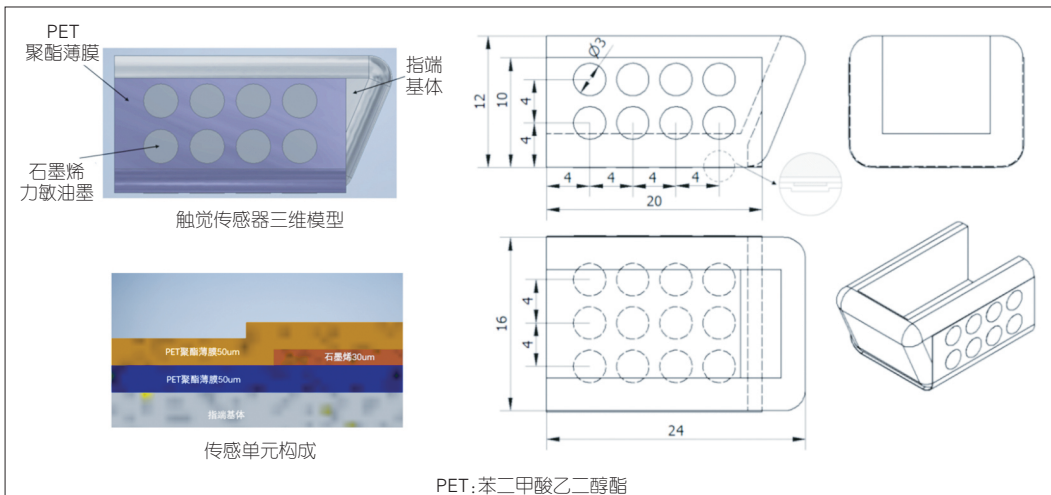
综上所述,本项目设计的触觉传感器切实可行,具有触觉压力场和滑移剪切力场三维的精确感知能力。

5 实验结果及参数指标

我们使用如图5所示的测力计作为标准压力信号值,从而将力信号的大小与传感器测量的电压信号一一对应。实验结果如图8中的绿点所示。为了对力与电压的关系进行建模,我们对测试数据进行曲线拟合,具体如图8中的蓝线所示。力与电压的关系为: $f(x) = p_1 \cdot x^3 + p_2 \cdot x^2 +$



▲图5 测力计



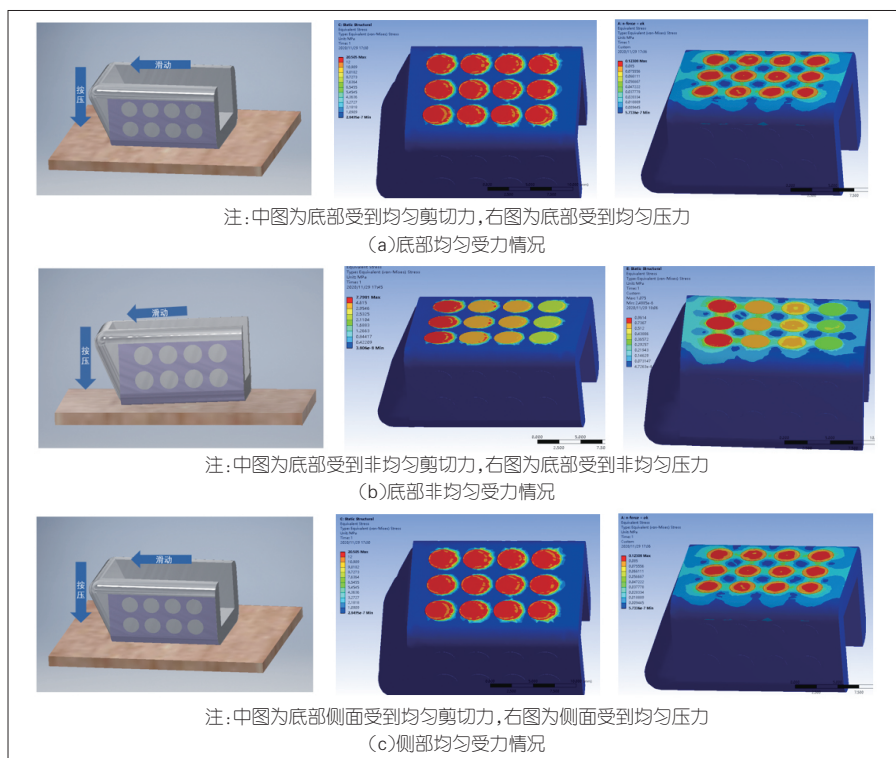
▲图6 传感器的三维模型、尺寸以及传感单元构成

$p_3 \cdot x + p_4$ 。其中, p_1 、 p_2 、 p_3 、 p_4 分别等于 4.121、-4.129、8.506、-0.8206。如图 8 所示, 当压力处于 20 N 以下时, 力与电压的关系是非线性关系, 可以认为此区间为传感单元非线性受力区间。从图 8 可以看出, 该传感器在 0~20 N 具有较高的灵敏度。当测力计以 0.1 N 稳定施加时, 传感器有稳定的反馈, 即所提出的触觉传感器测量精度为 0.1 N。触觉传感器的另一个关键性能参数是响应时间, 它表示测力计开始向传感器单元提供压力与计算机打印出电压信号的时间差。响应时间即响应延迟, 由周期延迟 (由 DSP 扫描过程引起) 和输出延迟 (由计算机引起) 组成。由于提出的触觉传感器阵列包含 4×7 个传感器单元, 因此 ADC 每次读取 4 行, 循环 7 次则可以完成一个周期的采集。其中, 一次循环延迟 \approx ADC 采集延迟 + 电压稳定延迟, 周期延迟 \approx (ADC 采集 4 行信号的采集延迟 + 电压稳定延迟) $\times 7$ (列), 多次测量获得平均响应时间为 23.8 ms。

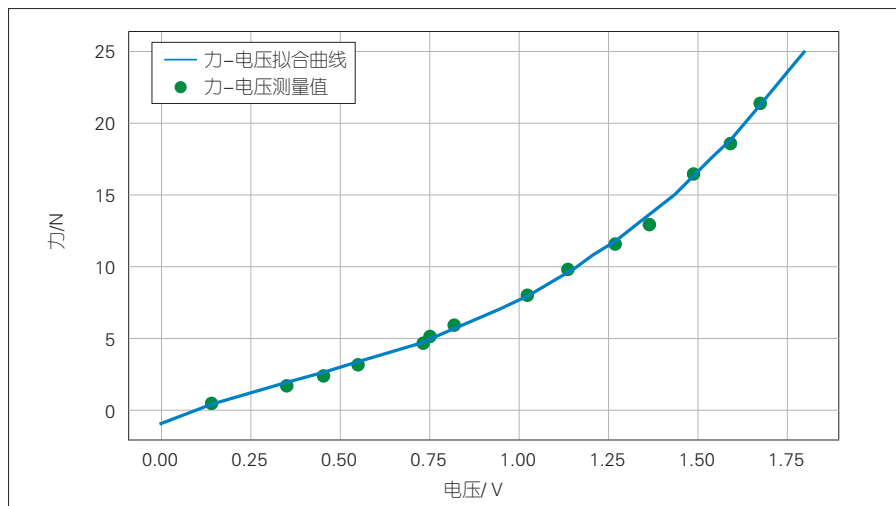
图 9 展示了传感器在不同受力情况下的触觉热力图。可以看出, 在传感器底部均匀受力情况下, 每一列上的压力大致相同; 在传感器底部受力不均匀情况下, 从左到右各列压力值依次减小。图 10 展示了传感器在物体表面滑动时的实时受力情况。在传感器上固定了一个橡皮擦, 然后拖动它滑过在桌面上固定好的牙签。可以看出, 本文提出的传感器准确地感知了整个过程中力的变化。

6 结束语

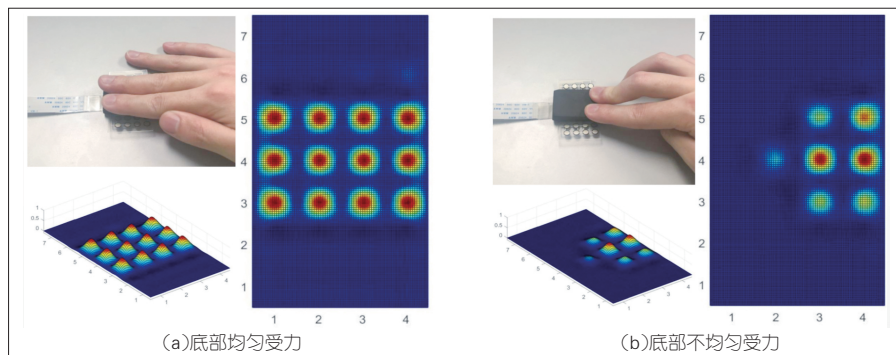
本文中, 我们提出了一种基于压阻原理的三维柔性触觉传感器阵列。该传感器通过测量所有单元的电阻变化来确定整个三维阵列的力。该传感器的关键性能参数在 0~20 N 的



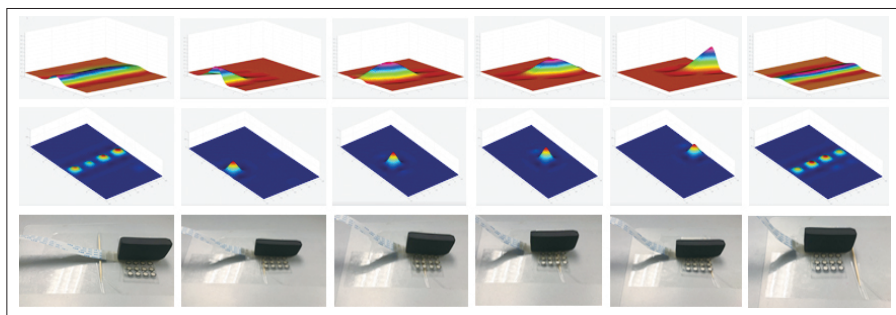
▲图 7 不同指端与接触面作用情况下的有限元受力分析



▲图 8 力-电压拟合图



▲图 9 传感器在不同受力情况下的触觉热力图



▲图10 对微小触觉信号的实时、准确感知

测量范围内,检测精度为0.1 N,平均响应时间为23.8 ms。该传感器可为多模式人机交互应用中的触觉算法提供数据基础,对仿人机器人抓取也有一定的指导意义。

参考文献

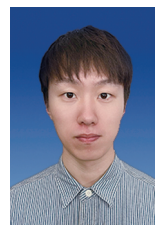
- [1] JACOBSEN S. Design of a multiple degree of freedom, force reflective hand master/slave with a high mobility wrist [EB/OL]. [2021-11-04]. <https://scholarsarchive.byu.edu/facpub/2108/>
- [2] HIRABAYASHI T, AKIZONO J, YAMOMOTO T, et al. Teleoperation of construction machines with haptic information for underwater applications [J]. Automation in construction, 2006, 15(5): 563-570
- [3] MOTOOKA W, NOZAKI T, MIZOGUCHI T, et al. Development of 16-DOF telesurgical force master/slave robot with haptics [C]// IECON 2010 - 36th Annual Conference on IEEE Industrial Electronics Society. IEEE, 2010: 2081-2086. DOI: 10.1109/IECON.2010.5675353
- [4] XU X, SINGH H, LIU Q, et al. A novel energy compensation scheme for quality enhancement in time-delayed teleoperation with multi-DOF haptic data reduction and communication [J]. IEEE transactions on haptics, 3379(99): 1. DOI: 10.1109/TOH.2021.3103379
- [5] KOVACS R, OFEK E, GONZALEZ FRANCO

- M, et al. Haptic PIVOT: on-demand hand-helds in VR [C]//Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology. ACM, 2020: 1046-1059. DOI: 10.1145/3379337.3415854
- [6] ZHOU R G, WANG D X, ZHANG Y R. Haptic rendering of tissue boundary for surgical training [C]//2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. IEEE, 2008: 949-954. DOI: 10.1109/AIM.2008.4601789
- [7] KAPPASSOV Z, CORRALES J A, PERDEREAU V. Tactile sensing in dexterous robot hands-review [J]. Robotics and autonomous systems, 2015, 74: 195-220. DOI: 10.1016/j.robot.2015.07.015
- [8] WETTEL N, SANTOS V J, JOHANSSON R S, et al. Biomimetic tactile sensor array [J]. Advanced robotics, 2008, 22(8): 829-849. DOI: 10.1163/156855308X314533
- [9] TESHIGAWARA S, TSUTSUMI T, SUZUKI Y, et al. High speed and high sensitivity slip sensor for dexterous grasping [J]. Journal of robotics and mechatronics, 2012, 24(2): 298-310. DOI:10.20965/jrm.2012.p0298
- [10] SCHMITZ A, MAIOLINO P, MAGGIALI M, et al. Methods and technologies for the implementation of large-scale robot tactile sensors [J]. IEEE transactions on robotics, 2011, 27(3): 389-400. DOI:10.1109/TRO.2011.2132930
- [11] SCHMIDT P A, MAËL E, WÜRTZ R P. A sensor for dynamic tactile information with applications in human-robot interaction and object exploration [J]. Robotics and autonomous systems, 2006, 54(12): 1005-1014. DOI: 10.1016/j.robot.2006.05.013
- [12] CHOI B, LEE S, CHOI H R, et al. Development of anthropomorphic robot hand with

tactile sensor: SKKU hand II [C]//2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2006: 3779-3784. DOI: 10.1109/IROS.2006.281763

- [13] JOHNSON M K, COLE F, RAJ A, et al. Microgeometry capture using an elastomeric sensor [J]. ACM transactions on graphics, 2011, 30(4): 1-8. DOI: 10.1145/2010324.1964941
- [14] SUNDARAM S, KELLNHOFFER P, LI Y Z, et al. Learning the signatures of the human grasp using a scalable tactile glove [J]. Nature, 2019, 569(7758): 698-702. DOI: 10.1038/s41586-019-1234-z

作者简介



王爽,大连理工大学在读硕士研究生;主要研究领域为机器人触觉抓取。



吴兵,大连理工大学在读硕士研究生;主要研究领域为机器人抓取、视触觉融合。



刘倩(通信作者),大连理工大学计算机学院副教授、IEEE触觉编码国际标准任务组秘书长;主要研究领域为视触觉融合人机交互、多模态感知与信息传输;发表论文20余篇。



触觉通信技术发展的若干思考

Reflections on Development of the Haptic Communication

陈鸣锴 /CHEN Mingkai, 周亮 /ZHOU Liang

(南京邮电大学, 中国 南京 210003)

(Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

DOI: 10.12142/ZTETJ.202106007

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211210.1631.002.html>

网络出版日期: 2021-12-13

收稿日期: 2021-11-15

摘要: 触觉通信技术被视为下一代移动通信中的杀手级应用。通过对触觉通信的发展背景、技术实质、应用服务, 以及技术难题与挑战的分析, 指出了触觉通信技术与通信网络建设相结合的应用价值。认为触觉技术与现有的多媒体技术的融合将给信号处理技术带来技术革新。

关键词: 触觉通信; B5G; 信号处理; 数据传输

Abstract: Haptic communication is considered as one of the most promising technologies in the next generation of mobile communications. The background, technical essence, application services, and the technical problems and challenges in haptic communications are analyzed. And the application value of combining the development of tactile communication technology and communication network construction is indicated. It is believed that the fusion of haptics and multimedia technology will bring a novel innovation to signal processing.

Keywords: haptic communication; B5G; signal processing; data transmission

1 触觉通信的发展

人在出生伊始, 听力、视力等感知体系尚未发育完全, 只有触觉感知神经系统已发育健全, 能够实现神经元与周身皮肤点对点的信息传递, 完成丘脑与外部信息的交互^[1]。因此, 触觉神经感知又被称为人类感知外界多元信息的第一入口。1988年, 朱利叶斯教授和帕塔博蒂安教授从基因层面揭示了人体对温度和机械触觉感知的生物科学原理^[2-3]。该项研究不仅获得了2021年诺贝尔生物学奖, 还引发了触觉生物学的研究热潮。随着触觉生物学研究的不断深入, 以电子仿生皮肤或智能织物为首的电子仿生材料研究也突飞猛进, 这将使由冰冷电子元器件堆砌出来的机器人转型成为具有温度和触感交互的“普通人”。

随着触觉生物学和仿生电子材料

研究的不断发展, 人们开始思考如何让目前仍处于本地孵化的触觉感知与交互技术向中远程迈进。特别是如何在现有以音视频为主的移动通信业务中加入触觉维度, 细化用户服务感受, 多层次化地刺激用户感知, 从而为用户带来更极致的互动体验和更丰富的场景体验。试想一下, 当你的汽车在行驶过程中出现故障时, 技术人员在4S店即可远程进行故障诊断甚至在线维修, 及时为你解决问题; 一位经验丰富的外科医生能给地球另一端的病人实施精准的医疗诊断甚至手术, 使得优质医疗资源的惠及面更广。更加实用的是, 我们在网上下单前就能摸到衣服的质地; 远在千里之外的奶奶在视频聊天时能轻抚孙女的头发, 甚至送上一个有温度的吻。

要实现上述场景, 就意味着需要闭合数据周期, 并利用比特和

字节构建的数据网络精确传递等同于人类触觉的感受。为此, 早在2012年, 德国德累斯顿技术大学GERHARDFETTWEIS教授就提出了“触觉互联网”的概念^[4]。这个概念虽然更多关注的是触觉信号的感知、传输等问题, 不能兼顾海量音视频的承载需求, 但却很好地展现了触觉通信和音视频通信之间的本质区别。这个区别体现在控制层面和传输层面:

(1) 控制层面。由于触觉是通过对环境施加运动, 并通过扭曲或反作用力来感知环境的, 因而触觉控制中的触觉反馈(动觉或振动-触觉)形成全局控制回路, 而非触觉控制中的只能是音频/视频的反馈, 并无控制回路的概念。

(2) 传输层面。触觉需要依靠触觉的即时性和力觉的反馈性, 因此, 触觉信号的传输不仅是双向的, 而且

对时延和可靠性都有着很高要求。否则,会造成触觉临场感系统的不稳定、触觉信息交互应用的服务水平大幅度降低等问题。由此可见,融入触觉数据的移动通信业务虽然能激发 5G 与 6G 的现象级新应用,但同时也对触觉通信的服务保障能力提出了很高要求。这使得触觉通信的服务质量预测和管控成为难题^[5]。

2 触觉通信的技术挑战

从触觉信号的本源、特点和需求出发,同时参照触觉信号的相关研究^[4],我们认为触觉信号具有三大特点:

(1) 多维度。神经生理组织、交互方式等多方面的影响导致触觉感知的信号刻画涉及摩擦力、温度、湿度等诸多特征参量。这些特征参量之间相互耦合,但还缺乏明确的映射模型。

(2) 交互性。一方面,触觉控制中的触觉反馈形成了全局控制回路;另一方面,触觉信号是双向反馈的,因此对传输时延和可靠性都有着非常高的要求。过大的延迟和丢包会给触觉系统带来不稳定的临场感,这使得触觉服务的用户体验直线下降。

(3) 上下文关联性。触觉信号在时间和空间上具有连续性,前后数据呈现出因果变化。其中,力学数据接收时的时空尺度关联特性更不能被忽视。另外,多维数据之间的突变过程也存在联动关系,这使得数据复原与处理技术难上加难。

目前,触觉通信主要面临 3 个主要的技术挑战:

(1) 多维信号采集与交织。受到神经生理组织、交互方式等多方面的影响,触觉感知的机理较为复杂。如图 1 所示,多种触觉属性共存导致触觉信号包含物体材质、软硬度和形变效果、表面粗糙程度、温度和湿度、力矩、力度大小等多维信息。各维数

据中会掺杂大量噪声信号,并会互相干扰。如何完成这样复杂的多维数据的采集和处理,尤其是如何利用精准电信号去衡量、描述多维力或非力属性的变化,是以往音视频编码研究中从未遇到过的情况。

(2) 交互信号传输与协同。依照采样面积、信号采集器件的精度,实时的触觉信号传输速率一般为 20 kbit/s ~ 10 Mbit/s。相较于音视频数据量,虽然触觉数据量有较大幅度的下降,但触觉数据的产生呈现出突发性。例如,人体的突然运动会使得待传输的触觉数据量大幅增加;但当人体或肢体保持静止时,数据传输又呈现出静默性^[5]。不仅如此,人体对触觉感知的灵敏度会远高于音频和视频信号,且触觉信号携带的多维信息要求传输具有协同一致性。因此,高突发、主从交互的触觉信号在传输过程中会对网络提出超低延迟和超高可靠性的要求,即要求延迟小于 10 ms,而可靠性高于 99.99999%^[6]。

(3) 层次数据处理与重构。触觉信号在接收端如何无失真重构是触觉通信中的又一难点。从信号处理的角度看,多层次数据重构需要从同步、

联动、效率等方面出发,而待处理的信号由于触觉信号的多维性呈现出多样化的表达形式,如矩阵形式、常数形式等^[7]。这意味着复原技术需要采用多层次解决方法,将这些多信号表征融合重构出原有的信号质感。

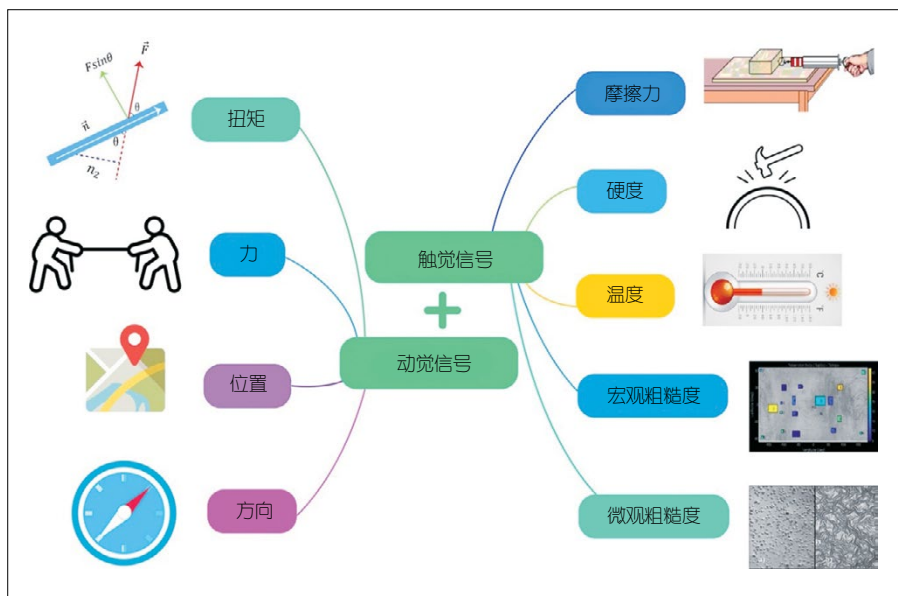
3 触觉通信的技术思考

从通信理论研究的角度出发,基于触觉通信传输的特殊性,我们需要重新思考通信的架构和理论,并围绕触觉通信需求完成网络结构与编码的设计。

3.1 触觉编解码技术

触觉数据有着多维度的特性,因此触觉编码多被设计成为多通道编码。从编码角度看,多通道建模又可划分为以下 3 个主要部分:

(1) 物体材质识别。一般而言,物体材质包括物体内部及表面的材料组成、粗糙程度、软硬程度等。正确识别这些属性是确定触感信号变化基调的前提,这对触觉信号的编码性能有着至关重要的影响。物体材质的识别首先需要组建包含图像、人工标注、



▲图 1 触觉信号主要属性

触感力学信号、敲击声音等大量材质数据的数据库,进而通过确定、优化机器学习或深度学习分类方法,区分触摸到的物体信号。

(2) 力信号采样编码。在触觉编码的过程中,力信号的电平转化与编码是重要环节,且与传统图像、语音数据采样有着本质的区别。一方面,力学信号的频谱范围是 $0 \sim 1.5 \text{ kHz}$, 采样频率一般为 $4 \sim 5 \text{ kHz}$, 这可能会导致触觉数据量较大。因此,我们需要对采集到触觉数据进行压缩。那么,如何设计压缩算法才能权衡压缩率与抗干扰性成为核心问题之一。另一方面,由于人体触感的着力点位置、角度、移动力道、速度等不断变化,摩擦力信号也随之变化,这对力学传感器的灵敏度和精度也提出更高要求^[7]。

(3) 恒定参量编码。虽然温度、湿度等恒定参量的感知与采样技术相对成熟,但信号在触觉编码过程中的突发变化要求恒定参量及时响应,且需要与其他信号的变化保持一致,避免带给用户迟滞感。这样一来,如何在数据突变态与静默态之间切换、如何确保多维信号编译码的同频同步,以及如何确保此类参量解码后实现小型化和无感化的重构终端设计,均是研究难题。

3.2 触觉通信传输技术

面对实时触觉信号的低延迟、低丢包率、低抖动,以及多维信号的协同一致性问题,人们希望借助 5G 和未来 6G 中的新型通信技术来推动触觉信息交互应用的发展。

(1) 面向网络切片的虚拟资源管理技术。面向网络切片的虚拟资源管理技术重塑了面向网络切片的虚拟资源管理体系化框架,能够在确保原有海量音视频数据流质量的前提下,为突发的小数据触觉信号流提供低延迟、

高可靠的稳定链路。以保障跨媒体业务的按需传输和资源高效利用为目的,该技术对虚拟化网络功能(VNF)进行编排、部署和调度:在编排阶段,提出面向兼容性和聚类需求的 VNF 实例聚类方法,以实现簇数目未知情况下的多维、高阶 VNF 实例聚类;在部署阶段,用服务功能链部署问题替换具有动态资源需求的 VNF 部署问题,以降低业务处理和传输时延,实现业务和部署同步;在调度阶段,提出资源和业务时延等多重约束条件下 VNF 最优调度,以将虚拟机选择和调度相分离^[8]。

(2) 基于通信、计算与存储协同的计算卸载与边缘缓存。对于远距离跨区域的触觉信号传输,为了确保主从设备之间交互信息的透明化、无感化,可发挥通信、计算与存储的协同优势,需要寻求路径上的边缘计算服务器来完成代理加速辅助^[7]。一方面,以最小化任务时延、最大化传输可靠性为出发点,通过研究要不要卸载、卸载多少、卸载什么、卸载到哪里等问题,提出基于移动性管理的协作式卸载策略;另一方面,以提升数据缓存利用率为出发点,将被动式缓存和主动式缓存相融合,提出一种混合式触觉数据缓存与更新方法^[9]。

3.3 触觉信号复原技术

触觉信号如何复原是提升用户沉浸式体验的关键。受触觉反馈方式和系统的限制,触觉交互应用开发成本高,而其增强效果大多停留在振动碰撞模拟阶段,触觉再现技术带来的交互体验和效率优化价值难以评判。因此,触觉复原的评价体系需要继续沿用图像中峰值信噪比(PSNR)或像素准确度判决等机制,但这些机制是否能达到触觉信号的重构效果,还有待考究。同时,人们开始尝试从生物神

经学或是行为心理学的角度对多维数据信号的融合效果进行深层次研究。

此外,网络的不确定性使得触觉信号到用户端的传输过程中,数据包随机出现甚至丢失。特别是当触觉信号不完整时,接收设备如何恢复将成为一个难题。为此,一般可利用触觉信号的上下文、时序因果等特性,来完成矩阵填充、上下文关联、计算差值补全等操作。对于大面积的力学信号缺失问题,我们还可以采用卷积神经网络(CNN)或生成式对抗网络(GAN)等人工智能算法来进行数据补全,并采用迁移学习和联邦学习等方法对数据进行协同的融合复原。同时,根据多维触觉信号的工作原理,可对触觉信号进行超分辨率增强、噪声消除处理、触觉信号切割、语义分析与处理等更加先进的信号处理与重构操作,并可与行为心理学、机器人工程学、柔性材料等形成学科交叉,共同推动触觉信号的处理和复现技术的工程示范落地。

为改善触觉信号单一模态信息的局限性,我们可通过挖掘音视触不同模态信号在数据与语义上的关联性,构建音视触多模态信号特征数据库和跨模态语义匹配数据库;通过挖掘本地信息的语义特征与上述数据库中信息的关联关系,提出多模态信号的超分辨率重建方法^[8];通过学习基本图形和一般图像的特征,提出基于图形和图像的多参量触觉渲染方法,解决音视触信号跨模态中不连续的问题,提升触觉展现的沉浸感。

