

分类号: TN929.53

单位代码: 10335

密 级: 公开

学 号: 12131008

# 浙江大学

## 博士学位论文



中文论文题目: 面向边缘智能网络的  
通信、感知和计算一体化研究

英文论文题目: Research on Integrated  
Communication, Sensing and Computing  
for Edge-intelligent Networks

申请人姓名: 齐俏

指导教师: 陈晓明 研究员

专业名称: 信息与通信工程

研究方向: 无线通信

所在学院: 信息与电子工程学院

论文提交日期 2023年6月9日



面向边缘智能网络的  
通信、感知和计算一体化研究



论文作者签名: 齐倩

指导教师签名: 陈明

论文评阅人1: 匿名

评阅人2: 匿名

评阅人3: 匿名

评阅人4: 匿名

评阅人5: 匿名

答辩委员会主席: 张朝阳 教授 浙江大学信电学院

委员1: 孟利民 教授 浙江工业大学

委员2: 厉力华 教授 杭州电子科技大学

委员3: 黄崇文 研究员 浙江大学信电学院

委员4: 刘雷 研究员 浙江大学信电学院

答辩日期: 2023年6月9日



**Research on Integrated Communication, Sensing  
and Computing for Edge-intelligent Networks**



Author's signature: Diao Di

Supervisor's signature: Xiaoming Chen

External Reviewers: Anonymous  
Anonymous  
Anonymous  
Anonymous  
Anonymous

Examining Committee Chairperson:  
Zhaoyang Zhang

Examining Committee Members:  
Limin Meng  
Lihua Li  
Chongwen Huang  
Lei Liu

Date of oral defence: Jun. 9, 2023



## 浙江大学研究生学位论文独创性声明

本人声明所呈交的学位论文是本人在导师指导下进行的研究工作及取得的研究成果。除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得 浙江大学 或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文作者签名：齐倩

签字日期：2023年6月9日

## 学位论文版权使用授权书

本学位论文作者完全了解 浙江大学 有权保留并向国家有关部门或机构送交本论文的复印件和磁盘，允许论文被查阅和借阅。本人授权浙江大学可以将学位论文的全部或部分内内容编入有关数据库进行检索和传播，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密的学位论文在解密后适用本授权书)

学位论文作者签名：齐倩

导师签名：陈晓明

签字日期：2023年6月9日

签字日期：2023年6月12日





## 致 谢

时光荏苒，转眼间五年的博士生涯已接近尾声。回首过去，我深感自己能够顺利地度过这段充满挑战和收获的日子，离不开导师、同窗、朋友以及家人关爱和支持。在此，我衷心地表达我最诚挚的感激之情。

首先，我要向我的导师陈晓明教授致以最崇高的敬意。何其有幸，成为晓明老师的学生，让我快乐且充实地度过了整个博士生涯。犹记得 18 年 3 月时和晓明老师的初次见面，一转眼五年就过去了。刚开始时，晓明老师手把手带我教我科研，为我答疑解惑，指点迷津。后来，晓明老师特意培养，使我在科研上能够逐渐独当一面。晓明老师是我见过最敬业、最勤奋的老师，也是我一生都要学习和追随的榜样。晓明老师以身作则，每天都是信电楼最早到的，每次找他讨论也永远在办公室，甚至晚上或者周末也会特意赶到学校只为回答你的一个问题。不管是科研、学习、生活还是就业方面，晓明老师永远都是能够站在我们学生的角度考虑，鼓励我们、帮助我们，给我们足够的支持和尊重。感恩之情，无以言表，永记于心！

我还要感谢张朝阳教授、钟财军教授、黄崇文教授、外导 Prof. Yuen Chau 和 Prof. Derrick Kwan Ng 在科研上对我的帮助和指导。感谢实验室的贾润东师兄、邵晓丹师姐、同届的俞光华、褚建杭师弟、田飞燕师妹、郭涔峰师弟、高智赫师弟、强意扬师弟、王琦师弟、李米丽师妹、应铭师弟和林晨岚师妹在科研和生活上对我的帮助，你们让我觉得实验室充满了温馨与快乐。这里还要特别感谢润东大师兄领我科研入门、小杭师弟近十年来的友情、琦琦师弟对我的信任和帮助以及光华、飞燕师妹和郭师弟一起快乐的时光。感谢隔壁的王得志师兄五年来对我的照顾和帮助，感谢隔壁的陶琴师姐对我就业的帮助。感谢硕士期间的室友陈凌畅、刘诗雨、孙怡琳以及博士期间的室友杨铎对我生活上的帮助。感谢学院机关的行政老师们对我的帮助和培养。感谢校医院推拿科黄融融医生三年多的治疗。感谢在新加坡留学期间房东 LinSan 和去了喵星的 Pukyi 对我的帮助和陪伴。感谢 CC98, my home! 五年来，我勤勤恳恳，每天都要刷 98! 98 让我开阔了眼界，也让我乐不可支。感谢竺可桢奖学金，让我遇上我的男朋友周璐瑜，感谢他在生活上对我的理解、支持、帮助、包容和陪伴。