4 触觉通信的应用场景

触觉技术将远距离感知控制增加了一个新的维度。触觉技术赋能的应用体系将是未来信息通信技术(ICT)领域一个重要研究方向。从应用场景的角度来看,触觉通信应用主要可分

为虚拟场景与现实场景,即在真实环境下的远程精准控制技术、在虚拟场景下的内容感知技术,如图2所示。

4.1 现实场景中的触觉应用

在真实场景中,包含触觉反馈的远距离控制技术可以实现工业远程精细化控制、高精度远程医疗手术、远程教育、电子商务等服务应用。

(1) 工业控制。虽然基于视频、音频信号的远距离工业控制应用已有一定规模的应用,例如,矿山开采、危险区域探索等对于操作精确度和时延要求不高的工程。这类工业控制不存在触觉反馈,因而交互形式单一。如果控制终端能够更加类人化,并能加入触觉信号,那么不仅远程终端控制的过程更加人性化,整个劳动效率可以得到提高,同时工业成本有所降低。

(2) 网络诊疗。2019年初,华为通过5G网络实现了首例远程的动物外科手术实验,但手术完全依赖于视频和音频信号,医生无法获取真实的手术触觉感受,从而花费了几倍于线下外科手术的时间。如果在医生端加入动触觉信号的捕捉和处理,加速手术交互感知的速度,就可以实现真正的沉浸式远程诊疗。如果在中医的诊疗过程中,加入动觉和触觉捕捉,那么针灸、推拿等中医疗法也有望通过网络实现。

(3) 远程教育。触觉信号也能应用于教育产业与电子商务。例如,培训过程中如果使用触觉信号,将会使学生们得到更加沉浸式的体验:学生可以设身处地地感受书法或绘画中笔锋、笔触的力道变化,无须老师们线下手把手教学,就能心领神会。

(4) 电子商务。在网络购物的环境下,消费者如果能直接触摸到所购买的商品,直观感受到商品的性能,便能减少商品的退换货率,间接提升产



▲图2 触觉通信在现实/虚拟场景中的应用

品销量。

4.2 虚拟场景中的触觉应用

在虚拟空间中,触觉信号的引入将给虚拟现实增强、数字孪生、网络游戏等产业带来革命性的变化。特别地,触觉信号技术可与3D建模、图形渲染、虚拟仿真等技术结合,赋予虚拟建模以生命,让其感知起来更加具有真实感^[10]。我们认为目前触觉通信可应用于以下几类虚拟场景中:

(1) 网络游戏产业。传统的方法主要通过算力和渲染效果的迭代来提升游戏玩家的视听享受。游戏厂商也尝试过加入触觉来增加游戏的真实感,但仅依靠游戏手柄中马达震动模拟出的部分触觉反馈是远远不够的。一旦加入触感的真实模拟,可以让玩家更身临其境地感受到游戏中的虚拟世界,这会让用户获得更多的代入感、沉浸感。

(2) 虚拟现实增强。目前的虚拟现实增强技术所面临的最大问题就是用户对虚拟后的图形、图像的感知不

够真实和直观。为了让用户尽最大程度地感受虚拟图形、图像,添加物体的各项属性必不可少,如材质感受、回弹、击打感,乃至气味、味道等。这样能够多维度增加人对虚拟世界探索的代入感。

(3) 元宇宙建设。2021年10月,Facebook重组改名为Metauniverse,致力于元宇宙建设,力图打造现实世界孪生的网络虚拟世界,建立虚拟现实交互的移动社交网络。元宇宙的建设离不开触觉通信、区块链、云计算、虚拟现实等技术的强力支撑。这些技术可以将用户与虚拟角色都将囊括进一个特征关联永续的、广覆盖的虚拟现实世界中。

5 结束语

触觉通信技术正处于发展的萌芽期。由于能显著提升诸如远程医疗、“全感知”实战实训、临场感办公等新范式服务的沉浸式体验,触觉通信技术被视为5G和6G时代的杀手级应用。不仅如此,随着“数字中国”建设的

加快,超前布局以触觉信号为代表的新业态战略性技术已成必然。由于多维、交互、关联等特点,触觉信号要求通信系统更加灵活、智能、高效与开放。随着触觉通信技术研究的深入,编码技术、通信传输技术、信号复原技术等方面的技术革新将会是颠覆式的,并将会开启通信智能的新时代。

参考文献

- [1] ANTÓN-BOLAÑOS N, SEMPÉRÉ-FERRÁNDEZ A, GUILLAMÓN-VIVANCOS T, et al. Prenatal activity from thalamic neurons governs the emergence of functional cortical maps in mice [J]. Science, 2019, 364(6444): 987–990. DOI: 10.1126/science.aav7617
- [2] JULIUS D, MACDERMOTT A B, AXEL R, et al. Molecular characterization of a functional cDNA encoding the serotonin 1c receptor [J]. Science, 1988, 241(4865): 558–564. DOI: 10.1126/science.3399891
- [3] COSTE B, MATHUR J, SCHMIDT M, et al. Piezo1 and Piezo2 are essential components of distinct mechanically activated cation channels [J]. Science, 2010, 330(6000): 55–60. DOI: 10.1126/science.1193270
- [4] STEINBACH E, STRESE M, EID M, et al. Haptic codecs for the tactile Internet [J]. Proceedings of the IEEE, 2019, 107(2): 447–470. DOI: 10.1109/JPROC.2018.2867835
- [5] SIMSEK M, ALJAZ A, DOHLER M, et al. 5G-enabled tactile Internet [J]. IEEE journal on selected areas in communications, 2016, 34(3): 460–473. DOI: 10.1109/JSAC.2016.2525398
- [6] MUMTAZ S, AI B, AL-DULAIMI A, et al. Guest editorial 5G tactile Internet: an application for industrial automation [J]. IEEE transactions on industrial informatics, 2019, 15(5): 2992–2994. DOI: 10.1109/TII.2019.2904202
- [7] ANTONAKOGLU K, XU X, STEINBACH E, et al. Toward haptic communications over the 5G tactile Internet [J]. IEEE communications surveys & tutorials, 2018, 20(4): 3034–3059. DOI: 10.1109/COMST.2018.2851452
- [8] ZHOU L, WU D, CHEN J X, et al. Cross-modal collaborative communications [J]. IEEE wireless communications, 2020, 27(2): 112–117. DOI: 10.1109/MWC.001.1900201
- [9] ZHOU L, WU D, WEI X, et al. Cross-modal stream scheduling for eHealth [J]. IEEE journal on selected areas in communications, 2021, 39(2): 426–437. DOI: 10.1109/JSAC.2020.3021543
- [10] WEI X, ZHOU L. AI-enabled cross-modal communications [J]. IEEE wireless communications, 2021, 28(4): 182–189. DOI: 10.1109/MWC.001.2000448

作者简介



陈鸣镭, 南京邮电大学通信与信息工程学院讲师; 主要从事无线多媒体通信、无线资源分配、触觉通信、跨模态通信等科学研究工作。



周亮, 南京邮电大学教授、博士生导师、通信与信息工程学院院长, 教育部宽带无线通信与传感网技术重点实验室主任; 先后获教育部“长江学者奖励计划”特聘教授、中共中央组织部“海外高层次人才专家”等荣誉称号, 获国家自然科学基金委员会“优秀青年基金”资助; 作为项目负责人和主要完成人, 主持并参与了多项国家级重点科技攻关项目; 发表论文多篇, 拥有授权专利 25 项。



800 Gbit/s 光模块技术及应用

Technologies and Applications of 800 Gbit/s Optical Module

王会涛 /WANG Huitao^{1,2}
张平化 /ZHANG Pinghua^{1,2}
苏展 /SU Zhan^{2,3}

(1. 中兴光电子有限公司, 中国 南京 210012;
2. 移动网络和移动多媒体技术国家重点实验室, 中国 深圳 518057;
3. 中兴通讯股份有限公司, 中国 深圳 518057)
(1. ZTE Photonics Technology Co., Ltd., Nanjing 210012, China;
2. State Key Laboratory of Mobile Network and Mobile Multimedia, Shenzhen 518057, China;
3. ZTE Corporation, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202106008

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211209.0931.002.html>

网络出版日期: 2021-12-12

收稿日期: 2021-11-10

摘要: 800 Gbit/s 光模块场景包括 SR (100 m 场景)、DR/FR/LR (500 m/2 km/10 km 场景)、ER/ZR (40 km/80 km 场景)。结合商用规律和技术成熟度两个维度, 提出了单模方案下沉、单波 200 Gbit/s 来临、相干下沉 3 个趋势判断。预测了 800 Gbit/s 主流模块接口形态, 具体包括基于直接调制激光器 (DML) / 硅光 (SiPh) 的 800 Gbit/s SR8、基于电吸收调制激光器 (EML) / SiPh 的 800 Gbit/s DR4、基于 EML 的 800 Gbit/s FR4、基于 EML 的 800 Gbit/s LR8 和基于相干的 800 Gbit/s ER/ZR。

关键词: 单模方案下沉; 单波 200 Gbit/s; 相干下沉

Abstract: Applications of 800 Gbit/s optical module contain SR (100 m scenario), DR/FR/LR (500 m/2 km/10 km scenario), and ER/ZR (40 km/80 km scenario). Based on the commercial law and technology maturity, three trends of single-mode sinking, single wavelength 200 Gbit/s coming, and coherent sinking are proposed. The interface forms of 800 Gbit/s mainstream modules are predicted, including 800 Gbit/s SR8 based on direct modulation laser (DML)/silicon photonics (SiPh), 800 Gbit/s DR4 based on electro-absorption modulation laser (EML)/SiPh, 800 Gbit/s FR4 based on EML, 800 Gbit/s LR8 based on EML, and 800 Gbit/s ER/ZR based on coherence.

Keywords: single-mode sinking; single wavelength 200 Gbit/s; coherent sinking

1 市场预测和应用场景

4K 虚拟现实 (VR)、物联网、云计算等新业务的出现, 对网络的带宽、并发率和实时性提出了更高的要求。根据 Omdia 预测^[1], 未来几年随着带宽需求的不断提升, 虽然 100、200、400 Gbit/s 光模块仍将保有最大的市场占有率, 但是 800 Gbit/s 光模块将在 2023 年实现商用, 在 2025 年实现规模部署。

如图 1 所示, 根据 800 GE 网络结构, 对于架顶交换机 (TOR) 到 Leaf 交换机的连接距离, 短的有几十米, 长的可能有几百米。在这部分连接上, 大型互联网公司普遍采用 100 Gbit/s 速率的连接技术, 并从 2021 年开始逐步换代到 200 Gbit/s 或 400 Gbit/s 的速率技术。一些领先的公司会在 2023

年开始试用 800 Gbit/s 技术。Leaf 到 Spine 交换机的连接, 或者 Spine 交换机到核心路由器的连接, 可能会解决一个园区内部或者相邻园区之间的互联问题。这种连接距离会达到 2 km, 甚至 10 km。接口速率也将从 2021 年开始由 100 Gbit/s 逐步换代到 200 Gbit/s 或 400 Gbit/s 速率。一些公司会在 2023 年开始试用 800 Gbit/s 技术。数据中心互联 (DCI) 一般是指相邻几个数据中心之间负载均衡或容灾备份的连接, 这种连接距离可能长达几十公里。对于这么远的距离, 由于光纤资源比较珍贵, 人们主要采用密集波分复用加相干通信的方式以尽可能复用光纤资源。我们把 800 Gbit/s 光模块的应用场景分为 SR (100 m 场景)、DR/FR/LR (500 m/2 km/10 km 场景), 以及 ER/ZR (40 km/80 km 场景)。

2 技术方案

2.1 方案概述

800 Gbit/s 技术方案演进包括 3 代。

第 1 代 为 8 光 8 电: 光接口 8×100 Gbit/s, 电接口 8×100 Gbit/s, 商用时间为 2021 年;

第 2 代 为 4 光 8 电: 光接口 4×200 Gbit/s, 电接口 8×100 Gbit/s, 商用时间预计为 2024 年;

第 3 代 为 4 光 4 电: 光接口 4×200 Gbit/s, 电接口 8×100 Gbit/s, 商用时间预计为 2026 年。

从长期来看 (5 年内), 光 / 电单信道 200 Gbit/s 技术将会实现普及; 从短期来看 (3 年内), 由于单信道 200 Gbit/s 的光电芯片器件和均衡技术目前尚不成熟, 产业界仍需要时间来突破相关技术瓶颈。

2.1.1 电接口及封装

从 100 Gbit/s 直调直检光模块发展来看,当电接口单通道速率与光接口单通道速率相同时,光模块的架构将达到最佳状态,并具有低功耗、低成本等优势。单通道 100 Gbit/s 电接口将是 8×100 Gbit/s 光模块的理想电接口,单通道 200 Gbit/s 电接口将会是 4×200 Gbit/s 光模块的理想电接口。在封装方面,800 Gbit/s 光模块可能存在双密度四通道小型可插拔(QSFP-DD800)、八通道小型可插拔(OSFP)等不同形式。由于存在模块内走线和连接器损耗等因素,基于 200 Gbit/s 电接口的可插拔光模块仍面临诸多挑战。

2.1.2 光接口

800 Gbit/s 光模块光接口架构主要有 3 种,如图 2 所示^[3]。

(1) 8×100 Gbit/s 4 电平脉冲幅度调制(PAM4)光模块: PAM4 收发器以 53 Gbd 运行,使用 8 对数模转换器(DAC)和模数转换器(ADC)、8 个激光器、8 对光收发器,以及 1 对 8 通道粗波分复用器(CWDM)或基于以太网通道的波分复用(LAN-WDM)复用器和解复用器(SR/DR 应用场景不需要)。

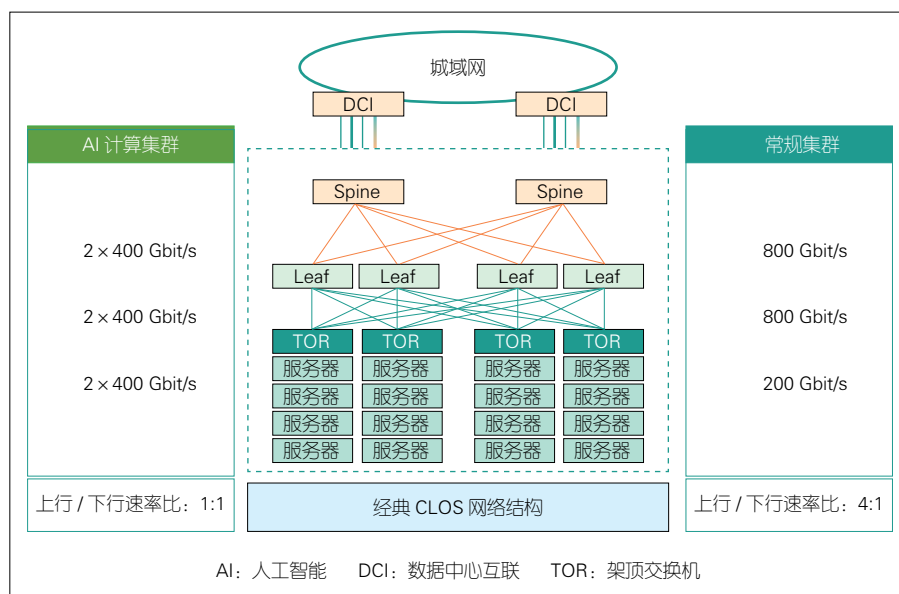
(2) 4×200 Gbit/s PAM4 光模块:

PAM4 收发器以 106 Gbd 运行,使用 4 对 DAC 和 ADC、4 对光收发器(包括 4 个激光器),以及 1 对 4 通道 CWDM 或 LAN-WDM(取决于光纤色散损失)复用器和解复用器(SR/DR 应用场景不需要)。

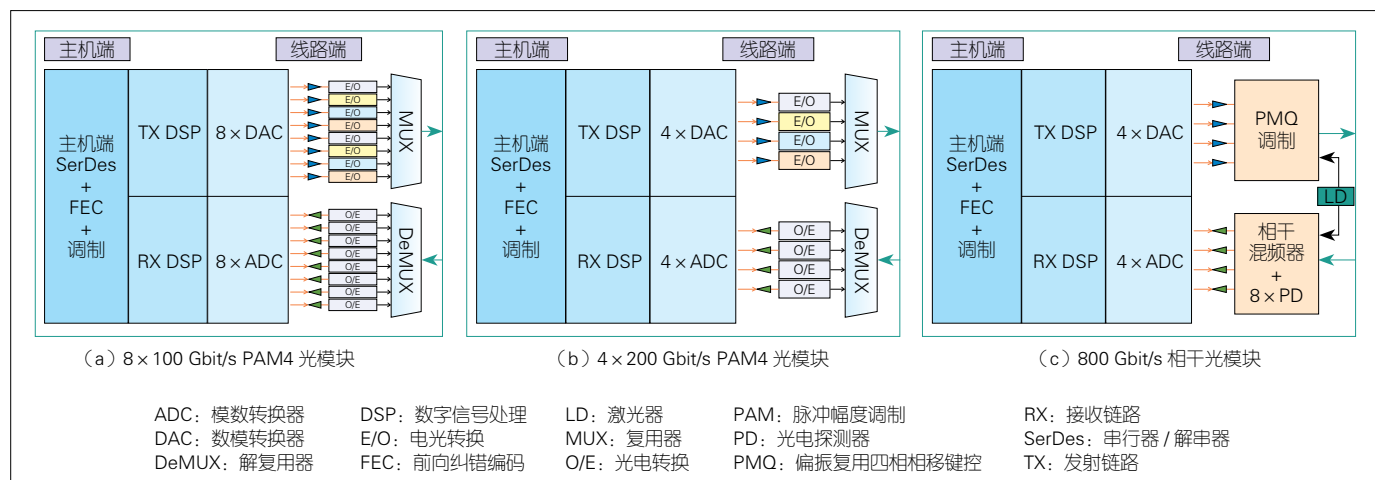
(3) 800 Gbit/s 相干光模块: 在双极化十六正交振幅调制(16QAM)下以 128 Gbd 运行。它使用 4 对 DAC 和 ADC、1 个激光器和 1 对光收发器,可以在数据中心相干光模块中使用固定波长激光器,以降低成本和功耗。

8×100 Gbit/s 直调直检方案可利

用已有技术架构,相关技术和标准比较成熟,供应链也较为完善。在 SR 场景下,垂直腔面发射激光器(VCSEL) 100 Gbit/s 技术面临挑战。提升多模方案性能和降低多模光纤成本,将成为该技术持续演进的关键因素。以硅光(SiPh)和直接调制激光器(DML)为代表的单模技术迅速发展。其中, SiPh 技术发展更为迅速,未来有望在 100 m 及以下传输距离的应用场景中与多模方案展开竞争。在 DR/FR 场景下,存在电吸收调制激光器(EML)、DML 和 SiPh 3 种方案。在 LR 场景下,



▲图 1 800 GE 网络架构图^[2]



▲图 2 3 种 800 Gbit/s 光模块光接口架构

有基于粗波分复用 (CWDM)、细波分复用 (LWDM) 和窄带细波分复用 (nLWDM) 的 800 Gbit/s LR8 方案, 这些方案目前仍处于研究阶段。在波长选择上, 由于 O 波段边缘波长的色散较大, LWDM8 在色散代价方面优于 CWDM8。目前, 10 km 及以上距离的直调直检方案主要面临“最坏情况”色散和狭窄的色散容限匹配挑战。构建新的波长体系并压缩多通道波长范围, 可使最坏情况色散相应变窄, 从而简化数字信号处理 (DSP) 设计, 降低理论功耗。例如, 8 × 100 Gbit/s PAM4 直调直检方案采用 800 GHz 间隔的 LWDM 方案时色散受限距离约为 10 km, 采用 400 GHz 间隔的 nLWDM 方案时色散受限距离可拓展至 20 km, 采用 200 GHz 间隔的 nLWDM 时色散受限距离可进一步拓展至 40 km。同时, 压缩零色散点分布或飘移范围, 缩小对应的色散范围, 也是解决方案之一。然而, 由于不同厂家光纤产品零色散点的分布并不统一, 大范围压缩仍存在难度。

对于 4 × 200 Gbit/s 直调直检方案, 单通道 200 Gbit/s 沿用 PAM4 调制码型, 可利用相对成熟的 PAM4 产业基础条件 (但也不排除新调制码型的可能性)。在 4 × 200 Gbit/s DR 和 FR 应用场景中, 目前有 4 路单模并行 (PSM4) 和 CWDM4 两种技术方案。这两种方案目前仍面临较多挑战, 需要进一步展开研究。对于 LR 应用场景, 有基于 CWDM、LWDM, 以及 nLWDM 的 800 Gbit/s LR4 方案。这些方案目前仍处于研究讨论阶段, 需要高带宽光电芯片器件、更强的均衡技术和前向纠错 (FEC), 以确保纠后的误码率 (BER)。

800 Gbit/s 相干光模块的器件带宽需要极大提升, 同时器件设计难以一步到位地实现带宽翻倍。基于 96 GBd 器件实现的 800 Gbit/s 相干光模块须采

用更高阶的调制码型。这种方法存在光信噪比 (OSNR) 低、传输距离和应用场景受限等缺点。基于 128 GBd 的双偏振 (DP) -16QAM 相干光模块, 拥有更好的 OSNR 和传输能力, 将成为 800 Gbit/s 相干的主流实现方案。

2.1.3 FEC

FEC 总体分为 3 类^[4]: 端到端 FEC、嵌套级联 FEC 和分段式 FEC。

业界普遍认为, 8 × 100 Gbit/s 直调直检方案在 40 km 以内传输距离的应用, 可由端到端 KP4 FEC 来实现。而 40 km 传输距离则有可能采用更强的 FEC。

4 × 200 Gbit/s 直调直检方案因速率更高, 须引入新的 BER 标准、新的 FEC 编码方式和更复杂的均衡器。IEEE 802.3 B400G SG (电气与电子工程师协会 802.3 后 400 Gbit/s 研究组)、800G Pluggable MSA (800 Gbit/s 可插拔多源协议) 工作组已开展相关讨论。级联方式可能成为 4 × 200 Gbit/s 直调直检方案的新路径。这种方式既保留了 KP4 FEC, 避免了主芯片集成新 FEC 所带来的额外成本, 又可通过光模块中轻量化、易实现的 FEC 为光链路提供额外保护, 降低解码带来的功耗和时延。在纠错性能上, KP4+BCH (144,136) 等多种级联内码均可在纠前误码率 $1 \sim 2E-3$ 的区间基础上, 使纠后范围小于 $1E-13$ 。同时, 目前对 800 Gbit/s 最强烈的诉求来自 OTT (指互联网运营商) 数据中心与高性能计算等场景应用。这些场景对时延敏感度要求较高。低时延 FEC 算法成为 800 Gbit/s 的核心诉求之一。

800 Gbit/s 相干包括 800 Gbit/s LR、800 Gbit/s ZR 两种。因此, 我们需针对不同应用场景进行 FEC 算法的设计。(1) 800LR 场景需要 10 km 园区网络, 对时延和功耗要求较高。目

前, 解决方案有 KP4+eHamming/eBCH 级联、空间耦合码 FEC(XR-FEC)、集群型 FEC (CFEC)、Zipper、轻量化开放 FEC (OFEC) 等。其中, 级联方案与 4 × 200 Gbit/s 直调直检级联方案有共通之处。两种路径的相通可进一步降低主芯片复杂度。(2) 800ZR 场景主要应用于 DCI, 是光互联论坛 (OIF) 400ZR 标准的延续。800ZR 采用 DP-16QAM 调制格式, 使 CFEC 纠错能力受到一定挑战, 可能需要多级编码 (MLC)、OFEC 等纠错能力更强的 FEC 方案。

2.1.4 均衡技术^[5]

为实现单通道 200 Gbit/s 的数据传输速率, 光电芯片都要进行性能升级, 例如需要 200 Gbit/s 的 SerDes、带宽高于 50 GHz 的光电芯片和器件等。从目前技术研究报道来看, 带宽高于 50 GHz 的光芯片相对容易实现。如何在带宽提升的前提下保证其他指标性能的最优是需要考虑的重点。目前, Driver 和 TIA 电芯片带宽还不能满足速率需求, 还需要具备均衡能力。在提升自身带宽的同时, 这些电芯片需要实现系统级信号优化的效果。高效的均衡技术可以更大限度地放宽系统对光电器件带宽的要求。

常见的均衡技术包括前馈均衡 (FFE)、判决反馈均衡 (DFE) 和最大似然序列均衡 (MLSE) 等。其中, 由于实现方式简单, FFE 被广泛应用于 SerDes 系统和光信号 DSP (oDSP) 芯片中。为了缓解单通道 200 Gbit/s 对光电器件带宽的需求, 一方面可以在发端采用 FFE 预均衡技术以补偿发射端器件带宽, 另一方面通过在 oDSP 施加更加强效的均衡技术来缓解带宽限制对系统性能的劣化影响。对于单波 100 Gbit/s 标准中采用的 5 抽头 FFE 均衡, 当速率提高到 200 Gbit/s 时, FFE

抽头数将增加。虽然更高性能的 MLSE 均衡算法也可作为解决方案,但 MLSE 实现方式更为复杂,所需运算量也很大,这将增加 oDSP 的功耗。

2.1.5 标准化进展

全球多个标准化组织竞相开展 800 Gbit/s 的标准化工作。IEEE 802.3、800G Pluggable MSA、100G Lambda MSA 和国际光电委员会 (IPEC) 等均已启动 800 Gbit/s 相关规范的制订工作,对 800 Gbit/s 光模块的应用场景、接口规格等进行了定义。800G Pluggable MSA 已先后发布面向低成本、100 m 传输距离需求的 8 × 100 Gbit/s PSM8 以及面向 2 km 传输距离需求的 4 × 200 Gbit/s FR4 规范。QSFP-DD800 MSA 对 QSFP-DD 封装向 800 Gbit/s 的演进进行定义,并将其合并至 QSFP-DD MSA 新发布的 QSFP-DD SPEC 6.0 文件中。

2.2 800 Gbit/s SR 场景技术方案

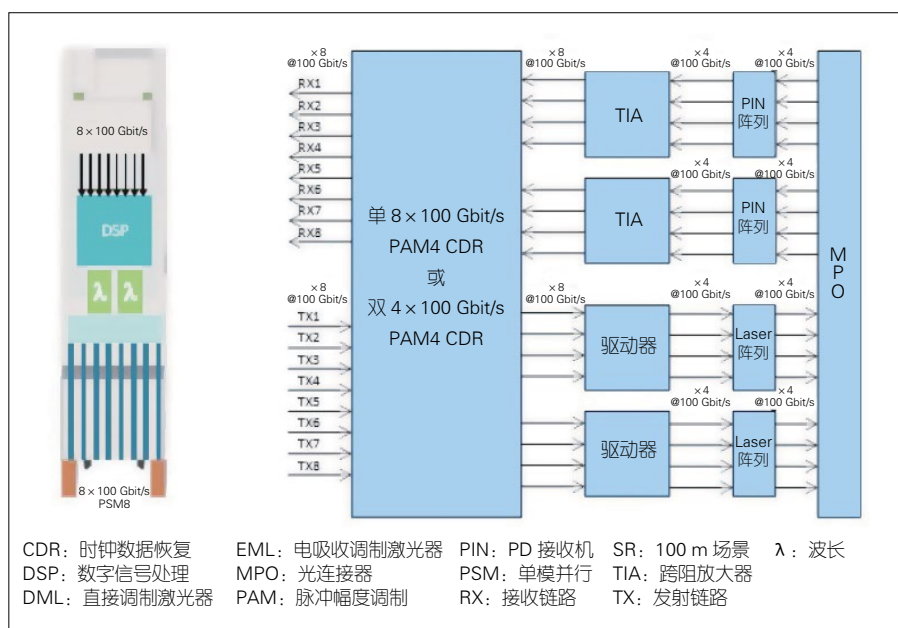
针对 800 Gbit/s 短距需求,受传统多模光纤带宽限制,高波特率信号的多模传输距离被进一步压缩到 50 m 以内,即使使用新型 OM4/OM5 多模光纤也很难实现 100 m 传输距离。根据 800G Pluggable MSA 工作组的定义,800 Gbit/s 在 100 m 传输上不再采用基于 VCSEL 的多模方案,而是采用并行单模光纤 (SMF) 传输 PSM8。一般可采用调制格式为 PAM4 (内置 DSP 芯片) 的方式来满足 800G SR 的场景需求。目前,800 Gbit/s SR 场景技术方案具体包括基于 DML/EML 的方案和基于 SiPh 的方案,如图 3 和图 4 所示。