此外，我要深深地感谢我的父亲齐荣生、母亲朱建娟、姐姐齐艳以及刚出生不久的外甥女小嘟嘟！感谢你们一直在背后默默地照顾我、鼓励我、支持我和治愈我，让我在求学路上没有后顾之忧。现在想来，人生真的很奇妙。曾经的我有股叛逆在身上，十年前选专业，父母说当医生或者老师好，我毅然选择了读工科。大三时，父母让我考研，我说不考。读硕士时，父母让我读博士，我说不读。现在，我即将博士毕业成为一名高校老师。读博期间，感受最深的就是家人和健康才是最最重要的。希望以后有更多的时间能够陪伴你们。

最后，也想感谢一下努力又幸运的自己。我常说我运气不好，但是在人生转折点时运气每每都意外的好。小学时生了一场大病，不能剧烈运动，各种忌口，妈妈带我在杭州求医数年，当时就已开始怀疑人生。从初中开始，逐步成为了所谓的尖子生。没有参加体育中考，中考也没考好，但还是擦底进入了唯一一所重点高中。高中稳步上升，但高考又又没考好，坚持留杭，意外选择了其实很好的杭电电子专业。大学稳步上升，考研时自信满满，但又又又没考好，却又擦底进了浙大，还遇上了这么好的导师。读研期间，算是风声水起，但没有博士名额辗转求职就业，最后还是顺利转博。博士快毕业时，就业环境恶劣，却又如愿找到了心仪的工作。二十多年的书总算是读完了，但是人生的路还很长。我想，未来也一定会更好的！

再次感谢一路上陪伴着我的家人、师长、同窗和朋友的支持与帮助，希望你们健康快乐，幸福美满！感谢勤奋乐观的信电学院，感谢海纳江河的浙江大学，感谢繁荣昌盛的祖国！

齐俏

23年6月于美丽的求是园

## 摘 要

近年来,随着物联网、车联网和工业互联网的广泛应用,网络边缘部署了大量设备和网元,以实现海量信息的高效感知、传输和处理。在这种情况下,无线网络逐渐向边缘智能网络演进,并具有通信、感知和计算一体化功能,从而为各行各业提供无处不在的实时智能服务。

本学位论文以边缘智能网络为应用场景,针对资源受限、干扰复杂和环境时变等挑战,以“泛在通信、精确感知和实时计算”为研究目标,以资源共享和多点协作为手段,对通信、感知和计算一体化的基础理论、关键技术、典型架构和核心算法展开研究。论文的主要内容及其创新性如下。

首先给出了通信、感知和计算一体化的基础理论和关键技术。针对边缘智能网络的特点,介绍了通信、感知和计算这三个基本要素的功能定义和性能评价指标。根据通信、感知和计算一体化的发展趋势,详细阐述了其在各个演进阶段的特点,并从资源、能力和业务的角提出了一种通用的一体化系统架构,为通信、感知和计算一体化的设计提供了有益的理论指导。基于边缘智能网络面临的挑战,分析了实现通信、感知和计算一体化的关键技术,为通信、感知和计算一体化的设计提供了重要的技术支持。

其次提出了一种面向感知的通信和计算一体化架构,利用非正交传输和边缘联邦学习实现海量感知数据的高效处理,将“先感知再通信最后计算”的模式转变为“基于感知数据的模型参数交互”的模式。针对信道衰落、干扰和噪声以及信道估计误差等不利因素对模型参数传输的影响,提出了一种基于空中计算的联合设备选择和收发机设计的鲁棒算法,提高了通信效率和计算精度。仿真结果表明所提算法在通信和计算方面均比现有算法表现稳定且性能更佳。