800 Gbit/s SR8 DML/EML 方案: 采用 8 × 100 Gbit/s DSP (未来或采用模拟时钟数据恢复单元)、同一波长 DML/EML 光芯片,收发两端各使用 8 根光纤 (PSM8 并行单模 8 通道),并且采用 24 芯或 16 芯 MPO 接头。

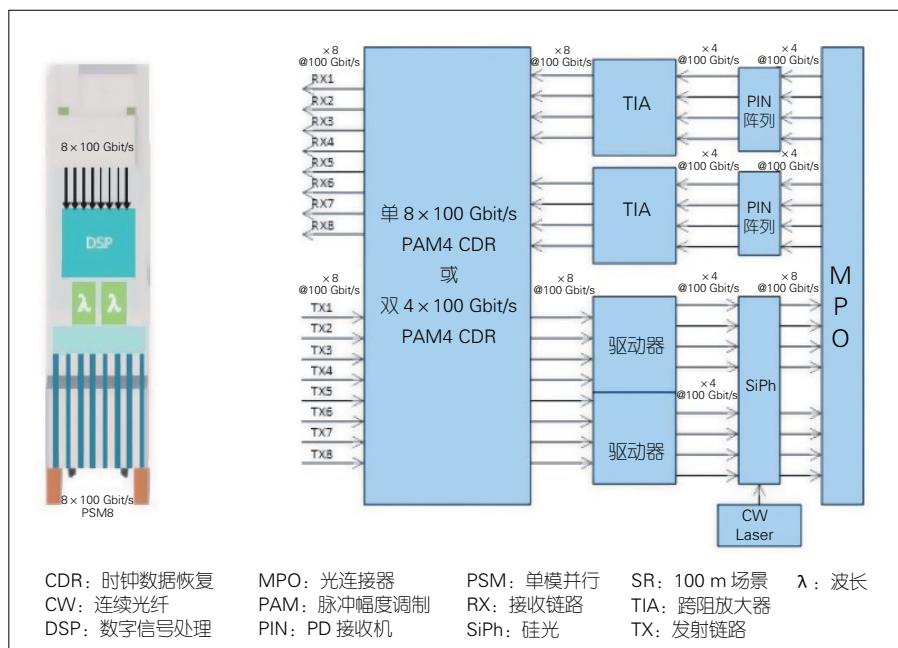
800 Gbit/s SR8 SiPh 方案: 采用 8 × SiPh 马赫·曾德尔调制器 (MZ) / 连续光纤 (CW) 激光器 (硅光作为发射端,同时调制器和光源分离),可以实现并行多路的共享光源架构。若插损控制得当,使用 1 ~ 2 个光源实现 8 路并行可使系统具有很好的成本优势。

2.3 800 Gbit/s DR/FR 场景技术方案

针对 800 Gbit/s 500 m 的互联需求,8 × 100 Gbit/s 方案中的 SiPh 方案相比于 400 Gbit/s 的 DR4 (SiPh),在降低成本方面比较有限。4 × 200 Gbit/s 方案具有更低的成本优势,同时 100 Gbd 器件的良率还有待考察。8 通道并行方案可能仍会被作为前期 MSA 方案。



▲图 3 8 × 100 Gbit/s SR8 DML/EML



▲图 4 8 × 100 Gbit/s PSM8 SiPh

如图5所示, 800 Gbit/s DR4(EML/SiPh)方案采用 4×200 Gbit/s DSP。光芯片采用 $4 \times$ EML/SiPh, 为同一波长。因带宽发展受限, 方案不采用DML。收发端各用4根光纤(PSM4并行单模4通道), 均为同一波长, 并采用12芯MPO接头。

在800 Gbit/s 2 km(FR) 互联需求方面, 基于单通道200 Gbit/s的PAM4技术是光强度调制和直接检测互连的下一代技术的代表, 将成为4通道800 Gbit/s光连接的基础。当速率从100 Gbit/s升到200 Gbit/s时, 波特率会翻倍, 灵敏度会恶化约3 dB, 因此, 需要更强大的FEC来保持接收器较高的灵敏度(-5 dBm)。

800 Gbit/s FR4 EML方案采用 4×200 Gbit/s DSP、CWDM4 EML光芯片。收发两端采用合分波器, 各用1根光纤(CWDM4), 同时采用双朗讯(LC)接头, 如图6所示。

2.4 800 Gbit/s LR/ER/ZR 场景技术方案

在800 Gbit/s 10 km 互联需求方面, 考虑到色散容限, 目前业界提出4种解决方案: 800 Gbit/s LWDM8/nLWDM8、800 Gbit/s LWDM4/nLWDM4、800 Gbit/s 自零差检测(SHD)相干、800 Gbit/s 相干。

如图7所示, 800 Gbit/s LR8方案采用 8×100 Gbit/s DSP、LWDM8 EML光芯片。收发两端采用合分波器, 各用1根光纤(LWDM8), 并采用双LC接头。

800 Gbit/s LR4方案主要采用 4×200 Gbit/s DSP、 $4 \times$ EML LWDM4波长。接收侧使用200 Gbit/s PAM4波导雪崩式光电二极管(APD)。如图8所示收发两端采用合分波器, 各用1根光纤(LWDM4), 同时采用双LC接头。

在波长的选择方面, 目前有LWDM4(800 GHz波长间隔)和nLWDM4(400 GHz波长间隔)两种方案。nLWDM4方案具备色散代价更小、DSP功耗和复杂度更低的优势, 但需要开发新的EML芯片。

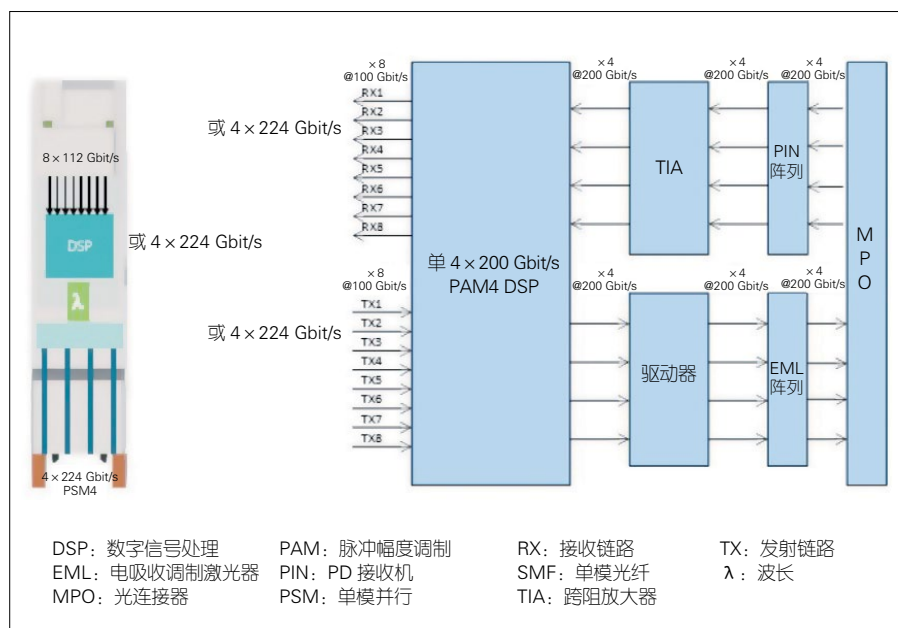
针对800 Gbit/s 40 km/80 km场景的互联需求, 业界采用的解决方案为800 Gbit/s 相干。该方案采用专用相干

DSP、128 Gbd IC-TROSA, 并采用双LC接头。

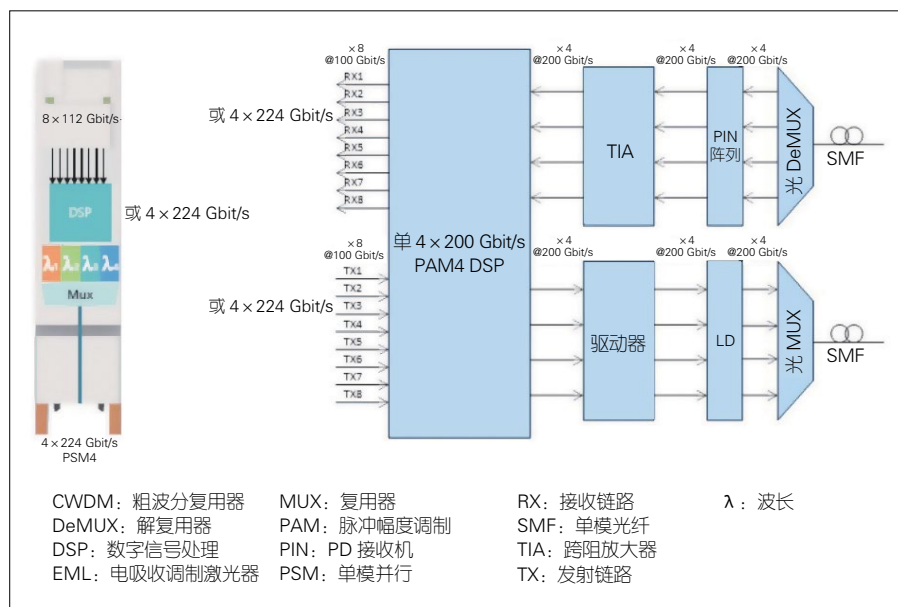
3 800 Gbit/s 的发展趋势及商用策略

800 Gbit/s的发展趋势包括3个方面: 单模下沉、单波200 Gbit/s来临、相干下沉。

(1) 单模下沉。受限于多模光纤



▲图5 4×200 Gbit/s PSM4 EML/SiPh



▲图6 4×200 Gbit/s CWDM4 EML

的带宽, 100 Gbit/s PAM4 VCSEL+ 多模光纤的传输距离为 50 m。如果采用 OM5 光纤, 那么系统成本就会增加。在未来, 单模光接口方案下沉是必然趋势, 而这将利好 SiPh 技术。

(2) 单波 200 Gbit/s 来临。虽然 112 Gbd EML 技术发展较快, 目前已经有样机诞生, 但是 55 GHz 的带宽资源略显不足。200 Gbit/s PAM4 速率等

级的 SiPh 调制器和硅基薄膜铌酸锂的应用前景非常广阔。

(3) 相干下沉。随着传输速率的提升, 相干技术方案在 80 km 传输距离的基础上将进一步向 40、20、10 km 等更短距离拓展应用。同时, 非相干方案也在努力向长距离应用拓展。因此, 两种方案在应用时会出现“相遇”的情况。两种方案的相遇将与技术方

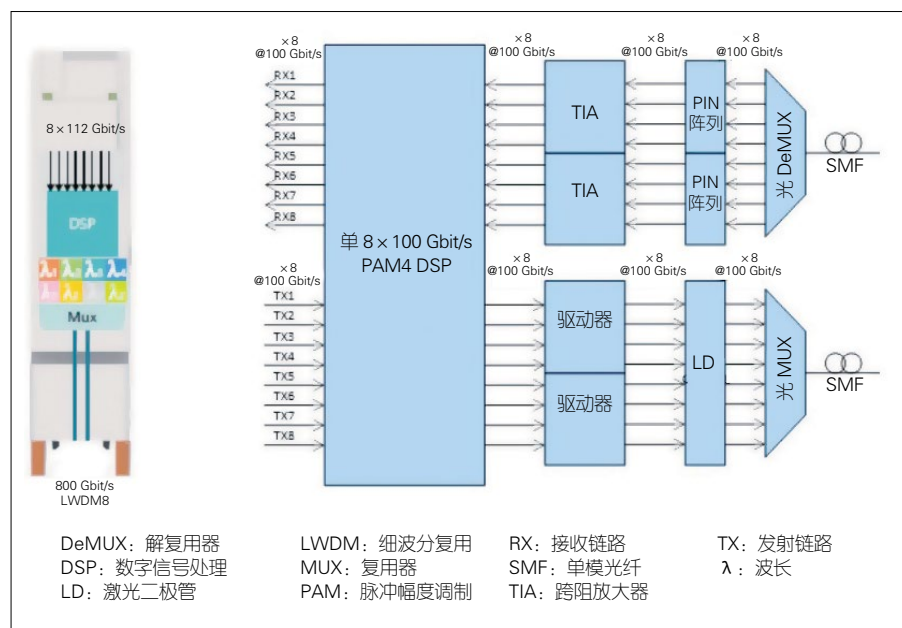
案单位比特率成本等因素密切相关。相干方案只需要一个激光器、调制器和接收器的这一事实, 将使自身具备与 PAM4 相媲美的成本竞争力 (即使光器件变得更加复杂)。PAM4 会用到 4 个简单的激光器、调制器和接收器。即使这些器件的在 800 Gbit/s 时比较复杂, 它们也足以快速降低整个系统成本, 使 PAM4 保持较强的竞争力。总体而言, 相干和 PAM4 传输的竞争已经开始, 未来结果如何, 还需要时间来证明。

800 Gbit/s 光模块电接口前期为 8×100 Gbit/s, 光接口有 8×100 Gbit/s PAM4、 4×200 Gbit/s PAM4、128 Gbd DP-16QAM 相干 3 种。800 Gbit/s 光模块应用场景方案预测如表 1 所示。

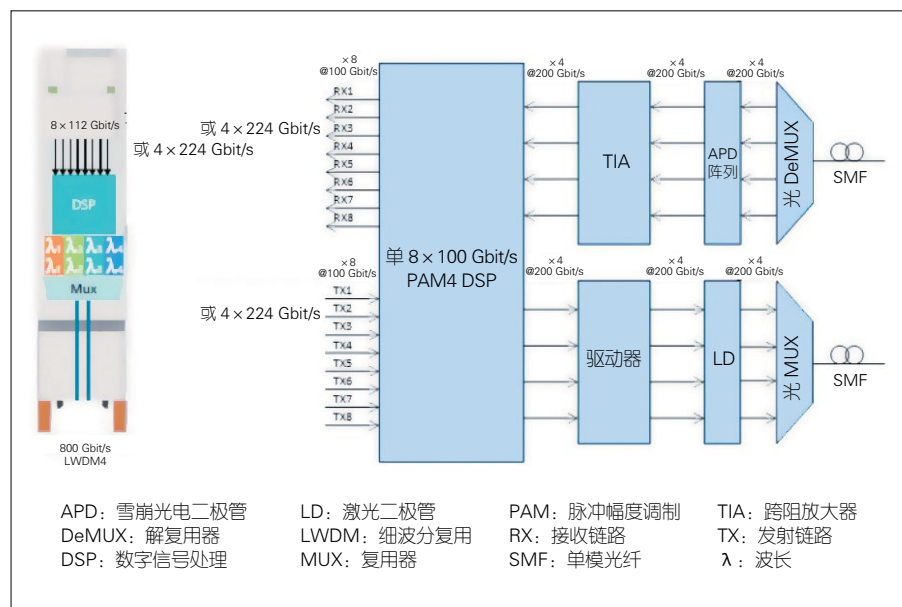
单模方案下沉有助于 800 Gbit/s SiPh 方案的光模块覆盖到海量 100 m SR 场景。800 Gbit/s-SR8/DR8 采用同平台方案并按指标筛选发货, 有助于降低系统平均成本。此外, SiPh 光模块有可能实现非相干领域弯道超车。100 Gbit/s 光模块产业链百花齐放, 呈现多元化现象。但是 200 Gbit/s 产业链多元化会降低 (仅有 EML 方案和 MZ 方案)。单波 200 Gbit/s 光接口 (800 Gbit/s-FR4) 将在 800 Gbit/s 和 1.6 Tbit/s 中使用。200 Gbit/s PAM4 EML 的供方有限, 并且带宽略显不足。200 Gbit/s PAM4 SiPh 的硅光调制器和 200 Gbit/s PAM4 薄膜铌酸锂在性能、成本和产业链多样化方面具有很好的竞争优势。

4 结束语

当交换芯片速率达到 51.2 Tbit/s 时, 800 Gbit/s 光模块需求将产生; 当交换芯片的速率达到 102.4 Tbit/s 时, 800 Gbit/s 和 1.6 Tbit/s 光模块需求均将出现。根据交换芯片的演进趋势、市场需求和技术成熟度, 800 Gbit/s 光模块



▲图 7 8×100 Gbit/s LWDM8 EML



▲图 8 4×200 Gbit/s LWDM4 EML

▼表 1 800 Gbit/s 光模块应用场景方案预测

800 Gbit/s 应用场景		商业方案预测
SR	基于 SMF 单模的 800 Gbit/s SR8 解决方案是可行的, 具有一定竞争力	
DR	800 Gbit/s DR8 在前期会成为商用选择, 后期会切换为 800 Gbit/s DR4	
FR	第 1 代会是 2×400 Gbit/s FR4, 800 Gbit/s CWDMM4 将成为主流模块。	
LR	第 1 代会是 2×400 Gbit/s LR4, 800 Gbit/s LR8、800 Gbit/s LR4、800 Gbit/s 相干将会出现竞争。其中, 800 Gbit/s LR8 将成为主流选择	
ER/ZR	800 Gbit/s 相干方案作为唯一技术方案。	
CWDM: 粗波分复用	FR: 2 km 场景	SR: 100 m 场景
DR: 500 m 场景	LR: 10 km 场景	ZR: 80 km 场景
ER: 40 km 场景	SMF: 单模光纤	

将于 2022 年进入市场, 在 2025 年左右实现规模应用。在早期 8×100 Gbit/s 将成为主流方案, 这将有助于 400 Gbit/s 向 800 Gbit/s 平滑演进。

当单通道电接口速率与光接口速率相同时, 光模块的架构将达到最佳, 并具有低功耗、低成本等优势。 4×200 Gbit/s 将是 800 Gbit/s 光模块的理想架构, 未来将成为 1.6 Tbit/s 的实现基础。因内部封装形态受到高频信号完整性的影响, 200 Gbit/s 每波长光模块有光电集成化的可能, 而这对成本的影响还有待分析。由于波特率的提高, 200 Gbit/s PAM4 DML 带宽受限, 200 Gbit/s PAM4 10 km 以上长距传输的灵敏度将受到影响, 而高带宽波导 APD 对中长距 800 Gbit/s 的成本影响尚不明确。超短距离采用 8×100 Gbit/s

方案, 短距离采用 4×200 Gbit/s 方案, 中长距离仍采用 8×100 Gbit/s 方案。受 PAM4 色散、多路径串扰 (MPI) 等影响, 长距离采用 800 Gbit/s 相干方案。

致谢

本研究得到中兴光电子技术有限公司沈百林、熊孝海等专家的帮助, 谨致谢意!

参考文献

- [1] Omdia. Datacom optical components forecast, 2020–26 [EB/OL]. [2021–10–15]. <https://omdia.tech.informa.com/OM017740/Datacom-Optical-Components-Forecast-202026>
- [2] 800G Pluggable MSA. 200G per lane for future 800G & 1.6T modules [EB/OL]. [2021–11–15]. <https://www.800gmsa.com/documents/200g-per-lane-for-future-800g-and-16t-modules>
- [3] WAYW. 800G: coherent versus PAM4 optical transceivers inside data centers [EB/OL]. [2019–09–09][2021–11–15]. [https://www.neophoton-](https://www.neophoton-ics.com/800g-coherent-versus-pam4-optical-transceivers-data-centers)

ics.com/800g-coherent-versus-pam4-optical-transceivers-data-centers

[4] 中国通信标准化协会. 800 Gbit/s 光收发合一模块研究: 2020B46 [S]. 2020

[5] IMT-2020 (5G) 推进组. 5G 承载与数据中心光模块 [R]. 2021

作者简介



王会涛, 中兴光电子技术有限公司规划总工; 长期从事光通信系统、光电子器件产品和技术的研究及规划工作; 曾获得中国通信学会科技进步奖一等奖、教育部技术发明奖二等奖; 拥有专利 10 余项。



张平化, 中兴光电子技术有限公司光模块规划工程师; 从事光模块规划工作。



苏展, 中兴通讯股份有限公司光模块系统工程师; 从事光模块的预研、研发和标准跟踪等工作。



TSN 与 5G 融合部署的需求和网络架构演进

Requirements and Network Architecture Evolution of TSN and 5G Integrated Deployment

摘要: 认为时间敏感网络 (TSN) 技术与 5G 网络系统的融合部署将成为 5G 网络向确定化演进的重要方向。从 TSN 技术现状入手, 分析了 TSN 与 5G 网络系统融合的必要性和可能性, 重点提出了 TSN 与 5G 网络系统融合部署的演进路径和 3 种融合部署形式。

关键词: 时间敏感网络; 5G 网络系统; 前传网络; 承载网; 核心网

Abstract: It is considered that the integration and deployment of time sensitive network (TSN) technology and 5G network system will become an important direction for the deterministic evolution of 5G network. The technical status of TSN technology is discussed. The necessity and possibility of the integration of TSN and 5G network system are analyzed. The evolution path and three integration deployment forms of TSN and 5G network system are highlighted.

Keywords: time sensitive network; 5G network system; fronthaul network; bearer network; core network

朱瑾瑜 /ZHU Jinyu
张恒升 /ZHANG Hengsheng
陈洁 /CHEN Jie

(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.202106009
网 络 出 版 地 址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20200630.1413.004.html>

网络出版日期: 2020-06-30
收稿日期: 2020-06-10

5G 将开启一个全连接的新时代, 将通信的作用从人与人之间的连接扩展到各行各业、万事万物之间的相互连接, 以非常好的业务体验、效率和性能使能垂直行业, 以用户为中心构建全方位信息生态系统, 为运营商和产业合作伙伴带来新的商业模式, 形成行业数字化新格局。与此同时, 垂直行业多种多样的业务需求场景要求 5G 网络具备差异化的网络定制能力, 并可实现业务数据的确定性传输。这种确定化的业务保障能力涵盖时延、时延抖动和丢包率等关键服务质量 (QoS) 指标。5G 网络通常借助移动边缘计算、网络切片和时间敏感网络 (TSN) 等多种关键技术, 为业务承载提供确定化的传输保障能力。

相关 TSN 技术基础协议的制定和应用推进主要由电气与电子工程师协会 (IEEE) 802.1 TSN 工作组负责。因

具备高精度时间同步、个性化精准流量调度、智能化网络管理机制等特性, TSN 可以广泛应用于存在互联互通、确定性高质量传输、全业务承载需求的业务数据传送场景。近年来, TSN 技术在音视频传输、工业、移动承载、车载网络等各个应用场景受到产业界更为广泛的关注, 众多机构和组织纷纷开展相关研究工作。

1 TSN 技术现状

1.1 TSN 发展背景

关于 TSN 技术的研究最早起源于 2006 年。IEEE 802.1 工作组在这一年成立音视频桥接 (AVB) 任务组。AVB 技术成功实现了数据在音视频网络中的确定时序和低时延传输, 并通过对流量的整形调度解决了音视频流量的实时同步确定性传输问题。随

后, 该项技术在网络带宽保证及有界传输时延方面的潜在能力被工业领域所关注。2012 年, AVB 任务组更名为 IEEE 802.1 TSN 任务组, 并对时间确定性以太网的应用需求和适用范围进行了扩展, 以覆盖音视频以外的更多领域, 如汽车、工业制造、运输、过程控制、航空航天以及移动通信网络等。IEEE 802.1 TSN 任务组聚焦 TSN 技术在各类通信场景下的基础共性技术及标准的研究, 已经推出以 IEEE 802.1AS 为基础的时间同步协议、以 IEEE 802.1Q 系列协议为代表的流量调度及网络管理相关协议。目前这些协议已经基本成熟。

IEEE 802.1 工作组致力于 TSN 的标准化工作。TSN 是当前实现确定性网络的技术方向。IEEE 802.1AS 时钟同步、IEEE 802.1Qcc 流预留、IEEE 802.1Qch 循环排队等技术可保障物理

层和链路层的确定性时延^[1-5]。

1.2 TSN 研究现状

随着 TSN 技术受到越来越多的关注, 各类通信标准和行业机构针对 TSN 相关技术在垂直行业的部署应用均展开了研究。相关的标准制定组织有 IEEE、国际电工委员会 (IEC)、国际互联网工程任务组 (IETF)、通用公共无线接口 (CPRI) 联盟、中国通信标准化协会 (CCSA) 等。参与 TSN 研究和推广的机构及产业联盟有中国工业互联网产业联盟 (AII)、美国工业互联网联盟 (IIC)、德国工业 4.0 实验室网络 (LNI4.0), 以及国际联盟组织音视频网络联盟 (AVNU)、用于过程控制的对象连接与嵌入 (OPC) 基金会、OpenStack 基金会等^[6-7]。

(1) 工业网络

新型工业网络不仅对带宽有很高的要求, 还必须同时保证控制数据的周期性可靠传输。例如, 每隔 100 μ s 设备与控制器之间就要进行数据交互, 在这期间不允许有其他数据对这种周期性数据造成阻塞, 也不允许有时快时慢的延时响应。TSN 能够符合工业网络中这种“确定时延、不被中断、可靠传输”的要求。同时 IEEE 又制定了相应的标准, 使得 TSN 更具中立性。2017 年, IEC 和 IEEE 联合成立了 P60802 工作组, 旨在定义 TSN 应用于工业自动化网络的方案类标准。这一举动得到了很多厂商的支持, 如贝加莱、西门子、施耐德、罗克韦尔、三菱、思科、霍斯曼、MOXA 等。随后, 这些主流自动化厂商开始陆续推出基于 TSN 技术的工业以太网产品。由此可见, TSN 成为工业通信领域厂商的共同选择已经是大势所趋。

随着工业企业的信息化转型, 生产系统之间的信息互通逐渐成为普遍需求。工业领域存在诸多私有协议,

要把不同厂商间各个终端节点的数据打通。这需要解决两个层面的问题: 一个是统一的网络架构和链路协议, 另一个是各厂商终端或系统间的数据交互接口。目前业内倾向于利用 OPC UA over TSN 的网络架构来实现这种互联需求, 即由 TSN 提供统一的链路层实时传输, 同时由 OPC UA 提供统一的信息交互架构。OPC 基金会下的 Shaper 工作组正致力于推动此项融合技术的落地。ABB、思科、倍福、华为等厂商是其中的主要参与者。

(2) 车载以太网网络

在传统的车载网络中, 各类总线的应用范围有着比较明确的区分, 各设备和应用对带宽的要求也相对较低, 如表 1 所示。

面向媒体的系统传输 (MOST) 总线技术在一定程度上可以满足在引入多媒体信息娱乐系统 (MIS) 和基于高清摄像头的高级驾驶辅助系统后大量视音频数据的传输对网络带宽和时延的新需求。然而, 随着自动驾驶技术的发展, 辅助驾驶系统 (ADAS) 将大量精密的传感器融入系统之中, 对车载网络大带宽、低时延、确定性的需求日益迫切。同时人机交互越来越多样化, 智能化数据与电气参数之间的联动也需要一个统一融合的网络^[8-10]。

TSN 基于标准以太网技术, 能够在二层网络提供高带宽、高可靠、确定性的流量承载和数据传输服务, 符合车载网络的发展需求。因此, 在 TSN 的支持下, 基于以太网发展下一代车载网络已经成为一种必然趋势。

目前, IEEE 802.1 TSN 任务组已经启动 P802.1DG 项目, 以进行 TSN 技术在车辆信息娱乐系统、驾驶辅助系统的应用研究。2016 年起, 芯片厂商恩智浦、英伟达、博通陆续研发 TSN 相关芯片, 并与知名车企联合, 进行了一系列基于 TSN 技术的车载网络应用测试及验证工作。2017 年初, 行业组织 AVnu 联盟发布了《车载以太网及 AVB 技术应用》白皮书, 推动了 TSN 技术在车载音视频领域的应用。

(3) 移动前传网络

4G 网络采用了集中化无线接入网 (C-RAN) 架构, 即分布式的射频拉远单元 (RRU) 共享集中部署的基带单元 (BBU) 的方式, 以便解决重负载和轻负载基站之间 BBU 的利用均衡的问题。BBU 和 RRH 之间连接的网络称为前传网络。CPRI 标准是前传网络中应用最为广泛的接口规范。5G 网络的带宽消耗大概是 4G 的 1 000 倍, 使得移动承载网络面临更大的传输压力。由于其实现机制的原因, CPRI 接口只有利用造价高昂的光设备才能满足相应的性能要求, 且维护难度大, 不能很好地满足 5G 时代前传网络的承载要求。

由爱立信、华为、NEC 和诺基亚共同在 CPRI 的基础上设计的增强型 CPRI (eCPRI) 更加适应 5G 网络。相比于 CPRI, eCPRI 具有高带宽、高扩展性、低设备成本等优势, 并且可以支持基于以太网的承载方式, 为与 TSN 技术进行集成提供了基础。2018 年 11 月, IEEE 和 IEEE 标准协

▼表 1 车载网络主要总线一览表

类别	总线	通信速度 / ($\text{bit} \cdot \text{s}^{-1}$)	应用场景
A 类	LIN	$10^4 \sim 1.25 \times 10^5$	车身 (大灯、灯光、门锁、电动座椅等)
B 类	CAN	$1.25 \times 10^5 \sim 10^6$	状态信息 (汽车空调、电子指示、故障诊断等)
C 类	FlexRay	$10^6 \sim 10^7$	实时控制 (引擎控制、ABS、悬挂控制、线控转向等)
D 类	MOST/1394	10^7 以上	汽车导航系统、多媒体娱乐系统等

ABS: 防抱死制动系统
LIN: 局域互联网络

CAN: 控制器局域网络
MOST: 面向媒体的系统传输

会 (IEEE-SA) 发布了针对前传时效性网络的 IEEE 802.1CM-2018 标准。新标准解决了使用以太网将蜂窝无线电设备连接到远程控制器的问题, 有望在 5G 小型蜂窝网络以及未来基于云的无线接入网络的技术设计中发挥重要作用。

(4) 运营商网络

2015 年国际标准组织 IETF 成立了确定性网络 (DetNet) 工作组, 与负责第 2 层操作的 IEEE 802.1 TSN 工作组合作, 为第 2 层和第 3 层定义通用架构, 致力于在第 2 层桥接段和第 3 层路由段上建立确定性数据传输路径。这些路径可以提供延迟、丢包和数据包延迟变化 (抖动) 以及高可靠性的界限。可以说, DetNet 是广义的 TSN 技术。目前 IETF DetNet 工作组已经完成整体架构、数据平面说明、数据流信息模型等的成果交付, 为 TSN 技术应用于以互联网协议 (IP) 或多协议标签交换 (MPLS) 技术为基础的运营商广域网奠定了基础。同时, IEEE 802.1 TSN 任务组已经启动 P802.1DF 研究项目, 以定义 TSN 技术在运营商服务网络中的应用场景和技术需求。

(5) 信息网络

StarlingX 是 OpenStack 基金会旗下的一个试点项目, 它提供了一个完整的云堆栈, 专门为在边缘部署云而设计, 包括远端或 Lastmile, 以及工厂的内部云、工业物联网、多接入边缘计算 (MEC) 和虚拟无线接入网络 (vRAN) 等用例。StarlingX 的开源同时响应 Linux 基金会的 Akraino 项目。最新发布的 StarlingX 3.0 版本引入了 TSN, 开始尝试利用 TSN 技术实现在云计算、边缘计算等信息领域网络中数据流量的确定性传输。

(6) 5G 网络

第 3 代合作伙伴计划 (3GPP) R16 将 5G 端到端时延目标定为 1 ms。

就现有 5G 超可靠低延迟通信 (URLLC) 标准而言, R16 主要用于实现无线终端与基站之间的传输, 其技术思路与 TSN 并不相同。3GPP R16 23.501 已经将 TSN 技术纳入 5G 标准, 用于满足 5G 承载网的高可靠、确定性需求, 并与 URLLC 形成确定性传输的技术接力。5G URLLC 技术主要关注在可靠性和时延方面的业务保证, 而 TSN 技术则将在时延抖动以及时间同步方面对 5G 网络进行进一步增强。

3GPP R17 提出 TSN 增强架构。这种增强具体包括: 实现 5G 核心网架构的增强, 使控制面设计支持 TSN 相关控制面功能; 实现 5G 核心网确定性传输调度机制, 而不依赖外部 TSN 网络; 通过用户面功能 (UPF) 增强实现终端间的确定性传输; 实现可靠性保障增强和工业以太网协议对接, 并支持多时钟源技术。

2 5G 网络的特点及关键需求

2.1 5G 网络特点

作为新一代移动通信系统, 5G 致力于与垂直行业协同实现万物实时全连接。国际电联无线电通信部门 (ITU-R) 定义的增强移动宽带 (eMBB)、URLLC 和海量机器类通信服务 (mMTC) 三大需求是实现这种协同的技术基础。这主要体现在两个方面: 一方面, 5G 网络架构实现了转发面与控制面的彻底分离, 与网络切片和边缘计算技术的结合使得面向行业的可定制化网络构建成为可能。这

不仅可以实现垂直网络服务的快速上线和差异化服务质量保证, 还可以促进 5G 网络向垂直行业应用的开放发展。另一方面, 5G 网络因具有大带宽 (0 ~ 10 Gbit/s)、低时延 (1 ~ 100 ms) 和高可靠性 (99.9999%) 等能力, 已经获得业界广泛认可, 并被应用于工业互联网、智能电网、车联网通信、远程手术等新的垂直行业场景。

在 5G 标准的演进过程中, 3GPP R15 阶段致力于满足 eMBB 和 URLLC 敏感应用需求。在这一阶段, 3GPP 5G 相关标准可用于制造专业 5G 设备和建设独立全新的网络。5G 网络系统可满足超高视频、虚拟现实 (VR) 直播等对移动宽带的要求。虽然 3GPP R15 标准已经进一步减少了传输延迟, 并且提高了稳定性, 但是这些性能并不能满足一些对时延要求严格的应用需求。

3GPP 于 2018 年 6 月发布的 5G 独立组网 (SA) 标准明确定义了 5G 大带宽、高可靠、低时延等特性的实现方式, 使 5G 网络系统具备了在工业互联网、车联网、医疗等垂直行业应用的核心属性。3GPP 正在制定的 R16 版本正式确定 15 个研究方向, 明确将进一步研究 URLLC 来满足诸如工业制造、电力控制等工业物联网场景需求, 以及基于 5G 新空口的车用无线通信技术 (V2X) 场景需求。在 3GPP R16 的网络架构文稿 23.501 定义的 5G 关键应用场景中, 工业物联网和车联网对于业务的端到端时延要求最为严格, 如表 2 所示。

▼表 2 工业物联网和车联网对业务的端到端时延要求

主场景	子场景	端到端时延要求 /ms	3GPP 文件
工业物联网	离散制造业	10	TS 22.261 [2]
	智慧物流	30	TS 22.261 [2]
	高压配电	5	
车联网	高级驾驶、协同换道	10	TS 22.186 [111]
	远程驾驶	5	

3GPP: 第 3 代合作伙伴计划

2.2 TSN 与 5G 融合部署需求

在以工业为代表的垂直行业业务中,安全可靠、确定性地传输数据是通信技术的关键要求之一。工业互联网、车联网等应用对 5G 网络架构设计及技术选型提出了三大挑战。

(1) 端到端极致确定性业务体验

在工业互联网和车联网场景下的应用系统中,典型的闭环控制过程周期可能低至毫秒级别。这种应用系统不仅对可靠性有着极高的要求,还对业务数据传输有着十分严格的确定性要求。实现整个 5G 网络系统中包括新空口(NR) RAN 核心网在内的各个环节的性能优化和系统整体处理效率的提升,才可能实现端到端的极致高可靠低时延业务应用。TSN 技术在现有以太网 QoS 功能基础上增加了时间片调度、抢占、流监控及过滤等一系列流量调度特性,能够根据业务流量的特点配合使用相关特性,可确保流量的高质量确定性传输。将 TSN 技术与 5G 网络的传输进行融合,可以更为有效地保证 5G 网络的端到端的高可靠低时延传输。

未来信息系统与生产设备之间的数据交互量将呈指数级增长。尤其是在部署了云平台 and 边缘计算节点的情况下,工业信息网络还将抵挡算力网络的流量冲击,对网络负载有较高要求。这就要求新一代工业网络能够实现大负载和确定性的高质量传送。

(2) 异构系统的精密协作

5G 网络系统将以业务为中心全方位构建信息生态系统,实现设备之间的全面连接和精密协作。以智能工厂为例,生产设备、移动机器人、自动导引车(AGV)等智能系统内部均存在异构网络连接,并且各个系统可能会通过不同的方式接入到 5G 网络中。要实现这些设备系统之间的密切协同和无碰撞作业,就需要业务系统彼此

之间能够做到互联互通。

TSN+OPC UA 的组合被认为是解决异构系统互联互通问题的最佳组合,它可以同时实现网络的互联和数据层面的互通。TSN 技术基于标准以太网协议解决数据报文在数据链路层中的确定性传输问题;OPC UA 则提供一套通用的数据解析机制,以用于业务系统端设备,解决数据交换和系统互操作的复杂性问题。

(3) 全业务承载差异化的传输质量保证

5G 网络全面使能垂直行业新业务模式。以智能工厂为例,工业增强现实可以通过音视频实现生产环境远程感知,以及在线生产指导。远程控制可以用于实现远程人机交互。在恶劣环境下使用机器人有助于实现安全生产。此外,大量的设备维护、原材料及产品数据信息都需要通过传感器、射频识别(RFID)、智能终端等方式上传到云端。上述业务涉及的音视频、控制信号、物联网数据的传输会采用不同的传输机制和质量标准。虽然分片技术可以用来实现不同业务之间的差异化业务,但是目前的分片技术仅可以在空口和核心网实现,对于承载网则没有特定的技术方案。TSN 基于软件定义网络(SDN)架构实现网络资源的集中管理和按需调度,同时配合精确时间同步、流量调度等功能,可为不同类型的业务流量提供智能化、差异化的承载服务。将 TSN 技术与 5G 承载网融合部署,或许可以为 5G 端到端分片提供一种解决思路。

3 TSN 与 5G 融合部署演进

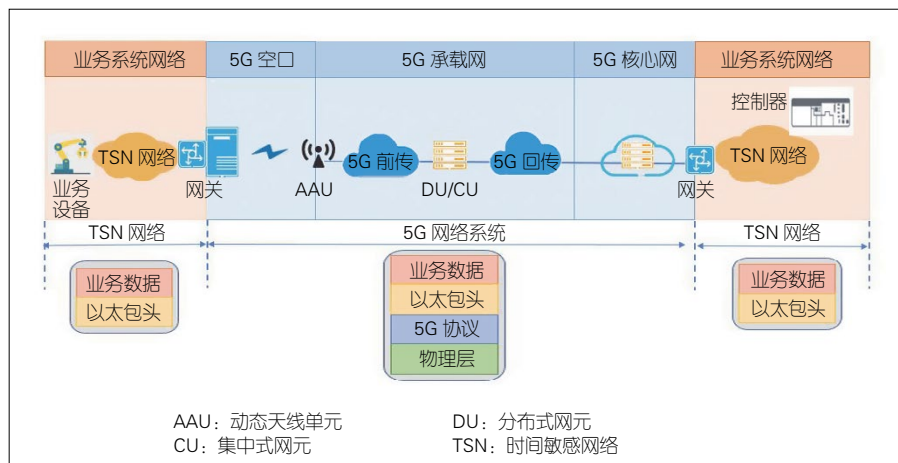
从 3GPP R16 相关标准中可以看出,端到端确定性传输是 5G 网络能够在工业互联网、车联网等对网络传输有极致要求的垂直领域落地应用的关键。5G 的确定性网络应该是一个端到

端的概念,它涉及业务系统网络、无线空口、承载网、核心网等多段网络。结合目前产业界需求及相关研究成果分析,我们认为 TSN 与 5G 的融合部署大体可以分为拼接式融合、5G 承载网融合和深度融合 3 个阶段。

3.1 拼接式融合

2018 年 6 月 3GPP 发布的 5G SA 标准定义了 5G 大带宽、高可靠、低时延等特性的具体实现方式,进一步降低了业务传输时延,提高了系统稳定性。在此基础上,5G 网络系统与 TSN 技术融合并提供网络服务的关键在于 5G 网络与 TSN 网络的互通,即把原有已经具备 TSN 特性的业务系统(如工业控制网络、车载网络等)与 5G 系统进行拼接,协同流量调度,通过分段实现业务传输的确定性来提升端到端业务传送质量,如图 1 所示。

在此类方案中,整个业务系统被看成一个用户设备(UE),TSN 中的流量分类要与 5G 网络系统的业务类型建立映射关系,同时 TSN 对于流量配置的相关标记需要被保留。流量在经过 5G 网络系统的远程传输后剥离 5G 封装,并且在进入到协同业务系统后,仍然按照 TSN 流量调度类型进行确定性传输。该方案的关键在于 TSN 网络与 5G 网络的边缘应部署对应网关。按照部署位置的不同,这种网关有两种:部署于 TSN 网络与蜂窝无线网络之间的 UE 侧网关和部署于 TSN 与 5G 核心网之间的核心网侧网关。TSN 与 5G 融合部署网关在业务系统网络侧的接口需要具备 TSN 的相关特性,并兼具将业务系统数据及包含 TSN 特性的以太网封装进 5G 传输包头中的功能。在封装过程中,网关还要将相关的业务流量标识映射到 5G 网络传输结构中。此外,UE 侧网关还需要具备有线网络向无线蜂窝网络转换的能力。



▲图1 TSN 与 5G 网络系统拼接部署示意

3.2 承载网融合

除了 5G 网络系统 NR 标准及新的核心网架构以外, 承载网络的重构也是一个重要研究方向。5G 网络系统中承载网络通常采用有线网络进行流量承载。在分布式网元 (DU) 和集中式网元 (CU) 合设的情况下, 承载网络通常可以分为前传和回传两部分, 如图 2 所示。

自 3G 开始移动回传网络通常采用包转发技术进行基站到核心网之间的流量承载, 例如无线接入网 IP 化 (IPRAN)、分组传送网 (PTN), 并借助 MPLS 转发技术实现业务流量的转发、调度和保护倒换。目前基本承

载技术已相对稳定成熟。5G 时代的回传网络一方面结合 SDN 和网络功能虚拟化 (NFV) 技术将驱动回传网络的智能化演进, 另一方面天然具备利用确定性网络技术 (MPLS Over TSN) 实现回传网络低时延、低抖动业务传输的能力^[11-13]。

在 4G 时代, 射频单元与基带单元分离的基站架构催生了前传网络的概念。CPRI 可作为前传网络主流接口。射频单元与基带单元的无线电单元通过光纤进行直连, 只需要将信号调制到有线物理载波上即可完成传输。为更好地满足 5G 网络大带宽、高可靠、低时延传输要求, 爱立信、华为、

NEC 和诺基亚在 CPRI 的基础上创建了更加适应 5G 网络的 eCPRI。eCPRI 接口用户面数据基于以太网进行传送。以太网具有带宽大、扩展性高、设备成本低等优势, 可以更好地与光传送网相匹配。可以说以 eCPRI 为接口的 5G 前传网络已经具备了与 TSN 融合部署的技术前提。事实上, IEEE 为此成立了两个相关项目组: IEEE 1914 下一代前传接口工作组和 IEEE 802.1CM 前传 TSN 工作组。

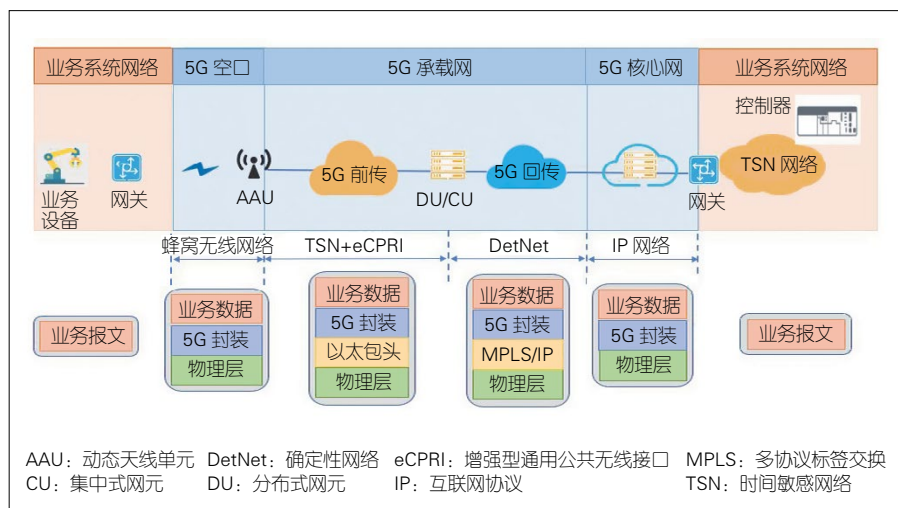
值得注意的是, 当 CU 与 DU 分离部署真正实现后, 前传的传输距离可能缩短, DU 与 CU 之间的中传网络将很可能采用包传输方式进行数据传输。因此, TSN 与 5G 中传网络的结合也将是融合部署的一个关键点^[14-16]。

TSN 与 5G 承载网的融合, 不仅存在利用 TSN 技术驱动承载网实现确定性传输的需求, 还具备从回传到前传再到中传部署 TSN 技术的基本技术前提。TSN 与 5G 承载网融合部署的实现, 将推进确定性传输方案从业务系统 TSN 网络与 5G URLLC 的拼接模式向 5G 网络系统内部承载网融合方向演进。

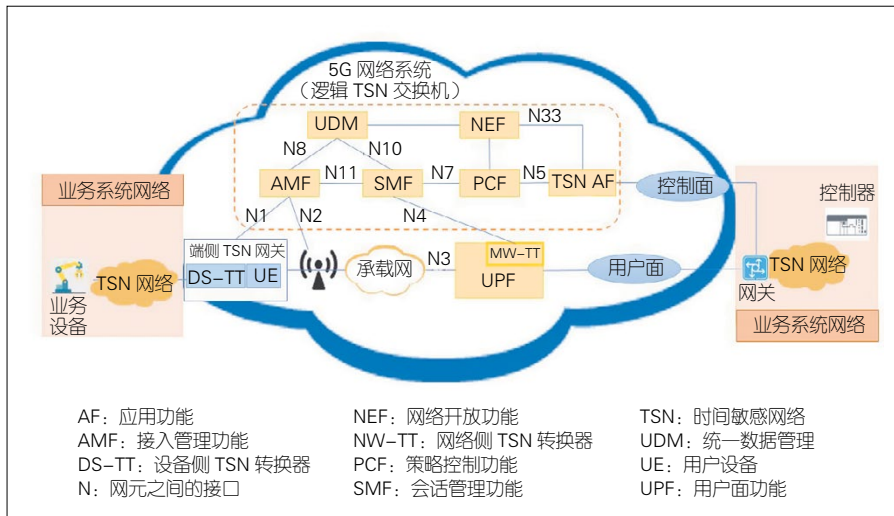
3.3 深度融合

在 TSN 与 5G 深度融合阶段中, 整个 5G 网络系统将升级为具备 TSN 特性的桥接系统, 以承载业务系统流量的远程确定性传送。3GPP R16 23.501 已经明确提出相关技术思路, 如图 3 所示。

在深度融合的架构下, 相对于业务系统 5G 网络被视为黑盒 TSN 交换机, 支持 TSN 集中式架构和时间同步机制, 并通过定义新的 QoS 模型 (流方向、周期、突发到达时间) 来实现精准的流量调度, 同时实现 5G 系统中 UE 到 UPF 之间多种确定性业务流量的共网高质量传输。其中涉及的主要关键技术有 4 个方面。



▲图2 TSN 与 5G 承载网络融合部署



▲图3 TSN 与 5G 深度融合部署

(1) TSN 技术与空口传输的融合: 在 URLLC 通信服务基础上增加时间同步、时延和时延抖动的有界性定义, 将 TSN 技术思路应用于无线空口。

(2) 设备侧 TSN 转换器 (DS-TT): 在 UE 侧部署 DS-TT 对相关端口、协议数据单元和 QoS 机制进行 UE 与业务系统之间的映射, 并支持 TSN 相关流量调度。

(3) 网络侧 TSN 转换器 (NW-TT): NW-TT 对相关端口、协议数据单元和 QoS 机制进行核心网与业务系统之间的映射, 并支持 TSN 相关流量调度。

(4) TSN 应用功能 (AF-TSN): 支持在同一 UPF 下的 UE 与 UE 之间确定性通信的能力开放。

在深度融合阶段, TSN 技术思路被进一步应用到 5G 网络系统中。除了承载网以外, 用户面的空口和核心网融入了 TSN 的流量调度特性, 同时控制面还结合了 TSN 的网络架构。

4 结束语

在 5G 进入规模化部署以后, 以工业互联网、车联网为代表的垂直行业对网络系统的确定性提出了更为严格要求。本文从 TSN 发展现状和技术

优势等方面分析了 TSN 适用的 6 种场景, 并结合 5G 网络系统特点及相关行业对于 5G 网络系统的新需求, 分析了 TSN 与 5G 网络系统融合部署的必要性, 重点提出了 TSN 与 5G 融合部署的演进思路和技术方案。总体而言, TSN 与 5G 网络系统的融合正处于探索阶段, 尚有许多技术细节需要进一步研究论证。但是可以预见, 这种融合将成为 5G 网络系统向确定化演进的重要技术方向。

参考文献

- [1] 3GPP. System architecture for the 5G System (5GS): 3GPP TS 23.501 [S]. 2019
- [2] 3GPP. Policy and charging control framework for the 5G System (5GS): 3GPP TS 23.503 [S]. 2019
- [3] 3GPP. Service requirements for the 5G System: 3GPP TS 22.261 [S]. 2018
- [4] Ericsson, Huawei, NEC, et al. Common public radio interface: eCPRI interface specification V2.0 [EB/OL]. (2019-05-10)[2020-01-15]. http://www.cpri.info/downloads/eCPRI_v_2.0_2019_05_10c.pdf
- [5] IEEE. IEEE standard for local and metropolitan area networks — bridges and bridged networks — amendment: IEEE 802.1Q-2018 [S]. 2018
- [6] IEEE. Enhancements to fronthaul profiles to support new fronthaul interface, synchronization, and synchronization standards: IEEE P802.1CMde [S]. 2019
- [7] IEEE. TSN profile for automotive in-vehicle ethernet communications: IEEE P802.1CDG [S]. 2019
- [8] IEEE. TSN profile for service provider networks:

IEEE P802.1CDF [S]. 2019

- [9] IETF. IETF Deterministic Networking (DetNet) working group [EB/OL]. [2020-01-03]. <https://datatracker.ietf.org/wg/detnet/>
- [10] 王卫斌, 朱进国, 王全. 5G 核心网演进需求及关键技术 [J]. 中兴通讯技术, 2020, 26(1): 67-72. DOI: 10.12142/ZTETJ.202001015
- [11] 黄韬, 汪硕, 黄玉栋, 等. 确定性网络研究综述 [J]. 通信学报, 2019, 40(6): 160-176. DOI: 10.11959/j.issn.1000-436x.2019119
- [12] 工业互联网产业联盟. 时间敏感网络 (TSN) 产业白皮书 (征求意见稿) [EB/OL]. (2019-10-31)[2020-01-03]. <http://www.aii-alliance.org/index.php?m=content&c=index&a=show&-catid=23&id=800>
- [13] 张蕾, 夏旭, 朱雪田. 基于 5G 确定化网络的行业应用研究 [J]. 电子技术应用, 2019, 45(12): 20-28
- [14] ITU. Functional architecture for supporting fixed mobile convergence in IMT-2020 networks: ITU-T Y.3131 [S]. 2019
- [15] 赵福川, 刘爱华, 周华东. 5G 确定性网络的应用和传送技术 [J]. 中兴通讯技术, 2019, 25(6): 62-67. DOI: 10.12142/ZTETJ.201905010
- [16] IMT-2020 (5G) 推进组. 5G 愿景与需求 [R]. 2014

作者简介



朱瑾瑜, 中国信息通信研究院技术与标准研究所产业互联网研究部工程师; 主要研究方向为工业互联网网络技术、时间敏感网络技术、移动承载网络技术等, 从事技术研究及相关标准研制工作; 参与工信部、科技部多个工业互联网

项目的研究与建设。



张恒升, 中国信息通信研究院技术与标准研究所产业互联网研究部副主任; 主要研究方向为工业互联网网络; 主持参与工业互联网总体网络架构等多个标准的制定工作, 参与工信部、科技部多个工业互联网项目的研究与建设。



陈洁, 中国信息通信研究院技术与标准研究所产业互联网研究部工程师; 主要研究方向为工业互联网网络架构、园区网络架构以及工厂外网架构等; 参与工信部、科技部多个工业互联网项目的研究。



基于典型无线场景库的 LTE-V2X 信道特性

Channel Characteristics of LTE-V2X Based on Typical Wireless Scenario Library

苏昭阳 /SU Zhaoyang¹

刘留 /LIU Liu¹

冯毅 /FENG Yi²

(1. 北京交通大学, 中国 北京 100044;

2. 中国联合网络通信集团有限公司, 中国 北京 100048)

(1. Beijing Jiaotong University, Beijing 100044, China;
2. China Unicom, Beijing 100048, China)

DOI: 10.12142/ZTETJ.202106010

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211210.1946.004.html>

网络出版日期: 2021-12-13

收稿日期: 2021-11-10

摘要: 分析了基于长期演进的车用无线通信技术 (LTE-V2X) 的关键技术、面临的挑战和和相关研究进展。详细介绍了北京交通大学和中国联合网络通信集团有限公司在上海开展的实地测试工作, 制定了测试方案, 并基于实际测量的数据对信道特性进行了深入分析。

关键词: 蜂窝车用无线通信技术; LTE-V2X; 无线信道; 测试

Abstract: The key technologies, challenges, and related research progress of long term evolution-vehicle to everything (LTE-V2X) are analyzed. The measurement campaigns carried out in Shanghai by Beijing Jiaotong University and China Unicom are introduced in detail, and the test scheme is given. Based on the actual measured data, the channel characteristics are analyzed.