再次提出了一种计算辅助的通信和感知一体化架构,利用多功能基站在同一频谱和相同硬件上实现了通信和感知的协同,有效提高了频谱和硬件的利用效率。针对资源共享导致的通信信号和感知信号的混叠问题,借助人工智能强大的计算能力,提出了一种基于深度学习的联合感知发射波形和通信接收波束设计算法,提高了通信和感知一体化的整体性能。仿真结果证明了所提算法的有效性且具有低复杂性和高鲁棒性的特点。

然后提出了一种通信支撑的感知和计算一体化架构，利用多节点协作通信实现感知信息和计算数据的高效传递和汇聚，显著提升了目标感知精度和数据计算效率。针对边缘智能网络中节点资源的有限性以及感知和计算任务的耦合性，提出了一种以最小化加权总能耗为目标的联合卸载选择和资源分配算法，增强了感知和计算一体化的整体性能。仿真结果表明所提算法可以在保证感知和计算任务的服务质量情况下，有效降低系统总能耗。

最后提出了一种功能复用的通信、感知和计算一体化架构，利用配置大规模天线阵列的基站所具有的空间自由度，在同一资源块上同时实现了多目标感知、多维计算和多流通信。针对功能复用情况下多业务的同信道干扰和资源竞争，提出了两种联合发射和接收波束成形设计算法，提升了系统的整体性能。仿真结果证实了所提的两种算法分别可以有效提升系统加权总性能和降低系统发射总功率。

**关键词：** 边缘智能网络，通信、感知和计算一体化，资源管理，波束设计，干扰协调

## Abstract

In recent years, with the widespread application of the Internet of Things, the Internet of Vehicles, and Industrial Internet, a large number of devices and network elements have been deployed at the network edge to achieve efficient sensing, transmission, and processing of massive information. In this context, wireless networks are gradually evolving into edge-intelligent networks, with the function of integrated communication, sensing, and computing (ICSC), to provide ubiquitous real-time intelligent services for various industries.

This thesis takes edge-intelligent networks as application scenarios. Facing various challenges such as limited resources, complex interference and time-varying environment, with the goal of ubiquitous communication, accurate sensing and real-time computing, by means of resource sharing and multi-node cooperation, the thesis conducts research on basic theory, key technologies, typical architectures and core algorithms of ICSC. The main contents and innovations are as follows.

Firstly, basic theory and key technologies of ICSC are described. With the characteristics of edge-intelligent networks, the functional definition and performance evaluation indicators of three basic elements of communication, sensing and computing are introduced. According to the development trends of ICSC, the features at each evolution stage are elaborated. From the perspectives of resources, capabilities and services, a general ICSC system framework is proposed to provide a useful theoretical guidance for the design of ICSC. Based on the challenges faced by edge-intelligent networks, the key technologies for implementing ICSC are analyzed, providing an important technical support for the design of ICSC.

Secondly, a sensing-oriented architecture of integrated communication and computing is proposed, which combines non-orthogonal transmission and edge federated learning to efficiently process massive sensing data, transforming the “sensing-communication-computing” mode into a “model parameter interaction based on sensing data” mode. To address the adverse effects of channel fading, interference, noise, and channel estimation errors on model parameter transmission, a robust over-the-air computation-based joint device selection and transceiver design algorithm is

proposed, which improves communication efficiency and computation accuracy. Simulation results verify that the proposed algorithm performs more stable and better both in communication and computing than existing algorithms.

Afterward, a computing-aided architecture of integrated communication and sensing is proposed, which utilizes a multi-functional base station to realize the integration of communication and sensing on the same spectrum with the same hardware, effectively improving the utilization efficiency of spectrum and hardware. To address the issue of communication signals and sensing signals overlapping caused by resource sharing, with the help of the powerful computing capabilities of artificial intelligence, a deep learning-based joint sensing transmit waveform and communication receive beamforming design algorithm is proposed to improve the overall performance of integrated communication and sensing. Simulation results demonstrate the effectiveness of the proposed algorithm and its low complexity and high robustness characteristics.