Keywords: C-V2X; LTE-V2X; wireless channel; test

如今汽车已经成为人们日常生活中不可缺少的出行工具。然而, 随着汽车保有量的增加, 一些潜在的问题也开始显现出来, 如交通安全、交通拥堵、环境污染等。在为人类带来便利的同时, 由汽车造成的负面影响已经不能忽视。于是车载无线通信技术 (V2X) 应运而生。V2X 包括车与车通信 (V2V)、车与物通信 (V2I)、车与人通信 (V2P)、车与网络通信 (V2N), 它可以显著提升驾驶的安全性与交通效率, 降低事故发生率^[1]。

V2X 面临着如何在无线传播环境快速时变与网络拓扑动态变化的情况下提供高可靠、低时延的通信服务等难题^[2]。目前, 世界上用于 V2X 的主流技术主要有两种: 一种是专用短程通信 (DSRC) 技术, 该技术采用电气与电子工程师协会 (IEEE) 802.11p 定

义的物理层协议; 另一种是基于蜂窝车用无线通信技术 (C-V2X), 该技术包括基于长期演进 (LTE) 移动通信技术形成的 LTE-V2X 技术和基于 5G 新空口 (NR) 演进形成的 NR-V2X 技术。其中, DSRC 可以提供多对多、低延时的通信, 但是可靠性较差^[2]; C-V2X 能提供更广的通信范围、更低的时延与更高的可扩展性^[3]。对于这两种技术, 中国选择使用 C-V2X 作为发展 V2X 的技术手段。

中国积极开展相关的研究与测试工作。C-V2X 产业化程度和产品接受程度都在逐步提升。一方面, 核心芯片、模组和终端产品的研发基本成熟; 另一方面, 交通运输行业和汽车厂商开始主动推广 C-V2X, 并开展了实地测试。2016 年, 工业和信息化部在重庆、武汉、长春进行了为期 3 年的 LTE-V2X 外场测试^[4]。2018 年, 中国率先为 LTE-V2X 直连通信分配

了 5 905 ~ 5 925 MHz 的专用带宽^[5], 并于当年 11 月在上海完成世界首例 C-V2X “三跨” 展示, 实现了通信模组、终端厂商、整车厂商 3 个不同方向的互联互通^[6]。2019 年, 在三跨的基础上, “四跨” 示范活动在上海举办, 四跨重点增加了跨安全平台的场景^[7]。2020 年, “新四跨” 活动在上海举办, 进一步推动了 C-V2X 相关技术和标准的测试验证, 同时增加了新技术元素, 以不断完善 C-V2X。

除中国外, 欧美国家和日本对相关技术的研究占据领先地位。美国主要基于 DSRC 开展研究工作, 并支持 C-V2X 作为备选技术。高通与福特开展了美国的第一个 C-V2X 试验。其中, 高通设计了一款针对 C-V2X 的芯片 (Qualcomm 9150)。欧洲从支持 IEEE 802.11p 逐渐转为技术中立态度。5G 汽车联盟与宝马集团等公司共同完成了第一个跨车型的 C-V2X 现场演示。此

基金项目: 国家自然科学基金重点项目 (61931001)

外,日本也将 C-V2X 作为备选技术。2018 年,日本多个通信和车辆公司联合进行了 C-V2X 直接通信测试^[8]。

随着 5G 的到来,C-V2X 面临着更多挑战。一方面,更加丰富的应用场景与业务种类要求通信技术能提供更低的时延与更高的可靠性;另一方面,与新技术如超大规模多输入多输出(MIMO)、毫米波通信、移动边缘计算等的结合也对 C-V2X 的发展提出了新要求。此外,V2X 通信的安全性也是一个亟待解决的问题,包括 V2X 通信中的功能安全、网络安全、隐私安全和数据安全^[2]。对此,人们需要构建 V2X 通信的安全防护体系。

面对这些挑战,进行 V2X 信道特性研究显得尤为重要。目前 V2X 信道特性的研究存在一些不足。相比于传统的蜂窝通信系统,V2X 信道具有工作频段高、发射端天线高度低、终端移动性高、受环境影响大的特点^[9]。通信双方的高速移动使信道状态变化剧烈,统计特性难以预测。虽然目前标准组织定义了 V2X 相关信道模型,但是这些模型有着场景覆盖不丰富、参数不合适等问题。因此,在更多场景下对 V2X 信道开展实地测量与建模是非常有必要的。

1 LTE-V2X 关键技术

1.1 LTE-V2X 无线通信特点

LTE-V2X 无线通信具有以下特点^[10]:

(1) 低延时高可靠通信

在 V2X 场景中,通信系统内个体处于高速移动状态,导致多普勒频移严重,无线传播环境复杂且快时变,网络拓扑具有高度动态性。LTE-V2X 依托现有 LTE 蜂窝通信系统,对物理层、资源分配、同步等关键技术进行改进,具备低延时高可靠的通信能力。

(2) 远距离数据传输可靠性高

相比于 IEEE 802.11p 采用的多跳中继进行远距离数据传输的方式,LTE-V2X 采用基站与云端服务器连接的方式传输远距离数据,可以有效避免中继节点的影响,提高远距离数据的可达性。

(3) 非视距场景传输可靠性高

由于利用了基站转发数据的方式进行通信,且基站天线架设高度一般较高,因此 LTE-V2X 可以很好地支持非视距场景,提高非视距场景下数据传输可靠性。

(4) 网络建设和维护成本低

由于是基于 LTE 蜂窝网络发展而来的,因此 LTE-V2X 可以在进行网络部署时在现有设备上升级扩展,方便快捷地实现 LTE-V2X 的网络建设,同时也可以利用现有商用网络进行安全证书的更新和设备的维护。

由此可见,LTE-V2X 能够最大程度利用已部署网络等资源,具有部署成本低、网络覆盖广的优势,并且在密集的环境下具有更远的通信距离、更大的容量、更佳的非视距通信性能和拥塞控制能力。

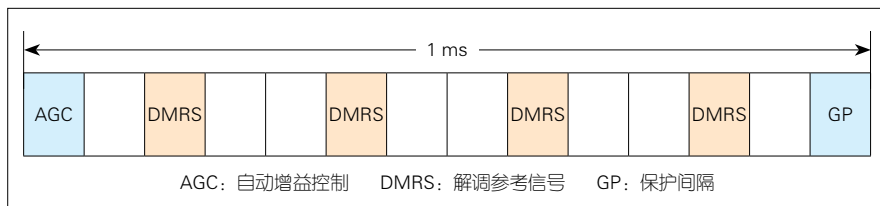
1.2 工作模式

LTE-V2X 针对车辆通信定义了直通模式和蜂窝模式两种通信模式^[10]。直通模式是指车辆之间的直接通信。这种模式引入设备到设备(D2D)PC5 接口,采用 V2X 专用频段,可以使邻近的终端在近距离范围内无需中心节点就可进行通信,从而实现车、路、人之间的短距离直连通信,有效达成

V2X 通信终端之间的低时延、高可靠传输,保证车辆安全驾驶。直通模式适合蜂窝网络覆盖有限的环境,但需要有良好的资源配置及拥塞控制算法。蜂窝模式使终端和基站通过 5G 空口(Uu)接口通信,工作在传统移动通信授权频段,由基站负责资源集中分配与数据转发,以具备集中控制、资源调度、统一协调的功能,可显著提高 LTE-V2X 的接入能力和组网效率。两种模式的协同工作、优势互补,可以使车辆在不同场景中实现低延时、高可靠通信^[11]。

1.3 物理信道设计

LTE-V2X 采用单载波频分多址接入技术,可以有效降低峰均功率比,在相同功放情况下有更大的发射功率。LTE-V2X 在频域上支持 10 ~ 20 MHz 的可变带宽,物理信道设计为子帧、资源块和子信道。在 LTE-V2X 中,如果继续使用传统蜂窝通信的帧结构,车辆的高速移动和高频段工作造成的多普勒频偏就会对信道估计产生严重影响,因此需要对帧结构进行增强设计。如图 1 所示,每个子帧长度为 1 ms,包含 14 个正交频分复用调制(OFDM)符号。其中,第一个符号用于自动增益控制(AGC),最后一个符号作为保护间隔(GP)。将原本子帧中的 2 列解调参考信号(DMRS)增加到 4 列,能够有效处理典型高速场景高频段的信道检测、估计与补偿^[12]。资源块是分配给用户的最小时频资源,在时域上占 1 ms,在频域上占 180 kHz,并包含了 12 个子载波,



▲图 1 基于长期演进的车用无线通信技术帧结构

是所有控制信令和数据信息的基本单元。子信道是具有相同子帧的资源块的组合,用于传输数据信息和控制信息两种信息。这两种信息被安置在同一子帧内,以达到降低时延的目的。

1.4 资源调度

LTE-V2X 支持 Mode 3 和 Mode 4 两种资源分配方式。合理进行资源调度可以有效避免资源碰撞,提高通信的可靠性。Mode 3 基于 Uu 接口,由基站对资源集中控制。用户需要在通信前向基站发送资源分配请求,基站根据用户的地理位置和资源利用情况将资源分配给用户。Mode 3 是一种基于终端地理信息的半持续调度方式。Mode 4 是基于 PC5 接口的“感知+预约的半持续调度”的方案,如图 2 所示。此方案充分利用 V2X 业务的周期性特点。终端会自行选择子信道接入,在资源池中感知资源占用情况。在选择合适的资源后,终端会在这些资源上周期性地发送一定次数,直到触发资源重选。Mode 4 的资源分配过程不需要基站参与,是一个完全自组网的方案,既可以承载周期性的 V2X 业务,又可以利用对资源状态的感知以避免冲突,提高资源利用率和传输可靠性^[10]。

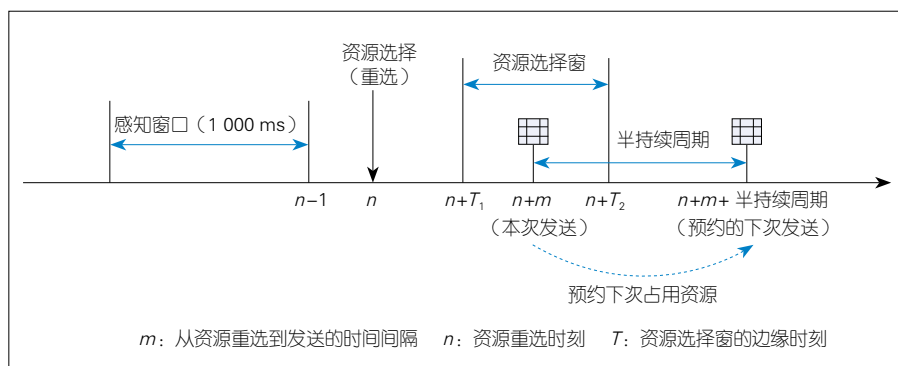
2 典型场景库分类

根据调研,V2X 通信主要涉及 11 个场景:城区直道、城区弯道、城区路口、城区环岛、高速直道、高速弯道、隧道、路堑、地上停车场、地下停车场、三岔路口。

各场景参数如表 1 所示。表中同时给出了各场景的示意图。

3 信道测量与特性分析

无线信道是无线通信系统设计的基础,会在很大程度上影响通信的性能。准确认识 V2X 信道特点是研究



▲图 2 “感知+预约的半持续调度”资源分配机制

▼表 1 典型场景库分类

场景	示意图	场景基本情况
城区直道		多径: 较多; 多普勒: 较强; RSU 部署位置: 路中或路侧
城区弯道		多径: 较多; 多普勒: 较弱; RSU 部署位置: 路中或路侧
城区路口		多径: 较多; 多普勒: 较弱; RSU 部署位置: 路中或路侧
城区环岛		多径: 较多; 多普勒: 较弱; RSU 部署位置: 路中或路侧
高速直道		多径: 较弱; 多普勒: 较强; RSU 部署位置: 路中或路侧
高速弯道		多径: 较弱; 多普勒: 较强; RSU 部署位置: 路中或路侧
隧道		多径: 较多; 多普勒: 较强; RSU 部署位置: 路中或路侧
路堑		多径: 较多; 多普勒: 较强; RSU 部署位置: 路中或路侧
停车场		地面停车场 多径: 较弱; 多普勒: 较弱; RSU 部署位置: 路侧 地下停车场 多径: 较多; 多普勒: 较弱; RSU 部署位置: 路中或路侧
三岔路口		多径: 较多; 多普勒: 较弱; RSU 部署位置: 路中或路侧

OBU: 车载单元 RSU: 路侧单元

V2X 通信的首要问题。信道测量是无线信道传播特征统计和参数提取的前提,有助于充分了解所关注场景的信

道特性。为了测量 LTE-V2X 信道并分析其特性,本文研究组在上海临港智能网联汽车综合测试示范区开展了实

地测量工作。

3.1 测试方案

外场测试地点为上海临港智能网联汽车综合测试示范区。选择了上述 11 个典型场景进行测试。测试系统如图 3 所示,基本参数配置如表 2 所示。

在外场测试中,要保证所使用的频段无其他频率干扰,收发信机之间要保持同步。发射端采用多载波信号序列作为探测信号,接收端使用频谱仪联调。测试过程中,还需要用全球定位系统(GPS)轨迹记录仪记录车辆的行驶状态,便于后续的数据分析与处理。图 4 给出了 V2I 高速直道场景与 V2I 城区直道场景下的测试实景图。

3.2 测试结果展示与分析

下面选择 V2I 高速直道场景与 V2I 城区直道场景进行路径损耗的对比,如图 5 所示。

本文研究组采取最小二乘法对实测路损进行拟合,并与自由空间路径损耗模型、WINNER 模型对比,得到了较好的拟合效果。经计算得知,V2I 与 V2V 的路损值分别为 2.27 与 2.67。城区直道路损值较大,这主要是因为城区直道环境反散射体较多,信号经过反射的次数也较多。

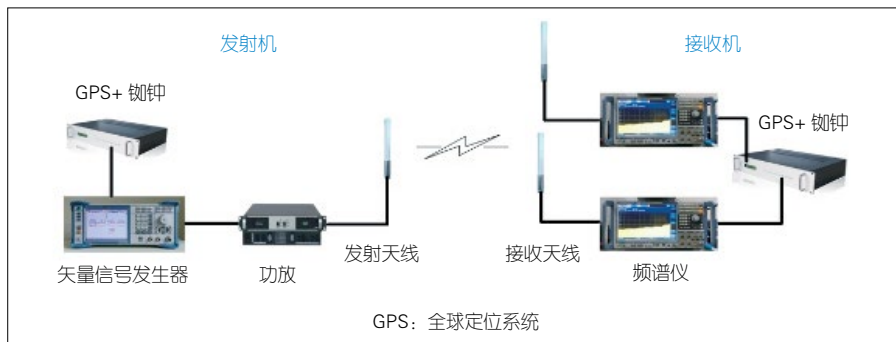
在对 V2I 高速直道场景有效径数进行分析后得到,平均有效径数为 1.81,如图 6 所示;而 V2I 城区直道场景的平均有效径数为 2.87,明显多于高速直道场景。这主要是因为城区场景下较多的反散射体使得发射信号在到达接收信号前发生了更多次的反射,从而产生了更多的多径。

从图 7 中可以看出,V2I 高速直道场景的多普勒功率谱存在快变特性。发生此快变的位置均为收发端距离最近的位置,且存在一定程度的多普勒扩展。

经过对测试数据的分析,我们初

步得出以下结论:对于大尺度衰落,V2V 场景下的路径损耗指数大于 V2I 场景下的;对于小尺度衰落,V2V 场

景的有效多径数大于 V2I 场景下的,且对向行驶场景大于同向行驶场景;无论是 V2V 还是 V2I,不同的测试场



▲图 3 信道测量系统参数配置

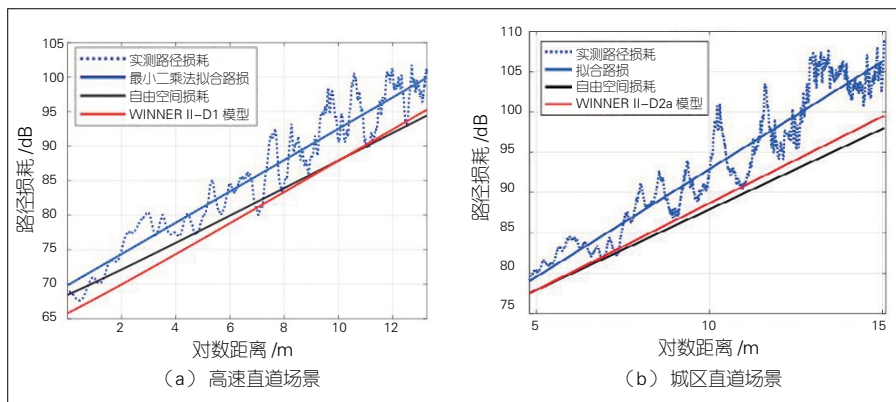
▼表 2 基本参数配置

参数	配置
测量频点	5.91 GHz (V2V), 5.92 GHz (V2I)
测试带宽	10 MHz
发射功率	27 dBm
天线配置	全向偶极子天线, 单发双收模式, 接收端天线间距为 10 cm
天线高度	V2V 及 V2I 中, 车载天线高度为 1.5 m; V2I 中, RSU 天线高度为 6 m
激励信号	多载波信号序列
时钟同步方式	GPS+ 铷钟

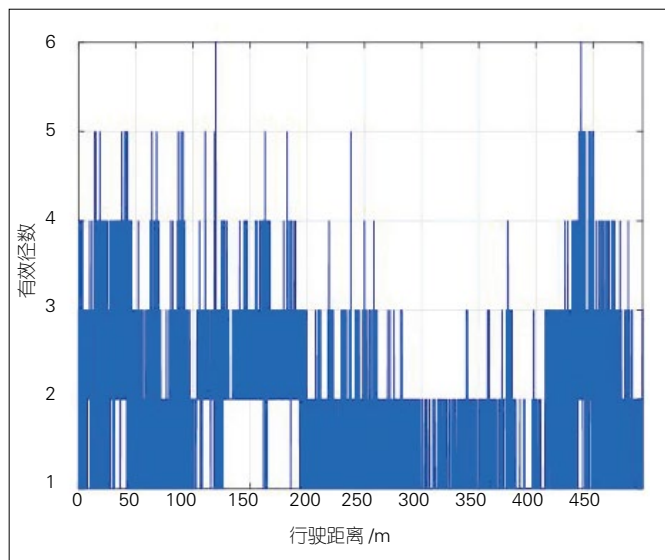
GPS: 全球定位系统 RSU: 路侧单元 V2I: 车与物通信 V2V: 车与车通信



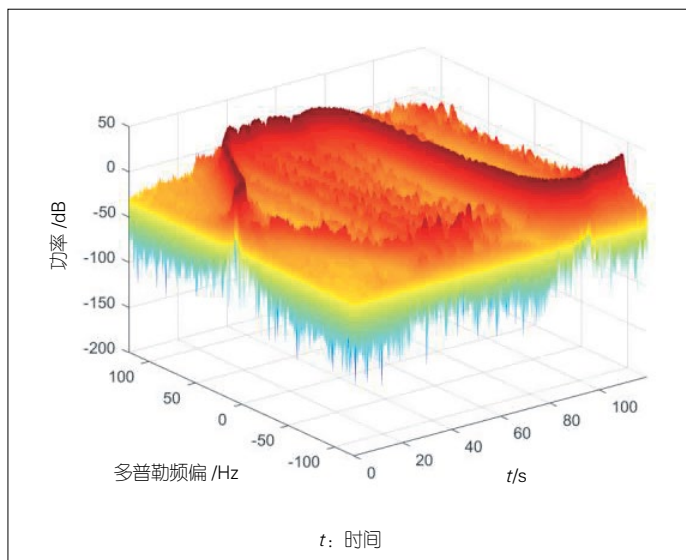
▲图 4 车与物通信高速直道测试场景



▲图 5 车与物通信路径损耗统计分析



▲图6 多径数量统计



▲图7 多普勒功率谱

景得到的结果均有较大差异，需要对每个测试场景进行深入分析。

4 结束语

本文从C-V2X的全球发展态势出发，介绍了LTE-V2X的关键技术、面临的挑战，详细描述了本文研究组在上海临港智能网联汽车综合测试示范区开展的实地测试工作，给出了相应的测试方案，并分析了测试结果。C-V2X作为中国车联网产业发展的关键技术，必将推动中国汽车产业、智能交通等领域的改革创新。由于目前C-V2X无线通信场景的实地测试较少，我们希望本文起到抛砖引玉的作用，以所做的测试为行业同仁开展相关工作提供参考，为车联网产业在中国的大规模部署提供帮助。

致谢

本研究得到了北京交通大学樊圆圆和中国联合网络通信集团有限公司邱佳慧、林晓伯工程师的帮助，谨致谢意！

参考文献

- [1] 陈山枝, 胡金玲. 蜂窝车联网(C-V2X) [M]. 北京: 人民邮电出版社, 2021
- [2] 陈山枝, 时岩, 胡金玲. 蜂窝车联网(C-V2X)综述[J]. 中国科学基金, 2020, 34(2): 179-185. DOI: 10.16262/j.cnki.1000-8217.2020.02.009
- [3] 郭荣斌, 朱永东, 朱凯男, 等. 车路协同C-V2X关键技术演进 [C]// 第十五届中国智能交通年会论文集. 深圳, 2020: 584-596. DOI: 10.26914/c.cnki-hy.2020.028421
- [4] 魏奎, 王庆扬. C-V2X蜂窝车联网标准分析与发展现状[J]. 移动通信, 2018, 42(10): 9-12. DOI: 10.3969/j.issn.1006-1010.2018.10.002
- [5] CHEN S Z, HU J L, SHI Y, et al. A vision of C-V2X: technologies, field testing, and challenges with Chinese development [J]. IEEE Internet of Things journal, 2020, 7(5): 3872-3881. DOI: 10.1109/JIOT.2020.2974823
- [6] 郭蓬, 袁俊肖, 戎辉, 等. LTE-V2X标准分析及发展现状的研究[J]. 中国汽车, 2019, 29(1): 59-62
- [7] 葛雨明. 我国LTE-V2X标准化及测试验证进展[J]. 移动通信, 2019, 43(11): 36-39. DOI: 10.3969/j.issn.1006-1010.2019.11.006
- [8] 崔硕, 姜洪亮, 戎辉, 等. 关于C-V2X的标准组成及研究现状分析[J]. 汽车电器, 2018, (9): 47-48+52. DOI: 10.13273/j.cnki.qcdq.2018.09.016
- [9] 董朔朔, 刘留, 樊圆圆, 等. 车联网信道特性综述 [EB/OL]. (2020-12-15) [2021-04-08]. <http://kns.cnki.net/kcms/detail/41.1185.TN.20201215.1702.001.html>
- [10] 陈山枝, 胡金玲, 时岩, 等. LTE-V2X车联网技术、标准与应用[J]. 电信科学, 2018, 34: 1-11
- [11] CHEN S Z, HU J L, SHI Y, et al. LTE-V: a TD-LTE-based V2X solution for future vehicular network [J]. IEEE Internet of Things journal, 2016, 3(6): 997-1005. DOI: 10.1109/JIOT.2016.2611605
- [12] CHEN S Z, HU J L, SHI Y, et al. Vehicle-to-everything (v2x) services supported by LTE-based systems and 5G [J]. IEEE communications standards magazine, 2017, 1(2): 70-76. DOI: 10.1109/MCOMSTD.2017.1700015

作者简介



苏昭阳, 北京交通大学电子信息工程学院在读博士研究生; 研究方向为车联网通信、信道测量与建模。



刘留, 北京交通大学电子信息工程学院教授、博士生导师; 研究方向为电波传播与无线信道建模、时变信道信号处理、5G关键技术、高铁宽带接入物理层关键技术等; 主持国家自然科学基金、北京市自然科学基金等多项科研课题; 2016年入选北京市科技新星; 发表论文130余篇, 申请专利30余项(授权20余项)。



冯毅, 中国联合网络通信集团有限公司智网创新中心总监、教授级高工, 享受国家级特殊津贴; 长期从事通信行业网络技术研究、建设规划、产品创新研发工作; 带领团队先后完成8项国际标准和9项中国标准; 发表论文10余篇, 获得省部级以上奖励10余项。



5G 电信云网络的容器演进方案

Evolution Solution for Container-Based 5G Telecom Cloud Network

摘要: 以云原生计算基金会 (CNCF) 定义的容器技术为出发点, 从技术方案和客户价值两个维度分析 5G 电信云网络引入容器技术的必要性。讨论了演进路线的兼容性和开放性, 并针对不同的部署场景和云平台新老供应商提出了相关演进方案。基于项目实践, 给出了分阶段部署应用和虚拟网络功能平滑迁移到基于虚拟机容器的云原生网络功能详细过程。

关键词: 5G 核心网; 容器; 云原生; 电信云; 微服务

Abstract: The container technology defined by the Cloud Native Computing Foundation (CNCF) is analyzed, and the necessity of introducing container technology is highlighted from the two aspects of technical solution and customer value. The compatibility and openness of the evolution roadmap are discussed, and relevant migration solutions are proposed for different deployment scenarios and suppliers for new and existing cloud platforms. Based on the project practice, the detailed process of phased deployment of applications and smooth migration of virtual network function to virtual machine container-based cloud native function is put forward.

Keywords: 5G core network; container; cloud native; telecom cloud; microservice

史庭祥 /SHI Tingxiang^{1,2}
徐法禄 /XU Falu²
章璐 /ZHANG Lu^{1,2}

(1. 中兴通讯股份有限公司, 中国 深圳 518057;
2. 移动网络和移动多媒体技术国家重点实验室,
中国 深圳 518057)
(1. ZTE Corporation, Shenzhen 518057, China;
2. State Key Laboratory of Mobile Network and
Mobile Multimedia, Shenzhen 518057, China)

DOI: 10.12142/ZTETJ.202106011

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20211119.1850.002.html>

网络出版日期: 2021-11-22

收稿日期: 2021-11-08

云计算从信息技术 (IT) 发展到通信技术 (CT), 网络功能 (NF) 从虚拟化演进到云化和云原生, 都经过了 10 余年的历程。2006 年 8 月 9 日, ERIC SCHMIDT 在搜索引擎大会上首次提出了“云计算”的概念。随后, 云计算在 IT 市场被广泛使用。

云原生是 MATT STINE 在 2013 年提出的一个概念, 具有良好的可伸缩性。目前, 云原生应用和服务受到了越来越多的关注。云原生是一个思想的集合, 既包含技术 (如微服务、敏捷基础设施), 也包含管理 (如 DevOps、持续交付、康威定律、重组等)。可以说, 云原生是一系列云技术和企业管理方法的集合。

为统一云原生接口标准, 2015 年 7 月云原生计算基金会 (CNCF) 成立。该基金会的会员包括 Google、IBM、Intel、Docker、RedHat 等国际著名公司, 代表了业界最领先的容器技术和云计算技术。CNCF 致力于通过技术优势 and 用户价值创造一套新的通用容器技术,

以推动云计算和相关服务的发展。

CNCF 中最知名的项目是 Kubernetes (以下简称 K8S)。K8S 是 Google 发起并维护的基于 Docker 的开源容器集群管理系统。K8S 容器集群能在私有云、公有云或者混合云之上运行。K8S 属于主从的分布式集群架构, 包含 Master 和 Node。其中, Master 作为控制节点, 可以调度和管理整个系统; Node 是运行节点, 可以运行业务容器。

根据 CNCF 对云原生的定义, 云原生网络功能 (CNF) 和 K8S 的设计与架构应满足相应要求^[1]。为此, 基于容器的云原生应用, 通常需要满足几个条件: 具有微服务体系结构、可拆分性和独立的业务逻辑, 不仅支持独立扩展, 还支持 K8S 应用程序接口 (API), 即具有 K8S 感知的网络功能。

1 电信云网络在 5G 阶段引入容器技术的必要性

随着全球 5G 网络建设浪潮的到

来^[2-5], 电信网络中以 5G 独立组网 (SA) 核心网为代表的业务网元, 需要引入容器技术。

1.1 第 3 代合作伙伴计划 (3GPP) 协议定义的服务化架构 (SBA) 和相关新特性

引入容器技术, 不仅可以满足 CNCF 提出的要求, 还能满足 Vodafone 在定义云原生时提出的容器技术需求: 微服务架构和开放接口, 其中开放接口用于微服务之间的通信; 轻量化虚拟技术, 如容器或其他轻量化云平台; 以微服务为单元提供网络功能 (NF) 生命周期管理, 包括初始化、弹缩、自愈等。

此外, 3GPP R15 标准满足增强移动宽带 (eMBB) 商用要求, 和 R16 一起为 5G 核心网定义 SBA, 并支持如下几方面的新特性^[6-12]。

(1) 微服务。5G SA 核心网重新定义各个 NF。大多数 NF 之间采用服务接口 (SBI), 同时 NF 采用微

服务架构, 以便支持无中断在线升级 (ISSU)。

从 2020 年上半年某运营商的 5G SA 核心网招标项目可知, SBA 架构、微服务组件、无状态设计、网络切片、自动化部署、2G/3G/4G/5G 全接入和全融合等需求, 必须以微服务为特性要素, 才能实现灵活敏捷的业务创新、网络的部署和运维, 以及资源的重复利用, 实现低成本、高价值的 5G 网络目标。

(2) 跨平台。无论是以 R15 标准为主的 eMBB 场景, 还是 R16 和 R17 正在规划的超可靠低时延通信 (URLLC) 和海量机器类通信 (mMTC) 场景, 抑或是公有云和私有云之间^[13], 都使得 5G 核心网不仅要依托现有云化基础设施, 还要按需混合部署在虚拟机、容器 (如虚拟机容器、裸金属容器等) 之上, 因为只有这样才能最大程度地实现资源共享, 节省基础设施投资成本^[14]。

此外, 跨平台部署的 5G 网络强调自动化和智能化。将不同平台和不同云架构的应用串联起来, 有助于实现诸如业务一键开通、新业务快速上线等功能。

(3) 易演进。为进一步提升资源利用率, 基于 SBA 架构的 5G 核心网能够提供服务化接口, 以便于合作方调用业务功能, 实现 NF 和网络功能服务 (NFS) 共享。依照 5G 标准和产业逐步发展的特点, 架构设计须侧重前瞻性。比如, 公共服务先期以 R16 基于服务的架构增强 (eSBA) 的服务框架呈现, 以降低长期投资成本^[15]。

1.2 从 5GC 技术落地实践中挖掘容器技术价值

虽然容器技术不是 SBA 架构和上述新特性的必选, 但是容器技术仍具有一些突出优势。

1.2.1 微服务要求每个 NF 的各个组件以更小的粒度占有资源

微服务之间通信协议采用的是面向服务的开放接口超文本传输协议版本 2 (HTTP2), 原有的虚拟机资源分配方式无法满足系统要求, 为此引入容器技术成为必然选择。

此外, 容器化 NF 不仅可以使生命周期管理 (LCM) 变得更加高效, 还能使自动化编排占有更少的资源。这些都将显著提升 NF 和 NFS 的服务水平, 改善网络性能, 提升用户体验, 增强对突发性业务的响应能力。

1.2.2 跨平台要求 NF 部署在各种云平台并支持多种部署方式

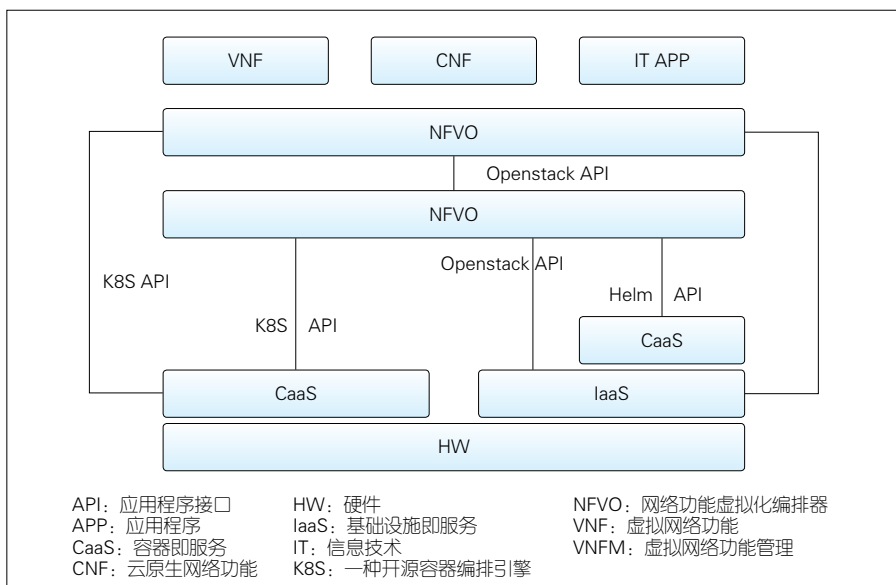
这里的云平台主要指与容器相关的云平台。在部署方式上, 对边缘数据中心 (DC) 有更高的时延要求, 也有更低成本和低功耗的设备要求。相应地, 只有采用轻量级的虚拟化技术代替传统虚拟机资源, 才能满足这些要求。容器平台是典型的轻量级虚拟化技术, 已广泛应用于 IT 领域。

对运营商网络而言, 虽然当前 NF 仍然以 VNF 形式进行部署, 但已处于

CNF 等容器应用的初期阶段。如图 1 所示, 结合近期全球运营商的需求, NF 的部署形式主要有三大类: VNF、CNF 和 IT 应用程序 (APP)。这 3 种部署形式对应 3 种云平台。基础设施即服务 (IaaS) 平台: 如主流的虚拟化管理系统 Openstack 和 VMware, 即虚拟机平台; 容器即服务 (CaaS) 平台: 如主流容器管理系统 K8S, 即容器或裸金属容器平台; IaaS+CaaS 平台: 容器应用包装在虚拟机资源池里运行, 即虚拟机容器平台。这 3 种平台支撑的 NF 形式分别为 (1) IaaS: VNF、IT APP; (2) CaaS: CNF、IT APP、部署在 CaaS 上的 VNF; (3) IaaS+CaaS: CNF。

1.2.3 易演进要求容器技术为基于 SBA 的 NF 和共享公共服务提供强大支撑

一方面, 更小规格的容器化 NF 有利于更为灵活的组件资源划分和动态调整。例如, 不同使用频度的 NFS 将采用不同的资源占用粒度: 高频使用的 NFS 采用规格较小的容器部署, 低频使用的 NFS 采用规格较大的容器。不同使用频度的 NFS 将具备各自的资



▲图1 多种云平台共存

源调整能力，易于业务快速扩展，还能兼顾资源管理开销和效率。

另一方面，更小规格的容器化 NF 便于将业务组件和基础组件分开配置资源。以往以大规格虚拟机方式部署的 NF，像没有应用集装箱技术的货物运输一样，资源使用率不高。

为更加清晰地表明以虚拟机方式部署的 VNF 和以容器方式部署的 CNF 之间的差别，如图 2 所示，我们做出如下几种假设：

(1) $VNF = VNFC1 + VNFC2$

其中，虚拟网络功能组件 (VNFC) 是 VNF 的一个组成单元。比如，操作维护管理单元 (OMU) 是一个 VNFC，通用业务处理单元 (GSU) 是另一个 VNFC，不同的 VNFC 有不同的特性。

(2) $NF = NFS1 + NFS2$

NFS 是 NF (或 CNF) 的一个组件。每个 NFS 都由一组相关的业务组件 (SC) 构成，包括操作维护管理 (O&M)、元数据管理等。相应地，对于归属某一类 SC1 的多个 SC，假设有 SC1-1、SC1-2 等；对于归属某一类 SC2 的多个 SC，假设有 SC2-1、SC2-2 等。于是有 $NFS1 = SC1-1 + SC2-1$ 、 $NFS2 = SC1-2 + SC2-2$ 。

(3) VNFC 采用虚拟机方式部署，SC 以 POD (K8S 定义的最小部署单元) 方式部署。假设每个 POD 都有一个容器，即最常见的单容器 POD。

为便于比较，设 SC1 和 VNFC1 是一类功能组件，SC2 和 VNFC2 是另一类功能组件。假设有两个节点各自分配虚拟机资源，并各自部署一类 SC。那么，VNFC1 对应容器化 NF 的功能由 SC1-1 和 SC1-2 实现，VNFC2 对应的功能由 SC2-1 和 SC2-2 实现，即两种方式的资源消耗相同。

容器方式的好处是基于更小粒度的 SC/POD 进行调度。其中，POD 是 K8S 中能够创建和部署的最小单元，是 K8S 集群中的一个应用实例。而虚拟机方式则基于 VNFC 实现调度。显然前者的调度效率更高，虚拟化管理更加精细化。此外，NF 以微服务方式对外呈现业务，以消息总线呈现接口。因此，基于容器技术的 NF 更容易发挥优势。

此外，容器技术应用于 NF 还具有其他突出优势：(a) 快速交付和部署，比如虚拟机容器方式可以采用已有的标准镜像包，并以容器方式部署，具备快速启动、扩容和缩容的特点，可大幅度节约时间；(b) 容器技术是

面向内核的虚拟化技术，不需要额外 Hypervisor 的介入，其虚拟化效率更高，需要额外开销的资源更少；(c) 基于共享 Linux 内核技术，自动化部署和维护更加简便。

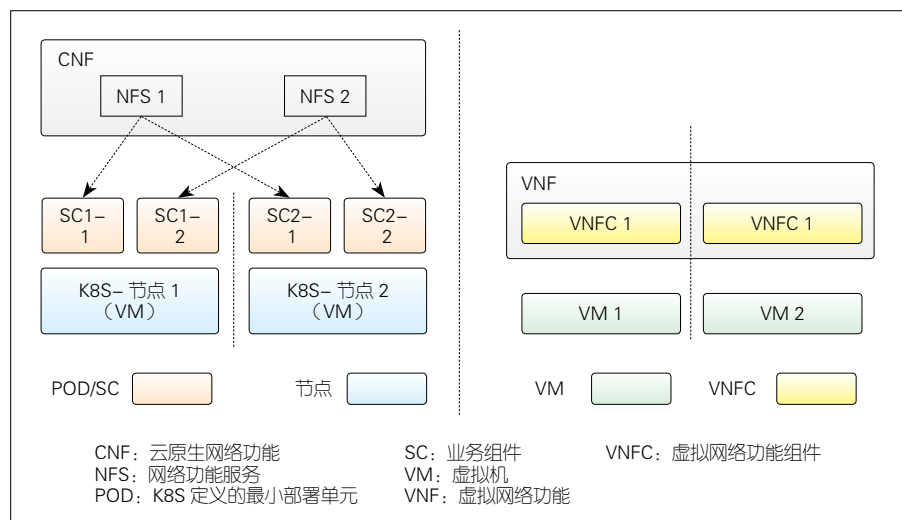
1.2.4 在 NFS 占有资源方面容器技术更具优势

一般而言，包含访客操作系统在内，采用虚拟机方式部署的中央处理器 (CPU) 消耗为 0% ~ 15%，内存消耗为吉比特量级，镜像文件规模在兆比特到吉比特之间。容器方式则可以共享 Linux 内核且通过进程隔离来实现微服务化的应用。容器方式的 CPU 消耗在 0% ~ 5%，内存消耗在兆比特量级，镜像文件规模在千比特到兆比特之间。图 3 给出了虚拟机架构和容器架构的对比。

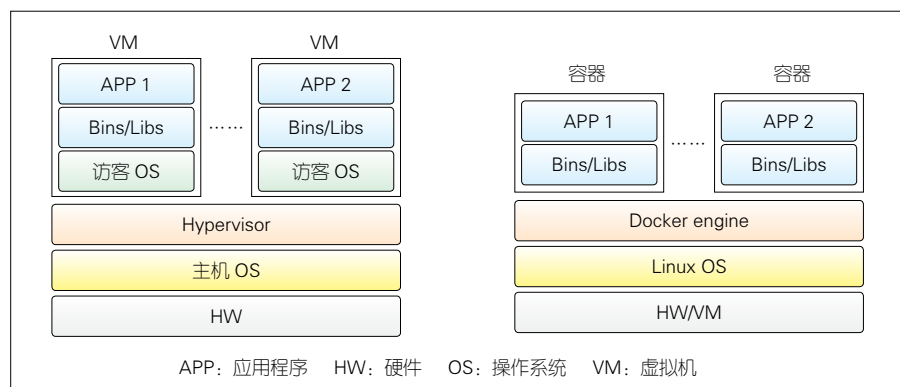
2 容器演进路线

无论是 CNCF 定义的云原生，还是 3GPP R15、R16 等版本定义 5G 核心网所需的 SBA 架构、微服务、跨平台等，容器技术必然被纳入 5G 电信云网络的技术规划中。电信云从虚拟化到云化，承载着 4G 网络改造的两大诉求：如何实现弹性、智能、可管可控的网络？在流量增收不增利的情况下，如何实现低成本建网和高效运维？

在 5G 阶段，电信云不仅要满足网络从 4G 向 5G 发展的业务侧要求，还要进一步提升网络性能，支撑 eMBB、mMTC、URLLC 等多场景需求，并完成“一云多构、一网多制”的转变。从 2G 到 5G 的发展过程看，无论是不断丰富的业务功能和日益完善的用户体验，还是网络侧从会话初始协议 (SIP) 到应用程序的认证、鉴权、计费框架协议 (Diameter)，都证实了一点：兼容和开放是电信网络发展的永恒主旨。



▲图 2 VNF 和 CNF 的结构差异



▲图3 虚拟机架构和容器架构

结合中国运营商的5G发展思路，我们从如下两个角度分析如何从目前基于虚拟机的VNF演进到基于容器的CNF。

(1) 兼容性

演进过程要考虑原有VNF长期共存的要求，这里我们以5G核心网为例进行说明。一方面，虽然分组数据网的网元将首次应用在5G网络，并定义微服务化NF，如接入管理功能（AMF）、会话管理功能（SMF）、用户面功能（UPF），但是相关协议对电路域（CS）和IP多媒体子系统（IMS）的定义尚未完成或成熟度不足。另一方面，存量VNF迁移到CNF，并没有形成平滑改造的共识。

(2) 开放性

容器应用在IT领域已成为主流，但是在电信领域还没有得到广泛应用。目前，从虚拟机向容器演进的方式有

虚拟机容器和裸机容器两种。虚拟机容器方式下CNF的编排接口没有纳入协议规范，同时供应商对裸机容器方式的具体实现方式存在争议。为此，支撑从VNF向CNF演进的容器技术、云管理系统和编排器，都需要具备充分的开放性，以避免日后在引入新应用或新技术时，面临重新部署甚至重构整个系统的被动局面。

当前，中国三大运营商已规模部署5G核心网（5GC）网络设备，同时各种5GC应用以VNF形式部署。如图4中的初期所示，各厂家即使引入容器技术也无须对外呈现，即不必对外提供容器管理功能。尽管如此，VNF方式仍有诸多优点：标准程度、商用成熟度和虚拟机隔离方案的安全性都很高，拥有成熟的转发面优化技术等。基于目前成熟的IaaS云平台方案和产品，VNF方式有利于5GC网络建设初

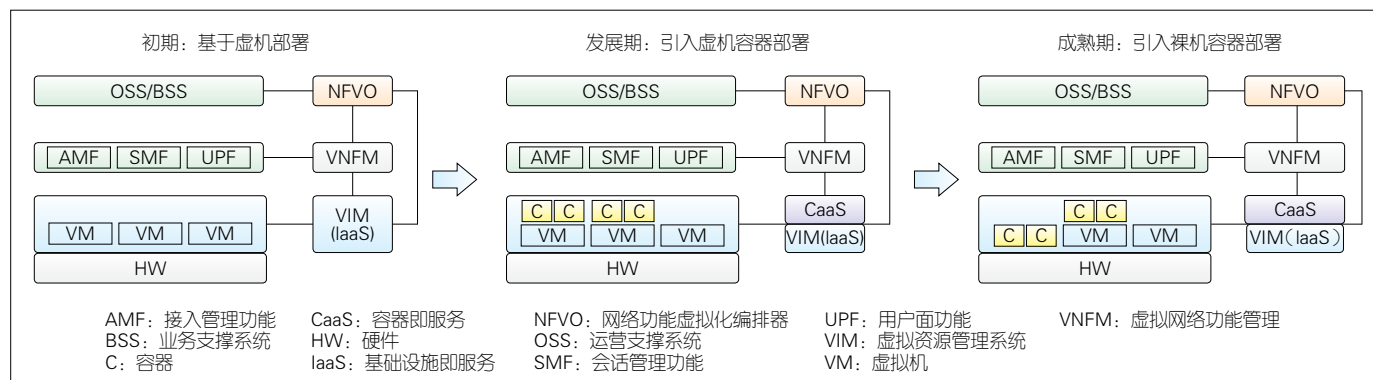
期的快速部署和运营。然而，没有引入开放的容器技术且不满足CNF要求的可感知K8S接口，是基于虚拟机方式部署VNF的最大不足。

如图4所示，发展期引入了虚拟机容器方案。该方案既支持容器化部署，满足运营商对容器的管理需求，又拥有虚拟机的安全能力，同时可应用于那些对安全隔离有较高需求的场景。影响该阶段快速发展的因素主要有两个：（1）相关标准尚不成熟，如欧洲电信标准化协会（ETSI）协议没有定义K8S容器管理系统和管理和网络编排（MANO）的接口；（2）IaaS云平台、硬件和APP的三方解耦转变为含有CaaS云平台的四方解耦，这使技术的选择难度和验证难度变得更大，商业成熟期变得更久。

随着容器应用进入成熟期，灵活的应用、弹性的网络和多层次的安全管控都将裸机容器平台推向主流。如2G/3G/4G和5G长期共存一样，该阶段仍要兼顾多种NF部署方式，最终才能实现“一云多构”的综合云平台。

3 容器演进方案

如前文所述，电信云网络从虚拟机部署转变为容器部署，需要经历多个阶段。从需求的角度看，业务网元在前，支撑它的云平台在后；从网络建设和演进的角度看，云平台基础设



▲图4 容器技术引入5G核心网的过程设想

施建设在前,相应的业务网元部署在后。由此可见,业务网元和云平台之间将相关作用、互相影响。

在云平台长期演进过程中,支持容器的云平台同时也要支持原有的虚拟机部署方式。容器部署包括虚拟机容器、裸机容器、未来“容器化的 VNF”等方式(VNF over Container)。对于业务网元来说,从基于虚拟机部署发展到基于容器部署是一个渐进过程,它要求虚拟机部署的网元逐步迁移到容器平台上。

按照容器平台提供者和虚拟机平台提供者的异同,以及部署场景的不同,我们提出如下容器演进方案。

3.1 边缘云部署先行

围绕 5G 核心网对电信云的新需求,比如控制面和媒体面分离(CUPS),媒体面下沉将带来本地分流和业务时延减少。为此,将 MEP、UPF 等网元下沉到边缘,意味着边缘云网络建设必将先行。

与中心 DC 相比,边缘 DC 的边缘云具备 3 个特点。(1)可改善系统性能,如具有硬件加速功能。(2)拥有多样性的云管理系统和 SDN^[16-17]。不同于重量级的中心 DC、功能丰富的云管理系统和 SDN,边缘云注重高效、轻量 and 易管理。(3)软硬件倾向高度集成,便于快速安装、上线和扩容。

因此,容器化应用和容器平台大多首先部署在边缘云。如图 5 所示,边缘云部署包括 3 个部分:

(1) 新建边缘云网络和相关应用

边缘云硬件基础设施和云管理系统应满足几个方面的需求:提供专有硬件支撑的硬件加速方案,以优化媒体面性能;制定虚拟化加速方案,以提升应用面的包转发能力;具有虚拟交换软件(OpenvSwitch)+数据平面开发工具集(DPDK)卸载能力,以

便采用专有硬件来优化 CPU 性能;增加 MANO 和网元管理系统(EMS)对专有加速硬件的统一编排和管理能力;云管理系统能够推荐紧凑型的容器平台或者虚拟机和容器融合的综合云平台。

(2) 实现边缘云和中心 DC 的云网络互通

中心 DC 的云网络和边缘云需要分开部署,拥有各自独立的交换网络、SDN、DC- 分组数据网关(GW),以实现业务面的互联互通。同时网络功能虚拟化编排器(NFVO)和 EMS 功能将纳入全网统一管理。

(3) 扩容中心 DC 的云网络

在边缘云部署的业务网元需要与中心 DC 的业务网元协同工作。例如,AMF 和 SMF 仍然需要部署在中心 DC 上,同时 SMF 能够按不同策略选择分

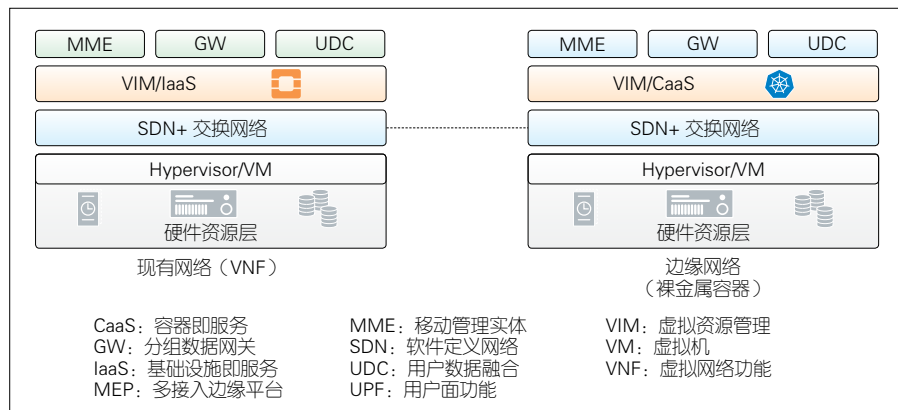
布式部署的 UPF。NFVO 和 EMS 功能也将纳入全网统一管理。

3.2 中心 DC 容器演进方案兼顾现网 VNF

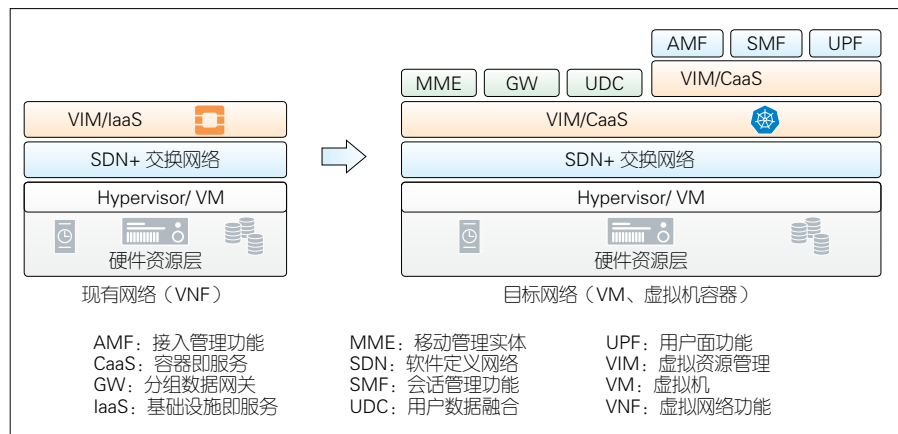
引入 CaaS 容器平台的方式不同,产生的容器演进方案也会不同。一种方式是在现有 IaaS 资源池部署容器应用;另外一种方式是,首先为容器平台和应用新建一个独立的资源池,然后在初期与现有 IaaS 虚拟机平台共存,最终实现虚拟机与容器融合的综合平台。

(1) 依托现有 IaaS 资源池的演进方案

将中心 DC 的虚拟机部署方式改造成虚拟机容器混合部署方式的高效方案是升级现有 IaaS 云平台。这种方案尤其适用于 IaaS 和 CaaS 平台均来自



▲图 5 边缘云部署



▲图 6 中心 DC 容器云网络依托现有 IaaS 资源池演进

同一个供应商的情况,如图6所示。

升级方案具体包括:(a)升级 IaaS 云平台和 SDN 支持 5GC 容器应用,如支持加速硬件和交换网络;(b)如果步骤 a 不能对 VNF 可见,则 VNF 将重新上电或进行重新部署;(c)资源池扩容,如部署 5GC 业务网元或扩容 4G 网元;(d)部署 CaaS 容器平台,如果是异厂家的容器平台,建议采用 IaaS 平台不感知的容器方案;(e)部署 5GC 容器应用。

(2) 新建容器资源池

一般而言,现有 IaaS 平台承载的 4G 业务是客户的收入来源。当 5G 网络处于初期或实验阶段时,在成熟网络上叠加 5G 网络并不是最佳选择。5G 网络所需要的硬件和网络基础设施与 4G 相比有显著差别,如支持 URLLC 的高性能硬件、切片网络所需的特殊软硬件等。此外,当容器平台来自新的供应商时,新建容器资源池具备一定优势。

与新建边缘云类似,5G 应用建设需要独立的容器云网络。这里我们以虚拟机容器方案为例进行说明。考虑到 4G 和 5G 融合部署的需求,5G 核心网要支持回落到 4G 网络的 5G 用户仍可以接入同一张核心网,即实现业务网元在新建容器云和现网的云网络之间的按需分布。

由于统一云平台需求有利于降低运维和新业务部署的费用,容器技术下一步演进方向是在容器云的基础上实现原 VNF 的迁移。这里迁移方式包括两种:现有 VNF 迁移到容器云上,保留原有 VNF 部署方式;基于已有虚拟机容器混合云或已部署的 CNF 来实现 VNF 功能的迁移。

第 1 种迁移方式适合容器云和现有 VNF 不是同一个供应商的情况。VNF 供应商希望在尽量不感知的情况下将 VNF 迁移到新平台,以便减少迁

移费用并维持业务平滑过渡。

第 2 种迁移方式一般用于第 1 种迁移方式不可行或者现有 VNF 已老旧甚至过保的情况。将现有 VNF 合并到容器云上的 CNF,有利于降低运维费用,提升资源利用率和运维效率。对于第 2 种方式,我们推荐使用裸机容器,以便 VNF 彻底迁移到 CNF。

4 容器演进实践

众所周知,目前业界的虚拟化核心网的网元主要以虚拟机方式部署。如何从 VNF 迁移到 CNF,成为运营商最为关心的话题之一。以 VNF 向虚拟机容器方案演进为例,综合考虑 2G/3G/4G 核心网向 5G 演进的相关设计和步骤,我们在本节说明容器演进的实践情况。

4.1 应用演进阶段

第 1 阶段:5G 非独立组网(NSA)部署阶段。这一阶段的诉求是:在不剥离 VNF 的情况下升级 VNF,以支持 5G NSA 网络。3GPP R15 定义的 NSA 网络已满足大部分 eMBB 业务需求。在不改变现有 IaaS 平台、不引入容器平台的情况下,平滑升级 VNF 是支持 5G NSA 的最佳方式。这样做的原因是,部署基于 R16 的 5G SA 网络相当于对现网的业务网元进行重构,无法在兼顾投入和收益的情况下满足不同国家和不同运营商的需求。

第 2 阶段:5GC 初期部署阶段。这一阶段具体包括:(a)基于虚拟机容器方式部署一些 2G/3G/4G/5G 融合的 CNF,具体包括分组域(PS)、用户数据管理(SDM)、策略控制功能(PCF)/策略与计费规则功能(PCRF)等。此时,原 IaaS 平台要改造成 IaaS 和 CaaS 的混合平台;(b)保持其他 VNF 不变,例如 IMS、CS 等;(c)升级 EMS,同时兼顾 VNF 和 CNF 管理。

第 3 阶段:5GC 大规模部署阶段。

(a)对 CNF 进行扩容以支持更大容量,同时引入新 CNF 以支撑大规模部署,并增强网络灵活性,如安全边缘保护代理(SEPP)和服务通信代理(SCP);(b)将 STP、DRA 和基于位置的服务(LBS)等从 VNF 迁移到 CNF;(c)由于没有进一步的业务功能需求,CS 仍采用 VNF 方式部署。

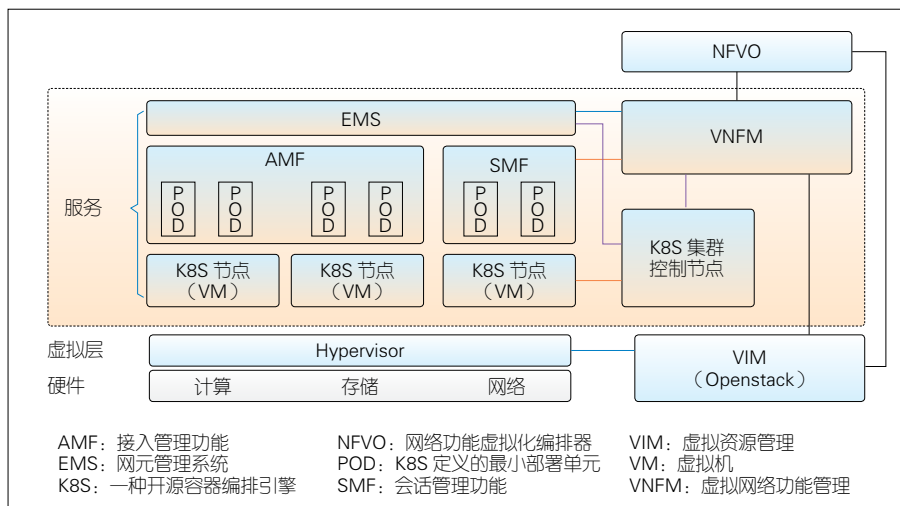
4.2 从 VNF 演进到 CNF 的具体步骤

CNF 是基于虚拟机容器的应用,而 VNF 只是基于虚拟机的应用,没有容器。对于 NF 而言,从 VNF 到 CNF 的迁移过程类似于把“大象”装进“冰箱”。下面我们将简单描述如何完成从 VNF 到 CNF 的迁移过程。

如图 7 所示,VNF 和 CNF 的架构与组件差异很大,没法通过软件升级方式实现业务功能迁移。这需要先在 IaaS 平台实例化 NF,然后再实现 VNF 的业务功能。具体操作步骤大致有 4 个。

(1)新建控制节点:提供支持基于虚拟机容器的 CNF IaaS 平台,如支持 5GC 和虚拟机容器应用的 OpenStack 平台。

(2)计算节点实现 NF 迁移,具体步骤包括:备份节点的 IaaS 平台数据和 VNF 配置数据;在空余硬件上装载支持 CNF 的 IaaS 平台,如 Hypervisor/ 分布式虚拟交换机(DVS)等;在管理节点上装载支持 CNF 版本的 MANO,并打通原 VIM 接口;在空余硬件上初始化含 K8S 组件的 CNF,使业务配置数据与 VNF 相同,同时打通相关组件的接口,在业务调试成功后投入 CNF 资源池,使 NF 层面和 VNF 保持负荷分担或容灾关系;将某 VNF 的业务功能迁移到该 CNF 上;释放原 VNF 占有的硬件资源;重复上述步骤,直至完成全部计算节点的 NF