Then, a communication-supported architecture of integrated sensing and computing is proposed, which utilizes multi-node cooperation communication to achieve the transmission and aggregation of sensing information and computing data, significantly improving the accuracy of target sensing and the efficiency of data computing. To address the issues of limited resources of nodes and the coupling of sensing and computing tasks in edge intelligent networks, a joint offloading selection and resource allocation algorithm with the goal of minimizing the total weighted energy consumption is proposed to improve the overall performance of integrated sensing and computing. Simulation results validate that the proposed algorithm can effectively reduce the total energy consumption while guaranteeing the service quality of sensing and computing tasks.

Finally, a function-multiplexing architecture of ICSC is proposed, which utilizes the spatial degrees of freedom provided by a base station equipped with a large-scale antenna array to simultaneously implement multi-target sensing, multi-dimensional computing and multi-flow communication on the same wireless resources. To address co-channel interference and resource competition among multiple services caused by function multiplexing, two joint transmit and receive beamforming design algorithms are proposed to improve the overall performance. Simulation results confirm that the two proposed algorithms can effectively improve the weighted overall system performance and reduce the total system transmit power, respectively.

**Keywords:** Edge-intelligent networks, integrated communication, sensing and computing, resource management, beamforming design, interference coordination.

## 插图

|      |  |    |
|------|--|----|
| 1.1  | 移动通信发展史.....                                 | 1  |
| 1.2  | 全球物联网设备连接数量和数据流量预测.....                      | 2  |
| 1.3  | 边缘智能网络示意图.....                               | 3  |
| 1.4  | MEC 架构.....                                  | 4  |
| 1.5  | 人工智能发展历程.....                                | 6  |
| 1.6  | 云计算、边缘计算和边缘智能示意图.....                        | 7  |
| 1.7  | 物联网技术成熟曲线（来源：Gartner）.....                   | 8  |
| 1.8  | 边缘智能网络架构.....                                | 10 |
| 1.9  | 边缘智能网络的感知能力.....                             | 11 |
| 1.10 | 功能分离和功能一体化架构比较.....                          | 12 |
| 1.11 | VIVO 通信研究院展示的通感一体化原型样机 <sup>[83]</sup> ..... | 16 |
| 1.12 | 全文内容和章节安排.....                               | 20 |
| 2.1  | 通信、感知和计算功能定义.....                            | 24 |
| 2.2  | 通信、感知和计算一体化发展趋势.....                         | 27 |
| 2.3  | 通信、感知和计算一体化系统架构.....                         | 29 |
| 2.4  | 边缘智能网络中无线感知的不同方式.....                        | 30 |
| 2.5  | 大规模MIMO下通信和感知一体化波束成形示意图.....                 | 33 |
| 2.6  | 通信和感知一体化存在的干扰情况.....                         | 34 |
| 2.7  | “先通信后计算”与“空中计算”架构比较.....                     | 35 |
| 2.8  | 边缘联邦学习模型.....                                | 36 |
| 3.1  | 面向感知的通信和计算一体化系统模型.....                       | 40 |
| 3.2  | 在所有设备参与FL的情况下，不同误差大小对FL测试精度的影响.....          | 42 |
| 3.3  | 在无传输误差的情况下，参与FL的设备数量对FL测试精度的影响.....          | 42 |

|      |  |    |
|------|--|----|
| 3.4  | 基于空中计算的模型聚合框图.....                     | 44 |
| 3.5  | 使用算法 2 求解稀疏诱导问题的收敛行为.....              | 53 |
| 3.6  | 聚合误差和广播误差的要求对平均选择设备数量的影响.....          | 54 |
| 3.7  | 不同算法在不同 $\gamma_D$ 要求下的平均选择设备数量对比..... | 55 |
| 3.8  | 不同信道误差范围和不同天线数量下的平均选择设备数量.....         | 55 |
| 3.9  | 仿真实验中训练 FL 模型所用的 6 层 CNN 架构.....       | 56 |
| 3.10 | 不同算法在 IID MNIST 数据集上的 FL 性能比较.....     | 56 |
| 3.11 | 不同算法在非 IID MNIST 数据集上的 FL 性能比较.....    | 57 |
| 4.1  | 计算辅助的通信和感知一体化系统模型.....                 | 60 |
| 4.2  | 通信和感知一体化信号流图.....                      | 60 |
| 4.3  | 问题 (4.9) 到问题 (4.28) 的转换过程.....         | 67 |
| 4.4  | 基于 DL 的优化算法设计框图.....                   | 68 |
| 4.5  | PDNN 架构.....                           | 69 |
| 4.6  | 不同 NN 的训练和测试性能比较.....                  | 72 |
| 4.7  | 所提的 DL 算法与基线算法的性能比较.....               | 73 |
| 4.8  | 基站和终端的发射功率对性能的影响.....                  | 73 |
| 4.9  | 终端数量和基站天线数量对性能的影响.....                 | 74 |
| 4.10 | 感知权重和终端数量对性能的影响.....                   | 75 |
| 4.11 | 不完美 CSI 和基站接收天线数量对性能的影响.....           | 75 |
| 5.1  | 通信支撑的感知和计算一体化系统模型.....                 | 78 |
| 5.2  | 通信支撑的感知和计算一体化系统流图.....                 | 79 |
| 5.3  | 算法 5 和算法 4 的收敛行为.....                  | 90 |
| 5.4  | 算法 5 与不同算法的性能比较.....                   | 91 |
| 5.5  | 基站天线数和感知任务要求对加权总能耗的影响.....             | 91 |
| 5.6  | 终端数量和 QoS 要求对加权总能耗的影响.....             | 92 |
| 5.7  | 计算任务大小和时延要求对加权总能耗和本地计算比例的影响.....       | 93 |
| 5.8  | 基站数量和感知任务要求对加权总能耗的影响.....              | 93 |
| 5.9  | 带宽在不同 QoS 要求下对加权总能耗的影响.....            | 94 |
| 6.1  | 功能复用的通信、感知和计算一体化系统模型.....              | 98 |
| 6.2  | 功能复用的通信、感知和计算一体化系统流图.....              | 99 |



---

|      |                            |     |
|------|----------------------------|-----|
| 6.3  | 感知精度对目标成像效果的影响示例.....      | 102 |
| 6.4  | 算法 6 的收敛行为.....            | 113 |
| 6.5  | 性能偏好系数和终端发射功率对系统性能的影响..... | 114 |
| 6.6  | 终端数量和基站天线数对系统性能的影响.....    | 115 |
| 6.7  | 算法 6 与基线算法的性能比较.....       | 115 |
| 6.8  | 算法 7 的收敛行为.....            | 116 |
| 6.9  | QoS 条件对发射总功率的影响.....       | 116 |
| 6.10 | 杂波数量和基站天线数对发射总功率的影响.....   | 117 |
| 6.11 | 算法 7 与基线算法的性能比较.....       | 118 |



## 表 格

|     |                                  |     |
|-----|----------------------------------|-----|
| 2.1 | 通信、感知和计算性能评价常用指标.....            | 26  |
| 2.2 | 通信、感知和计算一体化程度区别.....             | 28  |
| 2.3 | 基于传统波形共用的一体化波形案例.....            | 32  |
| 2.4 | 适用于空中计算的目标函数例子.....              | 35  |
| 3.1 | 算法 2 的仿真参数设置.....                | 53  |
| 3.2 | 在 IID MNIST 数据集上测试手写数字识别的例子..... | 58  |
| 4.1 | 所提 DL 算法的实际运行时间 .....            | 70  |
| 4.2 | 所提 DL 算法的仿真参数设置 .....            | 71  |
| 5.1 | 算法 5 的仿真参数设置.....                | 90  |
| 6.1 | 计算精度对图像识别效果的影响示例.....            | 103 |
| 6.2 | 算法 6 和算法 7 的仿真参数设置.....          | 113 |



## 符号对照表

---

|                               |                                      |
|-------------------------------|--------------------------------------|
| $a$                           | 标量 $a$                               |
| $\mathbf{a}$                  | 向量 $\mathbf{a}$                      |
| $\mathbf{A}$                  | 矩阵 $\mathbf{A}$                      |
| $(\cdot)^H$                   | 共轭转置                                 |
| $(\cdot)^*$                   | 共轭                                   |
| $(\cdot)^T$                   | 转置                                   |
| $(\cdot)^{-1}$                | 矩阵取逆                                 |
| $\ \cdot\ _0$                 | 0 范数                                 |
| $\ \cdot\ _1$                 | 1 范数                                 |
| $\ \cdot\ $                   | 2 范数                                 |
| $\ \cdot\ _F$                 | F 范数                                 |
| $ a $                         | 标量 $a$ 的绝对值                          |
| $ \mathbf{A} $                | 矩阵 $\mathbf{A}$ 的行列式                 |
| $ \mathcal{S} $               | 集合 $\mathcal{S}$ 的元素个数               |
| $\mathbb{E}\{\cdot\}$         | 期望                                   |
| $\mathcal{CN}(\mu, \sigma^2)$ | 均值为 $\mu$ 、方差为 $\sigma^2$ 的循环对称复高斯分布 |
| $\mathcal{O}$                 | 复杂度                                  |
| $\sum$                        | 求和操作                                 |
| max                           | 取最大值操作                               |
| min                           | 取最小值操作                               |
| $\emptyset$                   | 空集                                   |
| $[\cdot]^\downarrow$          | 对向量中的元素做降序操作                         |
| $\otimes$                     | 克罗内克积运算                              |
| vec( $\cdot$ )                | 矩阵向量化操作                              |
| diag( $\mathbf{a}$ )          | 对角线元素为向量 $\mathbf{a}$ 的对角矩阵          |