▲图7 中心DC容器云网络依托现有IaaS资源池演进

升级。

(3) 管理节点升级, 具体包括: 备份管理节点的数据配置; 建立新的管理节点; 从备份数据恢复数据配置; 完成新管理节点和全部CNF节点的业务调试; 原管理节点下电, 以释放资源。

(4) SDN和交换网络升级。一般认为交换网络(如SDN)和IaaS平台有接口, 无须和CaaS平台开通接口。构建CNF资源池的工作包括接管原VNF的交换功能, 这使得SDN和交换网络仍然有升级的需求: 升级SDN以支持CNF; 升级交换网络设备(如DC-GW、Spine和Leaf); 更新SDN和交换网络设备的管理接口。

5 结束语

随着基于NFV的5G核心网规模商用, 采用IaaS平台和虚拟机部署方式的VNF网络已成为5G SA网络建设的开端。VNF网络将支撑以eMBB为主的2C应用和以URLLC/mMTC为主的2B应用。然而, 基于容器技术的5G电信云网络将促使5G核心网应用产生更多价值。从目前全球IaaS网络建设的现状来看, 虚拟机容器方案是5G电信云网络演进的重点。同时, 从虚拟

机部署到虚拟机容器部署的演进具有的更好平滑性。本文中, 我们结合实践案例, 从应用、云平台和硬件等多个角度阐述5G电信云网络演进过程的要点和详细步骤, 希望为广大电信网络和技术工作者提供一些参考。

参考文献

- [1] CNCF. What is cloud native? [EB/OL]. (2018-06-11)[2021-11-06]. <https://www.cncf.io/about/faq/#what-is-cloud-native>
- [2] 陈佳媛, 王瑞雪, 班有容, 等. 中国移动面向5G的电信云基础设施技术研究和实践[J]. 移动通信, 2019, (1): 57-62
- [3] 陆平, 李建华, 赵维铎. 5G在垂直行业中的应用[J]. 中兴通讯技术, 2019, 25(1): 67-74. DOI: 10.12142/ZTETJ.201901011
- [4] 李珊, 张春明, 汪卫国. 5G商用起步, 融合应用蓬勃兴起[J]. 中兴通讯技术, 2019, 25(6): 2-7. DOI: 10.12142/ZTETJ.201906001
- [5] 陈亿根, 尹晓峰, 邵黎勋. 5G+工业互联网应用实践[J]. 中兴通讯技术, 2020, 26(6): 2-6. DOI: 10.12142/ZTETJ.202006002
- [6] ETSI. Network Functions Virtualisation (NFV); use cases: ETSI GS NFV 001 [S/OL]. (2018-07-03)[2021-11-06]. <https://www.etsi.org/technologies-clusters/technologies/nfv>
- [7] ETSI. Network Functions Virtualisation (NFV); architectural framework: ETSI GS NFV 002 [S/OL]. (2019-12-02)[2021-11-06]. https://www.etsi.org/deliver/etsi_gs/NFV/001_099/002/01.02.01_60/gs_NFV002v010201p.pdf
- [8] ETSI. Network Functions Virtualisation (NFV); terminology for main concepts in NFV: ETSI GS NFV 003 [S/OL]. (2019-12-02)[2021-11-06]. https://www.etsi.org/deliver/etsi_gs/NFV/001_099/003/01.04.01_60/gs_NFV-003v010401p.pdf

- [9] ETSI. Network Functions Virtualisation (NFV); virtualisation requirements: ETSI GS NFV 004 [S/OL]. (2019-12-02)[2021-11-06]. https://www.etsi.org/deliver/etsi_gs/NFV/001_099/004/01.01.01_60/gs_NFV004v010101p.pdf
- [10] ETSI. Network Functions Virtualisation (NFV); infrastructure overview: ETSI GS NFV-INF 001 [S/OL]. (2019-12-02)[2021-11-06]. https://www.etsi.org/deliver/etsi_gs/NFV-INF/001_099/001/01.01.01_60/gs_NFV-INF001v010101p.pdf
- [11] 3GPP. System architecture for the 5G system (Release 15): 3GPP TS23.501 [S/OL]. (2019-12-02)[2021-11-06]. <https://www.3gpp.org/>
- [12] 3GPP. Procedures for the 5G System (Release 15): 3GPP TS 23.502 [S/OL]. (2019-12-02)[2021-11-06]. <https://www.3gpp.org/>
- [13] 史庭祥, 田会芹. 云化网络多租户改造方案分析和实践[J]. 邮电设计技术, 2020, (4): 80-84
- [14] 史庭祥, 田会芹. 电信基础网络共享技术研究和实践[J]. 电信网技术, 2017, (7): 40-45
- [15] 史庭祥, 田会芹. 无线核心网的TCO分析方法研究[J]. 中兴通讯技术, 2016, 22(1): 50-53. DOI: 10.3969/j.issn.1009-6868.2016.01.013
- [16] 刘旭, 李侠宇, 朱浩. 5G中的SDN/NFV和云计算[J]. 电信网技术, 2015, (5): 1-4
- [17] 张朝昆, 崔勇, 唐嵩嵩, 等. 软件定义网络(SDN)研究进展[J]. 软件学报, 2015, 26(1): 62-81

作者简介



史庭祥, 中兴通讯股份有限公司高级工程师; 主要研究方向为云计算、核心网、虚拟运营及其关键技术; 发表论文10余篇, 获发明专利10余项(国际专利2项)。



徐法禄, 中兴通讯股份有限公司系统产品MKT及方案部副部长; 拥有多年通信行业的从业经验, 曾从事CDMA、FDD LTE、5G等产品的研发和规划; 获得国家科技进步奖二等奖、深圳市科技进步奖一等奖、IF设计奖、GTI移动业务应用创新奖、IF红点奖等奖项。



章璐, 中兴通讯股份有限公司电信云与核心网产品线产品规划总工、高级工程师; 研究方向为电信云与核心网的组网及其关键技术; 发表论文10余篇, 拥有专利10余项。



基于对称密钥算法的 5G 网络 防伪基站技术

Anti-Fake Base Station in 5G Network Based on Symmetric Cryptography Algorithm

摘要: 分析了 5G 网络场景下可能的伪基站攻击方式,提出了一种基于对称密钥算法的伪基站防护技术,包括密钥协商和传递流程、保护区信息提供流程和基站小区的真实性验证流程。设计了用于加密信息的密钥的产生、传递和使用方法,以及用于验证小区广播或单播消息真实性的密钥产生方法,实现了针对小区发送的广播消息或单播消息的真实性验证,达到了伪基站识别的效果。

关键词: 伪基站; 对称密钥算法; 保护区; 共享根密钥组

Abstract: The possible attack methods of the fake base stations in 5G network are analyzed. A technical solution for anti-fake base stations based on symmetric key algorithm is proposed, including protection key agreement and transmission procedure, protection area information provision procedure, and cell authenticity procedure. The generation, transmission, and use of key used to encrypt information are designed, and the key generation method used to verify the authenticity of the cell is also designed, realizing the authenticity verification of the broadcast or unicast message sent by the cell to achieve fake base station recognition.

Keywords: fake base station; symmetric key algorithm; protection area; shared root key group

毛玉欣 /MAO Yuxin
游世林 /YOU Shilin
闫新成 /YAN Xincheng

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

DOI: 10.12142/ZTETJ.202106012

网络出版地址: <https://kns.cnki.net/kcms/detail/34.1228.TN.20200728.0940.002.html>

网络出版日期: 2020-07-28

收稿日期: 2020-07-10

移动通信经历了从 1G 到 4G 的发展过程,如今正大力发展 5G,以实现万物互联。传统移动通信主要用于满足人们语音通信的需求。4G 让人们享受到无处不在的高速网络接入,5G 的到来将使移动通信网络赋能于更广泛的行业。通信技术的发展使得网络承载的信息和价值越加丰富,但由此引发的网络安全风险也随之增长。2G 网络和终端间采用单向认证,使终端无法对网络进行合法性验证。这导致了伪基站问题的爆发^[1]。随着通信技术的升级换代,终端和网络间引入双向认证,使网络安全性得到增强。然而,伪基站问题仍然无法完全避免。例如,攻击者通过干扰终端的 3G 和

4G 通信信号,使得手机无法正常使用 3G、4G 通信信号。为了保障通信,终端就会主动降频到 2G。如果终端使用 2G 网络进行通信,伪基站就可以再次发挥作用^[2]。

由此可见,移动蜂窝网络演进的确使伪基站问题得到一定程度的缓解,但是由于巨大的利益驱动,伪基站技术也在发展演进。本文中我们重点探讨 5G 网络场景下的伪基站问题和应对解决方案。

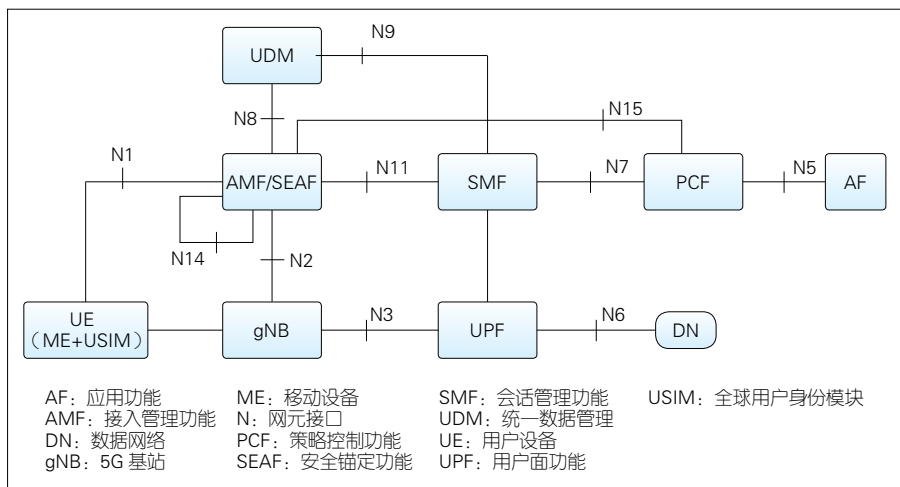
1 5G 网络场景下的伪基站威胁分析

第三代合作伙伴计划(3GPP)定义了 5G 网络架构^[3]。5G 网络包括无线接入网络(RAN)和 5G 核心网(5GC),如图 1 所示。RAN 主要由 5G 基站(gNB)组成。5GC 主要包含

统一数据管理(UDM)、接入管理功能(AMF)/安全锚定功能(SEAF)、会话管理功能(SMF)、用户面功能(UPF)、策略控制功能(PCF)等网络功能。用户设备(UE)一般由移动设备(ME)和全球用户身份模块(USIM)组成。gNB 和 AMF/SEAF 通常位于 UE 的服务网络,UDM 等则位于用户的归属网络。

5G 网络下的广播消息或单播消息有可能导致伪基站攻击出现。当 UE 处于连接模式或进入连接模式时,在安全上下文激活完成之前,可能会发生伪基站攻击。小区周期性广播同步信号和系统信息,UE 则基于同步信号检测小区。如果检测到的小区信号质量高于定义门限,则 UE 根据小区选择标准对所有候选小区进行选择,并驻留到信号质量最好的小区。

基金项目: 工业和信息化部新一代宽带无线移动通信网重大专项(2017ZX03001018)



▲图1 5G 通信网络架构

由于 UE 只检测信号质量，因此伪基站可以很容易地利用高质量信号来诱使 UE 驻留^[4]。

gNB 将业务信息，例如公共警告系统（PWS）消息、地震和海啸预警系统（ETWS）消息、车用无线通信技术（V2X）消息等，通过系统信息块（SIB）发送给驻留 UE。伪基站可篡改合法广播消息中携带的这些服务信息或者伪造虚假的服务信息。如果 UE 驻留到伪基站上，就会存在如下安全威胁：

（1）PWS 携带的信息可能包含虚假信息、垃圾短信等，进而误导用户点击欺诈网站的链接。此外，攻击者还可通过伪基站来传播虚假信息以制造社会恐慌。

（2）UE 驻留到伪基站后，无法获得正常的广播服务，如检测紧急呼叫、ETWS 消息、V2X 消息等。

（3）虚假广播消息也可能对终端的软件系统实施攻击，例如通过构造复杂消息来触发软件系统的已知漏洞，使操作系统或者应用软件被劫持或阻塞。

（4）降级攻击。gNB 向 UE 发送消息，请求 UE 的接入安全能力。UE 在安全模式（SMC）完成前给 gNB 反

馈 UE 能力信息。此时，伪基站可以作为“中间人”捕获该消息，将 UE 的真实接入安全能力调低并反馈给真实的 gNB，以降低 UE 接入到网络的安全能力。

（5）5G 虽然使用用户签约隐藏标识（SUCI）来替代用户签约永久标识（SUPI）发送给网络，但这种替代并非强制。当 SUPI 的保护模式设置为“null-scheme”时，SUPI 仍然使用明文在网络中传递^[5]。如果此时网络并没有建立安全传输模式，例如接入层安全（AS）、非接入层安全（NAS），则 SUPI 有可能在传输时暴露，从而造成隐私泄露。

（6）拒绝服务（DoS）攻击。伪基站可通过不断向 UE 反馈拒绝消息来拒绝 UE 的附着请求，从而制造 DoS 攻击。

在上述威胁中，预防（1）和（3）需要对 SIB 消息实施鉴权，预防（2）需要 UE 驻留小区时进行伪基站检测，预防（4）需要对 UE 和网络之间的能力协商消息进行保护，预防（5）需要防止 UE 驻留到伪基站上，预防（6）需要验证基站向 UE 返回的消息的真实性。

综上分析，5G 系统如果要防止伪

基站带来的安全威胁，就需要对单播或者广播消息进行保护，从而防止 UE 驻留到伪基站上。

2 基于对称算法的防伪基站技术方案

2.1 方案概述

目前移动通信网络主要使用非对称密钥算法为无线信令提供数字签名方案以实现伪基站防护^[6-7]。但非对称密钥算法对计算资源开销过大，这对终端和无线设备等计算能力有限的设备而言势必增加了处理负担。本文中，我们研究了一种基于对称密钥算法的伪基站防护技术，其核心思想为：在归属网络的协助下，服务网络向 ME 动态提供加密的用于验证无线信令真实性的密钥。gNB 为无线信令使用对称密钥算法生成一个消息认证码（MAC-I），然后由 ME 在 USIM 的协助下使用验证无线信令真实性的密钥来验证 MAC-I，以判断无线信令是否是伪基站发出的信令。技术方案包含如下关键步骤：

（1）服务网络向 ME 提供用于验证无线信令真实性的密钥；

（2）用于验证无线信令真实性的密钥是加密的，无法被 ME 解密，只能由 USIM 完成解密；

（3）USIM 存储的加密密钥具备和长期密钥（LTK）相同的安全要求；

（4）加密密钥和用于验证无线信令真实性的密钥需要周期性更新；

（5）ME 将用于验证无线信令真实性的加密密钥和无线信令发送给 USIM 验证。

通过对消息设置完整性校验来实现对广播消息（例如 MIB、SIB 和告警消息等）和下行绝对无线频道编号（DL_ARFCN）的保护。当 UE 驻留到 gNB 的一个小区时，服务网络仅给 USIM 提

供有限数量的密钥。这使得 UE 只能处理密钥对应 gNB 发送的经过完整性保护的广播消息,可有效减少 UE 处理发送消息的 gNB 数量,减轻 UE 处理负担。这些 gNB 覆盖的区域被称为保护区(PA),具体可参考本文 2.4.1 节内容。当攻击者捕获 PA 中某个小区的广播消息时,攻击者只能模拟该小区使用相同的 DL_ARFCN 频率向同一 PA 下的某个 UE 广播无法更改的消息,从而使攻击变得困难,同时使攻击者(伪基站)容易被定位和被网络检测到。当 UE 在一个重放系统消息的伪小区驻留时,UE 将使用本文 2.5.2 节中描述的单播消息真实性验证流程检测到该小区是伪小区,然后重新选择其他小区进行驻留。在同一 PA 中广播捕获的告警消息也可能是一种安全威胁。对此,可通过在告警消息中添加时间和区域等信息,以避免重放威胁。

2.2 技术框架

图 2 描述的是基于对称密钥算法的 5G 网络场景下防伪基站技术方案,包括 4 个流程:(1)保护密钥协商(PKA)流程、(2)保护密钥传输(PKT)流程、(3)PA 信息提供(PAIP)流程和(4)小区真实性验证(CA)流程。其中,流程(1)和(2)一并执行,流程(2)始终在流程(1)之后,流程(3)为服务网络向 ME 提供 PA 信息的流程,流程(4)包括针对广播消息和单播消息的真实性验证过程。

CK_p 是用于加密服务网络向 ME 提供信息的密钥,由归属网络和 USIM 协商产生,并且在 USIM 上存储时具备与根密钥 LTK 相同的安全需求,即 USIM 应防止 CK_p 泄露。

服务网络向 ME 提供的 PA 信息包括被 gNB 用于无线信令完整性保护的密钥。PA 信息中的密钥在被 CK_p 加密后提供给 ME。ME 接收之后必须由

USIM 完成解密。ME 将 PA 信息和无线信令发送给 USIM,然后由 USIM 进行小区真实性验证。CK_p 和被 gNB 使用的密钥必须定期更新,以便密钥在被攻破时即时失效。此外,该技术方案还具有另外的作用,比如在 5G 安全上下文激活之前,可对无线信令上传的敏感信息进行加密传输。

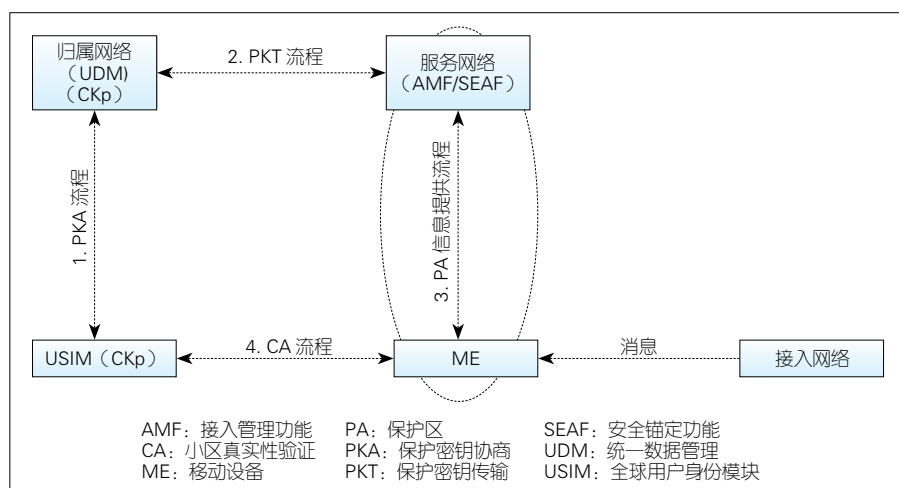
2.3 PKA 和 PKT 流程

图 3 展示了在 UE 向网络发起注

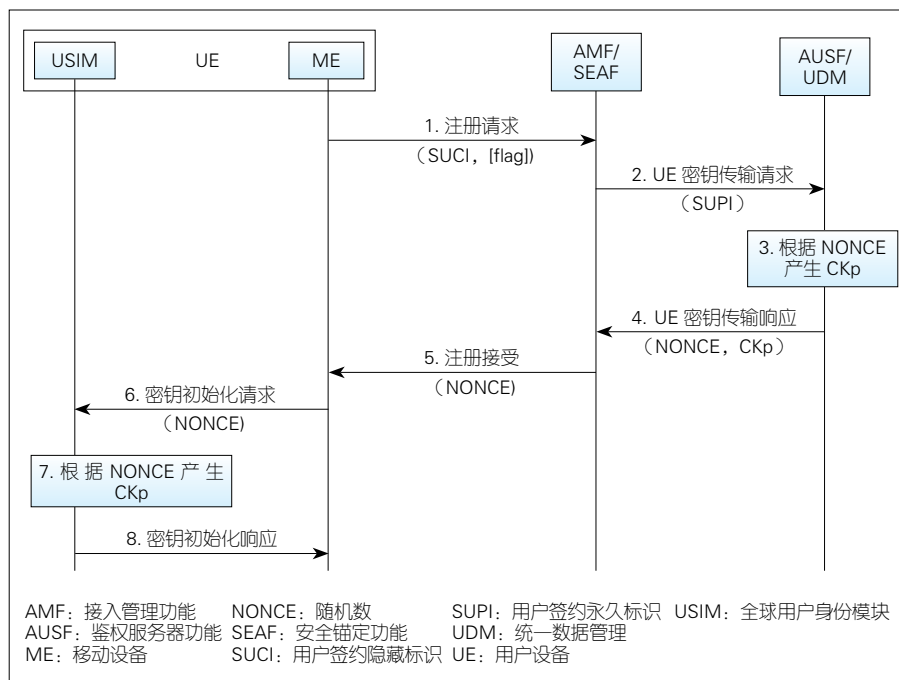
册的过程中完成的 PKA 和 PKT 流程。假设在漫游场景下,AMF/SEAF 所在的漫游服务网络 and AUSF/UDM 所在的归属网络之间的漫游接口具备完整性、机密性和抗重放保护能力。

(1) UE 发送网络注册请求消息给服务网络 AMF/SEAF。如果 UE 决定更新 CK_p,则请求消息需要携带一个说明密钥协商的标记。

(2) 如果 UE 注册到网络而发起注册流程,或者 AMF/SEAF 收到 Flag



▲图 2 基于对称密钥算法的伪基站防护方案



▲图 3 保护密钥协商流程和保护密钥传输流程

标记, 则发起 PKT 流程。AMF/SEAF 向 UDM 发送 CKp 请求消息 (UE 密钥传输请求), 并携带 SUPI。

(3) UDM 收到请求消息后, 会根据 LTK 产生一个 CKp 和一个随机数 (NONCE)。

(4) UDM 向 AMF/SEAF 返回 NONCE 和 CKp, 以完成 PKT 流程。AMF/SEAF 上报 CKp 给防伪基站功能 (AFBF)。AFBF 可以是运营和管理 (OAM) 平台功能, 也可以是专用网络功能。AFBF 通过 CKp 加密 K_{RBS} (参考 2.4.1 定义) 得到加密的 EK_{RBS} 。

(5) AMF/SEAF 向 UE 发送注册成功消息, 并携带 NONCE。

(6) UE 接收 NONCE。如果 UE 支持 AFBF, 则 ME 将 NONCE 发送给 USIM。

(7) USIM 启动 CKp 初始化过程。USIM 使用 NONCE 和 LTK 计算产生密钥 CKp。这种计算方法与 UDM 计算 CKp 的方法相同。同时 USIM 应能够防止 CKp 外泄。

(8) USIM 产生 CKp 之后向 ME 返回指示。

2.4 PAIP 流程

2.4.1 PA 保护区域

我们将每个 PA 下的 gNB 进行分组, 并将组称为共享根密钥组 (SRKG)。每个 SRKG 由 bsGKI 标识, 同时 SRKG 下的每个 gNB 由 bsGNI 标识。如图 4 所示, 而每个 gNB 仅属于一个 SRKG, 每个 SRKG 可以属于一个或多个 PA。

在对 gNB 分组的情况下, 每个 SRKG 具备一个密钥 K_{RBS} 。该密钥被 SRKG 下的每个 gNB 共享。基于 K_{RBS} 派生出 SRKG 下每个 gNB 的密钥 K_{BS} 可用于基站保护。服务网络向 ME 提供的 PA 信息中仅包含 K_{RBS} 的加密密

钥 EK_{RBS} , 以减少提供给 ME 的加密密钥数量。

任何注册区域 (RA) 均被一个且唯一一个 PA 完全覆盖, 因此在 UE 注册过程中, 服务网络可以向 ME 提供 PA 的加密密钥。图 5 描述了 RA 与 PA 之间的关系。

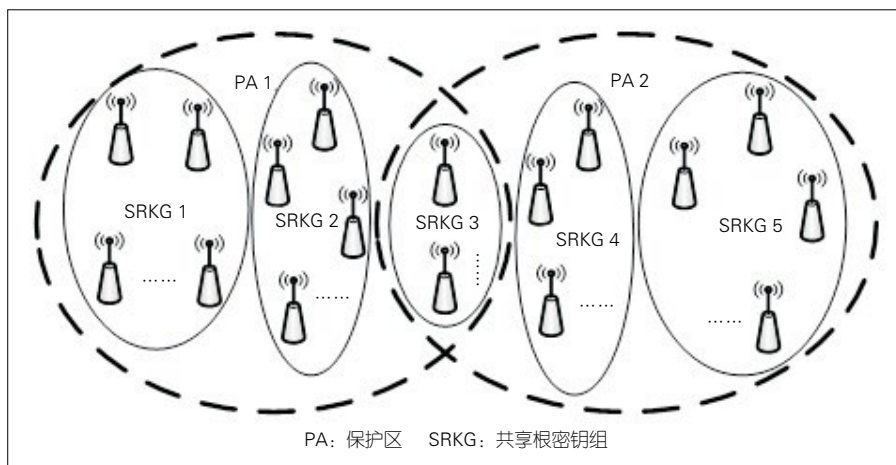
在注册期间 (包括初始注册、移动性或周期性注册) 或者其他初始 NAS 消息处理过程中, AMF 应向 ME 提供 PA 信息, 其中 PA 信息包括 bsGKI 列表、对应的 EK_{RBS} 列表以及密钥有效期。有效期用于防止伪基站破解 gNB 的 K_{BS} 去欺骗 UE。当给密钥加上有效期时, 即使上述密钥被破解, 密钥也已到期。因此, 设置的有效期应足够短, 以使密钥在一定期限内不

被破解。

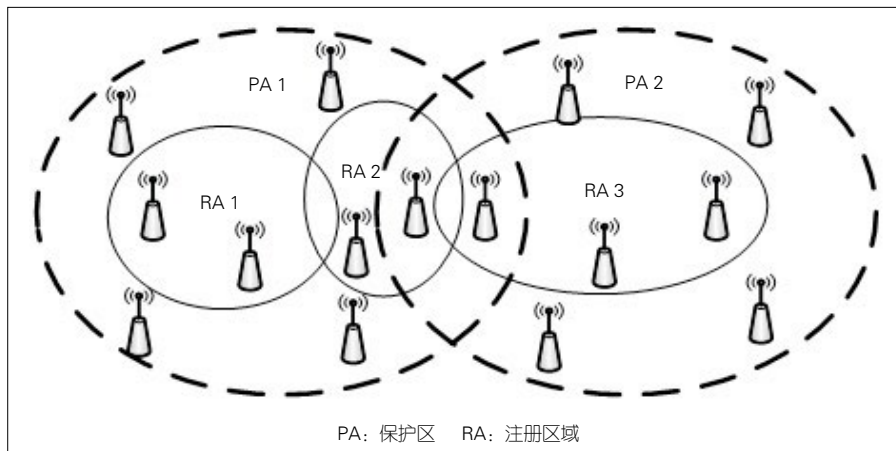
在密钥到期前, 服务网络应更改 SRKG 的根密钥 K_{RBS} 。为了提高性能, 一个 PA 应包含和其全覆盖下的跟踪区毗连的 gNB。这样, 当 UE 位于跟踪区边缘时, UE 也能获取每个毗连 gNB 的密钥。同时 PA 划分的范围也不宜过大, 这是因为太大的 PA 会很难减轻基于重放广播无线信令的伪基站攻击。

2.4.2 PA 信息表示

结合 2.4.1 对 PA 的定义, 如图 6 所示, PA 信息是一组 gNB 信息的结合, 表示形式为: $PA = (<IDa, Enc(CKp, K_{RBS}[a])>, <IDb, Enc(CKp, K_{RBS}[b])>, \dots)$ 。



▲图 4 SRKG 和 PA 之间的关系



▲图 5 RA 和 PA 之间的关系

每个 gNB 信息都由 gNB 标识 (ID_i , $i=a,b,c, \dots$) 和使用 CKp 加密的根密钥 ($K_{RBS}[i]$, $i=a,b,c, \dots$) 组成。在移动过程中, UE 获取的 PA 信息在不断更新。当 UE 从一个 PA 移动到另外一个 PA 时, ME 向服务网络发出 PA 更新请求。服务网络在响应消息中将更新的 PA 信息提供给 ME。

2.4.3 PAIP 流程

服务网络向 ME 提供 PA 信息的预置条件有两个:

(1) 服务网络通过 OAM 平台或专用网络功能提供每个 gNB 对应的

bsGKI 和 bsGNI;

(2) 服务网络通过 OAM 平台或专用网络功能管理 SRKG 的密钥及其其他相关信息。

K_{BS} 的产生方式为: $K_{BS} = \text{HMAC-SHA-256}(K_{RBS}, \langle \text{bsGKI}, \text{bsGNI} \rangle)$ 。其中, $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 是两个参数的串接, 根据 2.4.1 节中的定义, 为 gNB 标识。 K_{RBS} 为 gNB 的根密钥, 即 gNB 所在 SRKG 的密钥。每个 gNB 根据标识和 K_{RBS} 使用密钥派生算法派生得到 K_{BS} 。 K_{BS} 用于对无线信令进行加密和完整性处理。

图 7 描述的是 UE 与服务网络之

间的共享密钥提供流程, 具体包括 5 个步骤:

(1) 在 PKT 流程中, CKp 从归属网络发送到服务网络。

(2) UE 向 AMF/SEAF 发送注册请求, 该消息能够触发 PKA 和 PKT 过程。

(3) 如果 AMF/SEAF 支持 AFBF, 那么 AMF/SEAF 就可以获取 PA 信息, 即覆盖 UE 当前驻留跟踪区的 PA {bsGKI, EK_{RBS} , expiry} 列表。其中, EK_{RBS} 使用 CKp 加密的 K_{RBS} ; expiry 指示 EK_{RBS} 的有效期, 即在有效期到期后, EK_{RBS} 必须更新。

(4) AMF/SEAF 向 UE 返回注册成功消息, 并且消息中携带 PA 信息。

(5) 如果 ME 支持 AFBF, 那么 ME 将保存 PA 信息。

2.5 CA 流程

2.5.1 广播消息真实性验证流程

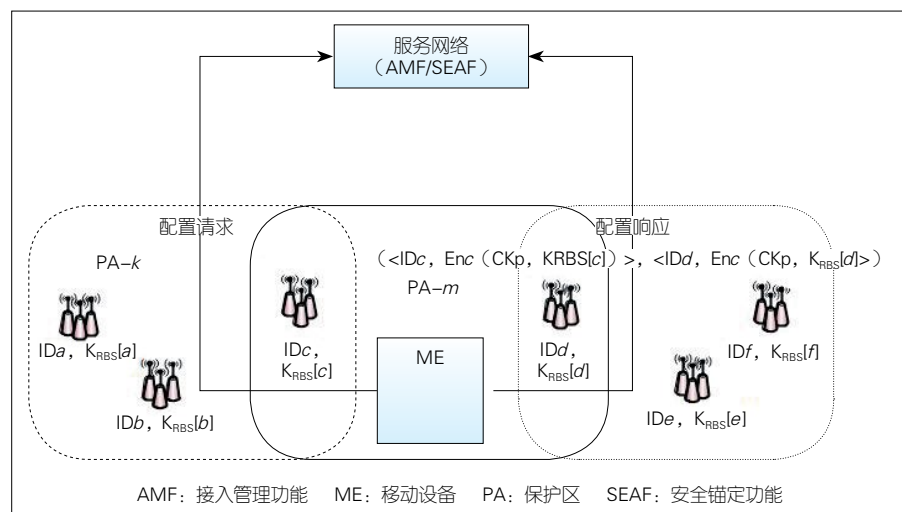
图 8 是广播消息真实性验证流程, 具体包括 6 个步骤:

(1) gNB 决定通过一个小区广播一条消息, 例如 MIB 或 SIB。该消息包含 gNB 标识 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 。如果消息长度加上 DL_ARFCN 长度后大于 32 B, 那么 gNB 产生网络侧哈希值 (HASH_{NW}) 的方式为: $\text{HASH}_{NW} = \text{SHA-256}(\text{DL_ARFCN} \parallel \text{message})$ 。

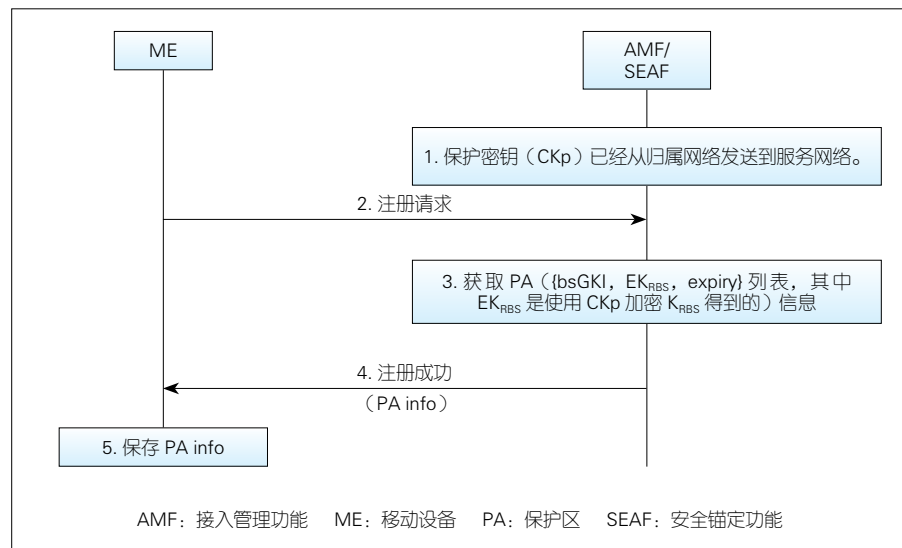
gNB 根据 K_{BS} 计算出 MAC-I, 以及 DL_ARFCN 串接消息 (其长度小于等于 32 B), 或者 HASH_{NW} (其长度大于 32 B)。

(2) gNB 广播消息, 同时消息携带 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 和 MAC-I。

(3) ME 通过小区接收广播消息, 同时消息携带 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 和 MAC-I。ME 检查 bsGKI 是否在保存的 PA 信息列表中。如果检查失败, ME 在缓存中将该小区标记为可疑小区;



▲图 6 ME 移动过程中的 PA 信息动态更新



▲图 7 PA 信息提供流程

如果检查成功,则继续步骤 4。

(4) 如果 DL_ARFCN 长度加上接收消息的长度大于 32 B, 则 ME 将首先计算 HASH_{MS} (与 gNB 计算 HASH_{NW} 的方式相同), 然后发送 EKR_{BS} 、 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 到 USIM。如果已经计算 HASH_{ME} , 则 ME 也会将 HASH_{MS} 发送给 USIM; 否则, ME 将把 DL_ARFCN 和接收到的消息串接发送给 USIM。EKR_{BS} 和 bsGKI 对应表示由 bsGKI 标识的 SRKG 密钥。

(5) USIM 先根据存储的 CK_P 解密 EKR_{BS} 并得到 KR_{BS}, 然后从 KR_{BS} 和 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 中派生出 K_{BS}, 再根据 K_{BS} 计算出 XMAC-I。USIM 计算 XMAC-I 的方式与 gNB 基于 K_{BS} 的计算 MAC-I 方式相同。

(6) USIM 将 XMAC-I 反馈给 ME。ME 会比较 XMAC-I 和 MAC-I, 如果两者相等, 则 ME 会处理广播消息; 否则, ME 会在缓存中将该小区标记为高危小区。

对网络认证成功后, UE 应检查缓存中的可疑或高危小区是否在 PA 中。如果不在, 则缓存中的小区被标记为可疑小区, 否则 UE 将进行小区验证。如果验证失败, UE 在缓存中将该小区标记为伪小区; 如果验证成功, 该小区将从缓存中被移除。UE 可以将缓存中的可疑、高危和伪小区告知给服务网络。

2.5.2 单播消息真实性验证流程

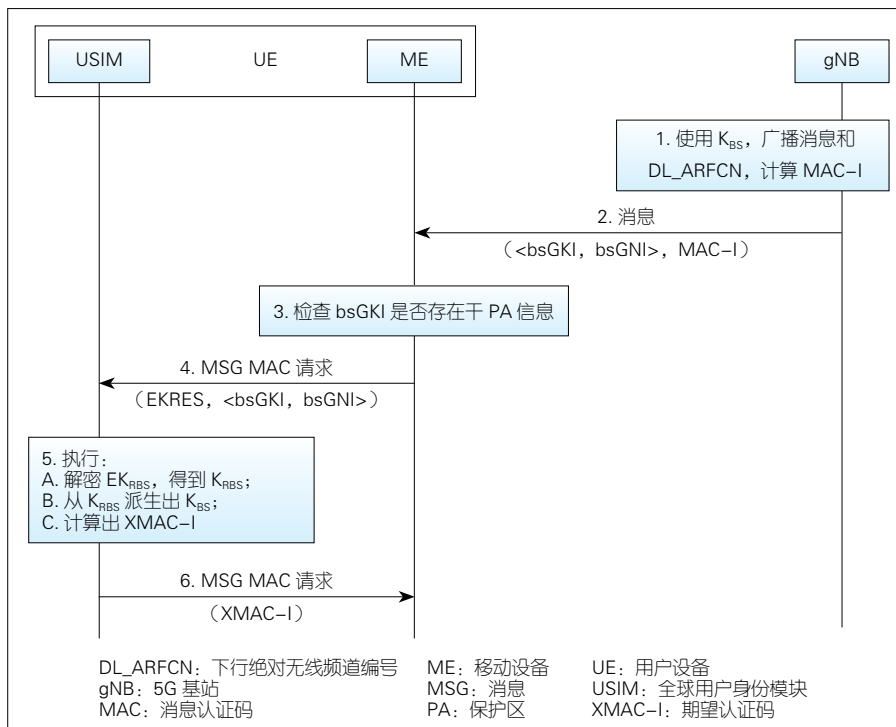
只有在 UE 和 gNB 无法获取 AS 安全上下文的情况下, 单播消息真实性验证流程才会生效。图 9 是单播消息真实性验证流程, 具体包括 7 个步骤。

(1) ME 尝试驻留在一个小区上, 并决定向该小区发送消息 1。ME 产生一个随机字符串 NONCE。NONCE 可被作为连续下行消息的重放保护参数。

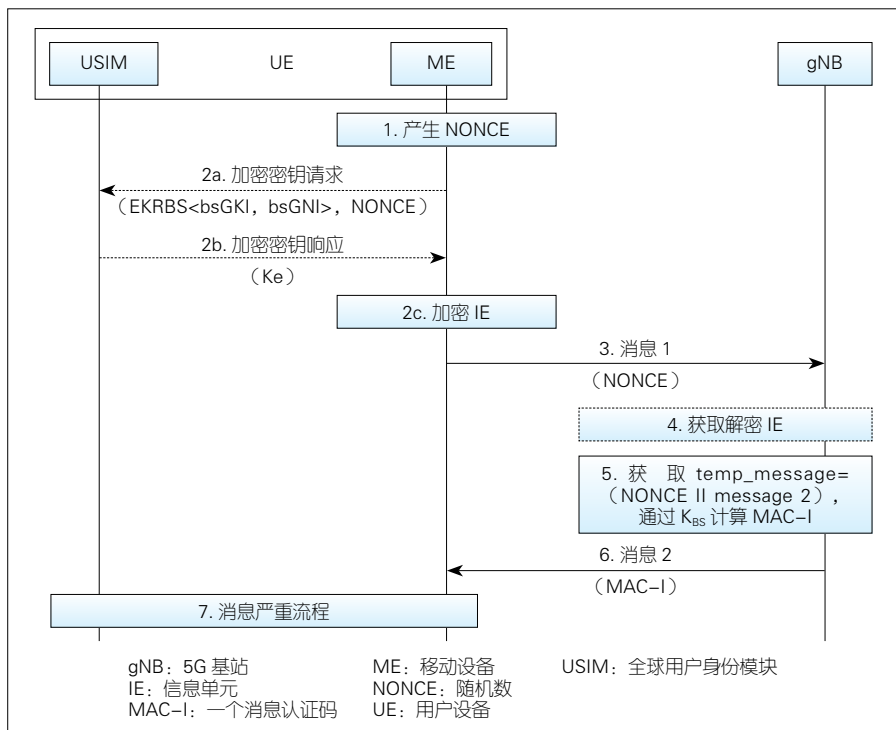
(2) ME 可以决定对消息 1 中

的敏感信息进行保护。ME 将 EKR_{BS} 、 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 、敏感信息和 NONCE 发送给 USIM。其中, $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 是从小区发送的广播消息 (比如 SIB

或者 MIB) 中获取的。同时, EKR_{BS} 和 bsGKI 相对应。USIM 解密 EKR_{BS} 得到 KR_{BS}, 并根据 KR_{BS} 和 $\langle \text{bsGKI}, \text{bsGNI} \rangle$ 派生得到 K_{BS}。然后 USIM 根据



▲ 图 8 广播消息真实性验证流程



▲ 图 9 单播消息真实性验证流程

K_{BS} 派生出 K_e , 并将 K_e 返回给 ME。ME 根据 K_e 和 NONCE 生成密钥流, 并对敏感信息和密钥流进行异或计算, 以实现敏感信息的加密。

(3) ME 向小区发送消息 1。其中, 消息 1 携带 NONCE, 并且消息 1 是部分加密的。

(4) gNB 收到消息 1。如果消息 1 是部分加密的, 则 gNB 解密加密部分以获得明文信息。

(5) gNB 决定向 UE 发送消息 2。gNB 获取 $\text{temp-message} = (\text{NONCE} \parallel \text{message2})$, 并按照 2.5.1 描述的方式计算 temp-message 的 MAC-I。

(6) gNB 将消息 2 发送给 UE。其中, 消息 2 携带 MAC-I。

(7) ME 接收消息 2, 并按照 2.5.1 中的步骤 4—6 进行消息真实性验证。验证过程用 $\text{temp-message} = (\text{NONCE} \parallel \text{message2})$ 替代消息 2。

3 结束语

本文描述了 5G 网络场景中由攻击广播消息或单播消息产生的伪基站问题, 进而提出了针对广播信息和单播信息进行真实性验证的方法以抵御

伪基站攻击。此外, 本文还研究了使用对称密钥算法对广播消息和单播消息进行机密性或者完整性防护的实现过程, 包括用于保护密钥的协商和传递过程、PA 信息表示和提供过程, 以及针对广播信息或单播信息的真实性验证过程。相对于使用非对称算法对无线消息进行防护的方法, 对称密钥算法具有对网络资源要求少的优势, 是值得推荐的技术方案。随着技术的不断演进, 伪基站攻击方法还会不断发展, 相应的防护方案需要进行不断完善。

参考文献

- [1] 李赟, 赵玉萍, 孙春来, 等. 一种基于伪信令的伪基站抑制方法研究与分析 [J]. 信息安全, 2014, (9): 12-16
- [2] 刘长波, 张敏, 常力元, 等. 5G 伪基站威胁分析及安全防护建议 [J]. 移动通信, 2019, 43(10): 58-61
- [3] 3GPP. System architecture for the 5G system: 3GPP TS 23.501 [S]. 2020
- [4] 李宁. 基于 GSM 基站信息的伪基站侦测系统 [D]. 成都: 西华大学, 2016
- [5] 3GPP. Security architecture and procedures for 5G system: 3GPP TS 33.501 [S]. 2020
- [6] IEEE. IEEE standard for identity-based cryptographic techniques using pairings: IEEE 1363.3 [S]. 2013

- [7] IETF. Elliptic curve-based certificateless signatures for identity-based encryption: IETF RFC 6507 [S]. 2012

作者简介



毛玉欣, 中兴通讯股份有限公司资深系统架构师; 主要研究方向为 5G 网络安全和虚拟化安全; 参与多项国际技术标准的制定; 拥有发明专利 70 余项、国际标准提案 50 余项。



游世林, 中兴通讯股份有限公司资深技术预研工程师; 从事 IMS、LTE、5G 安全方面的标准预研和产品研发工作; 参与多项国际技术标准的制定; 拥有发明专利和标准提案数十项。



闫新成, 中兴通讯股份有限公司网络安全首席系统架构专家、移动网络与移动多媒体技术国家重点实验室未来网络研究中心副主任, 教授级高工; 从事电信行业 20 年, 曾主持国家科技重大专项 5G 安全课题; 获得多项科技奖励, 拥有专利 40 余项。

《中兴通讯技术》第 27 卷总目次

卷·期·页

卷首特稿

筌路蓝缕 玉汝于成 王喜瑜 27-1-01

专题

专题：视频技术和用户体验评测

专题导读	江涛, 陆平	27-1-04
点云编码综述	李厚强, 李礼, 李竹	27-1-05
AVS3 视频编码关键技术及应用		
.....	张嘉琪, 雷萌, 马思伟	27-1-10
下一代通信助力实时分布云渲染		
.....	陆平, 盛斌, 朱方	27-1-17
视频质量增强模型加速算法	杨文哲, 徐迈, 白琳	27-1-21
基于图神经网络的视频推荐系统		
.....	高宸, 李勇, 金德鹏	27-1-27
脑启发视频用户体验评测关键技术		
.....	陶晓明, 杜冰, 段一平	27-1-33
超高清内容清晰度用户体验质量评价		
.....	朱文瀚, 翟广涛, 陶梅霞, 杨小康, 张文军	27-1-37
交互式视频质量评价方法研究进展		
.....	李继龙, 赵雪, 杨铀	27-1-44
HTTP 自适应流媒体直播系统中的用户体验质量优化		
.....	宋靳裸, 张远, 王博	27-1-48

专题：6G 愿景及技术挑战

专题导读	张平, 张建华	27-2-01
6G 网络的一些思考	杨峰义, 刘洋, 杨蓓	27-2-02
面向 6G 的智能物联网关键技术		
.....	Wai CHEN, 鲍媛媛	27-2-06

卷·期·页

面向 6G 的车联网关键技术	郎平, 田大新	27-2-13
太赫兹技术发展及其应用于 6G 的挑战分析		
.....	刘利平, 营梦楠, 陈艺戡	27-2-17
6G 背景下超奈奎斯特技术的机遇		
.....	王亚峰, 金婧, 王启星	27-2-25
一种基于公平性的无人机基站通信智能资源调度方法		
.....	吴官翰, 赵建伟, 高飞飞	27-2-31
非线性波数域的高分辨扫频干涉层析成像		
.....	谢胜利, 蔡树银, 白玉磊	27-2-37

专题：边缘计算与算力网络

专题导读	赵慧玲	27-3-01
边缘计算与算力网络综述	雷波, 赵倩颖, 赵慧玲	27-3-03
算力感知网络架构与关键技术		
.....	姚惠娟, 陆璐, 段晓东	27-3-07
算力网络实现一体化服务的探索与实践		
.....	雷波, 赵倩颖, 凌泽军	27-3-12
基于可编程网络的算力调度机制研究		
.....	李铭轩, 曹畅, 杨建军	27-3-18
基于 SRv6 的算力网络资源和服务编排调度		
.....	黄光平, 史伟强, 谭斌	27-3-23
算力网络：以网络为中心的融合资源供给		
.....	李少鹤, 李泰新, 周旭	27-3-29
多层次算力网络集中式不可分割任务调度算法		
.....	巩宸宇, 舒洪峰, 张昕	27-3-35

专题：高铁智能通信技术与应用

专题导读	艾渤	27-4-01
智能化高铁车载缓存技术		
.....	崔新雨, 刘玲, 周一青, 潘振岗	27-4-02

用于超高移动性信道的正交时频空调制	刘梦晓, 周晶, 张文逸 27-4-09
智能反射面在高铁通信下的应用研究	王靖瑜, 鞠宏浩, 方旭明 27-4-14
基于 5G-R 业务的高速铁路异构网络接入技术	李翠然, 谢健骊, 高文娟 27-4-18
高速移动条件下毫米波信道测量	左世元, 范戎飞 27-4-24
高速铁路高架桥场景中的复合无线信道特性	张逸康, 王公仆, 叶如意 27-4-30
智能超表面技术在智能高铁通信场景的应用探讨	赵亚军, 章嘉懿, 艾渤 27-4-36

专题：低轨卫星通信技术与应用

专题导读	郭庆 27-5-01
空天地一体化网络技术展望	田开波, 杨振, 张楠 27-5-02
面向低轨卫星的频谱认知智能管控	李高, 王威, 吴启晖 27-5-07
面向低轨卫星星座的多星协作信号检测技术	王虎威, 叶能, 安建平 27-5-12
NGSO 互联网星座用户链路同频干扰分析	刘子威, 李嘉颖, 张更新 27-5-18
DTN 辅助的低轨卫星网络路由技术	张培颖, 王超, 吴胜 27-5-23
大规模低轨卫星网络移动性管理方案	吴琦, 郭孟泽, 朱立东 27-5-28
基于多维梯度的卫星集群高可靠协同计算方法	马步云, 任智源, 李赞 27-5-36
卫星地面融合网络的窃听威胁与物理层安全解决方案	韩帅, 李季蹊, 李静涛 27-5-43

专题：触觉通信技术

专题导读	周亮 27-6-01
触感技术的现状与发展	曾繁荣, 房颖, 赵铁松 27-6-02

触觉通信中的信号获取与恢复关键技术	刘恒发, 魏昕 27-6-08
面向 6G 的触觉互联网关键技术与挑战	尤子硕, 李强, 唐冰, 葛晓虎 27-6-15
面向触觉通信的预测窗口和计算卸载策略优化	吴岩, 吴丹, 王嵩 27-6-21
一种面向视触融合人机交互的柔性触觉传感阵列	王爽, 吴兵, 刘倩 27-6-29

专家论坛

小视频内容分析技术发展探讨	薛向阳, 李斌 27-1-54
6G: 继续体系结构的变革	李少谦 27-2-43
对 6G 关键技术发展的思考	葛建华, 李靖 27-2-45
夯实云网融合, 迈向算网一体	唐雄燕, 张帅, 曹畅 27-3-42
零触碰与零信任	李军, 胡效赫 27-3-47
铁路新一代移动通信的挑战与思考	钟章队, 官科, 陈为, 艾渤 27-4-44
大规模低轨星座卫星通信网发展展望	孙智立, 李天儒 27-5-48
触觉通信技术发展的若干思考	陈鸣锴, 周亮 27-6-35

企业视界

构建智能实时网络, 使能 5G 视频业务繁荣	吕达, 郑清芳 27-1-60
5G NR 定位技术及其部署方案	张诗壮, 李俊强, 陈诗军 27-2-49
数据中心网络架构和协议演进分析	魏月华, 陈晓, 张征 27-3-51
5G 半导体产业发展趋势和创新——持续就是竞争力	刘新阳 27-4-51
未来网络可信通信技术	闫新成, 周娜, 蒋志红 27-5-52
800 Gbit/s 光模块技术及应用	王会涛, 张平化, 苏展 27-6-40

技术广角

面向视频云微服务系统的智能运维技术	徐代刚, 姜磊, 梅君君 27-1-68
用于人工智能的硅基光电子芯片	白冰, 裴丽, 左晓燕 27-1-77
面向移动用户的 RIS 波束追踪与覆盖增强算法	郭雅婧, 章嘉懿, 鲁照华, 王明慧 27-2-54
基于机器学习的无线网络流量预测与增长潜力评估	邢旭东, 高晖, 顾军 27-2-60
共建共享下边缘云建设思考	黄倩, 黄蓉 27-3-56
边缘计算使能星地协同网络下的服务部署机制	卢华, 段雪飞, 李斌 27-3-62
5G 毫米波通信—物理层安全预编码	倪云云, 陈伯庆, 李刚 27-4-53

Nature Flow: 新转发架构赋能未来数据中心网络	商志彪, 雷波, 郭茜 27-4-60
未来卫星通信和 5G-NR 深度融合架构设计	杨立, 窦建武 27-5-60
TSN 与 5G 融合部署的需求和网络架构演进	朱瑾瑜, 张恒升, 陈洁 27-6-47
基于典型无线场景库的 LTE-V2X 信道特性	苏昭阳, 刘留, 冯毅 27-6-53
5G 电信云网络的容器演进方案	史庭祥, 徐法禄, 章璐 27-6-58
基于对称密钥算法的 5G 网络防伪基站技术	毛玉欣, 游世林, 闫新成 27-6-65

专题预告

《中兴通讯技术》2022 年专题计划

期次	专题名称	策划人
1	新型网络技术	中国联通研究院副院长 唐雄燕
2	自然语言预处理模型	中国工程院院士 郑纬民
3	智能超表面技术	中兴通讯技术预研总工 赵亚军 北京理工大学教授 费泽松
4	多频段协同通信	电子科技大学教授 李少谦
5	通信感知一体化	中国科学技术大学教授 卫国
6	网络内生安全	北京航空航天大学教授 刘建伟

《中兴通讯技术》杂志（双月刊）投稿须知

一、杂志定位

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

二、稿件基本要求

1. 投稿约定

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

2. 内容和格式要求

- (1) 稿件须具有创新性、学术性、规范性和可读性。
- (2) 稿件需采用 WORD 文档格式。
- (3) 稿件篇幅一般不超过 6 000 字（包括文、图），内容包括：中、英文题名，作者姓名及汉语拼音，作者中、英文单位，中文摘要、关键词（3 ~ 8 个），英文摘要、关键词，正文，参考文献，作者简介。
- (4) 中文题名一般不超过 20 个汉字，中、英文题名含义应一致。
- (5) 摘要尽量写成报道性摘要，包括研究的目的、方法、结果 / 结论，以 150 ~ 200 字为宜。摘要应具有独立性和自明性。中英文摘要应一致。
- (6) 文稿中的量和单位应符合国家标准。外文字母的正斜体、大小写等须写清楚，上下角的字母、数据和符号的位置皆应明显区别。
- (7) 图、表力求少而精（以 8 幅为上限），应随文出现，切忌与文字重复。图、表应保持自明性，图中缩略词和英文均要在图中加中文解释。表应采用三线表，表中缩略词和英文均要在表内加中文解释。
- (8) 所有文献必须在正文中引用，文献序号按其在文中出现的先后次序编排。常用参考文献的书写格式为：
 - 期刊 [序号] 作者. 题名 [J]. 刊名, 出版年, 卷号 (期号): 引文页码. 数字对象唯一标识符
 - 书籍 [序号] 作者. 书名 [M]. 出版地: 出版者, 出版年: 引文页码. 数字对象唯一标识符
 - 论文集中析出文献 [序号] 作者. 题名 [C] // 论文集编者. 论文集名 (会议名). 出版地: 出版者, 出版年 (开会年): 引文页码. 数字对象唯一标识符
 - 学位论文 [序号] 作者. 题名 [D]. 学位授予单位所在城市名: 学位授予单位, 授予年份. 数字对象唯一标识符
 - 专利 [序号] 专利所有者. 专利题名: 专利号 [P]. 出版日期. 数字对象唯一标识符
 - 国际、国家标准 [序号] 标准名称: 标准编号 [S]. 出版地: 出版者, 出版年. 数字对象唯一标识符
- (9) 作者超过 3 人时，可以感谢形式在文中提及。作者简介包括：姓名、工作单位、职务或职称、学历、毕业于何校、现从事的工作、专业特长、科研成果、已发表的论文数量等。
- (10) 提供正面、免冠、彩色标准照片一张，最好采用 JPG 格式（文件大小超过 100 kB）。
- (11) 应标注出研究课题的资助基金或资助项目名称及编号。
- (12) 提供联系方式，如：通讯地址、电话（含手机）、Email 等。

3. 其他事项

- (1) 请勿一稿多投。凡在 2 个月（自来稿之日算起）以内未接到录用通知者，可致电编辑部询问。
- (2) 为了促进信息传播，加强学术交流，在论文发表后，本刊享有文章的转摘权（包括英文版、电子版、网络版）。作者获得的稿费包括转摘酬金。如作者不同意转摘，请在投稿时说明。
- (3) 编辑部地址：安徽省合肥市金寨路 329 号凯旋大厦 1201 室，邮政编码：230061。
- (4) 联系电话：0551-65533356，联系邮箱：magazine@zte.com.cn。
- (5) 本刊只接受在线投稿，欢迎访问本刊投稿平台：tech.zte.com.cn/submission。

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、刘新阳、
陆 平、史伟强、王会涛、熊先奎、
赵志勇、朱 方、朱晓光

双月刊 1995 年创刊 总第 161 期
2021 年 12 月 第 27 卷 第 6 期（卷终）

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

总编辑：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、卢丹、朱莉、任溪溪
设计排版：徐莹
发行：王萍萍
编务：王坤

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

发行方式：自办发行
印刷：合肥添彩包装有限公司
出版日期：2021 年 12 月 15 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