|  |   |
|--|---|
| $\text{diag}[a_1, a_2, \dots, a_N]$    | 对角线元素为 $a_1, a_2, \dots, a_N$ 的对角矩阵                       |
| $\mathbf{I}_M$                         | 维度为 $M \times M$ 的单位矩阵                                    |
| $I(\mathbf{A} \mathbf{B}; \mathbf{C})$ | 在已知 $\mathbf{C}$ 的前提下, $\mathbf{A}$ 和 $\mathbf{B}$ 间的互信息量 |
| $\propto$                              | 成正比关系   |
| $\frac{\partial y}{\partial x}$        | $y$ 对 $x$ 的偏导数  |
| $\nabla$                               | 梯度算子  |
| $\nabla_x$                             | 就 $x$ 的一阶导数   |
| $\nabla_x^2$                           | 就 $x$ 的二阶导数   |
| $\Omega_N = \{1, 2, \dots, N\}$        | 从 1 到 $N$ 的自然数集合  |
| $\text{tr}(\cdot)$                     | 矩阵的迹  |
| $\text{Rank}(\cdot)$                   | 矩阵的秩  |
| $\Re(\cdot)$                           | 复数的实数部分   |
| $\Im(\cdot)$                           | 复数的虚数部分   |

---

## 缩写词列表

---

|        |   |             |
|--------|---|-------------|
| 1/2/3G | 1st/ 2nd/ 3rd Generation Mobile Communication     | 第一/二/三代移动通信 |
| 4/5/6G | 4th/ 5th/ 6th Generation Mobile Communication     | 第四/五/六代移动通信 |
| IoT    | Internet of Things                                | 物联网         |
| AI     | Artificial Intelligence                           | 人工智能        |
| MEC    | Mobile Edge Computing                             | 移动边缘计算      |
| ESTI   | European Telecommunications Standards Institute   | 欧洲电信标准协会    |
| QoS    | Quality of Service                                | 服务质量        |
| DL     | Deep Learning                                     | 深度学习        |
| CNN    | Convolutional Neural Network                      | 卷积神经网络      |
| RNN    | Recurrent Neural Network                          | 循环神经网络      |
| MLP    | Multi-Layer Perceptron                            | 多层感知机       |
| FL     | Federated Learning                                | 联邦学习        |
| GPU    | Graphics Processing Unit                          | 图形处理器       |
| CUDA   | Compute Unified Device Architecture               | 统一计算设备架构    |
| FPGA   | Field Programmable Gate Array                     | 现场可编程门阵列    |
| CPU    | Central Processing Unit                           | 中央处理器       |
| MIMO   | Multiple-Input Multiple-Output                    | 多入多出        |
| OFDM   | Orthogonal Frequency Division Multiplexing        | 正交频分复用      |
| MISO   | Multiple-Input Single-Output                      | 多入单出        |
| IEEE   | Institute of Electrical and Electronics Engineers | 电气与电子工程师协会  |
| 3GPP   | 3rd Generation Partnership Project                | 第三代合作伙伴计划   |
| ITU    | International Telecommunication Union             | 国际电信联盟      |
| IMT    | International Mobile Telecommunications           | 国际移动通信      |
| OTFS   | Orthogonal Time Frequency Space                   | 正交时频空       |
| FMCW   | Frequency-Modulated Continuous Wave               | 频率调制连续波     |

|       |   |             |
|-------|---|-------------|
| LFM   | Linear Frequency Modulation                     | 线性频率调制      |
| MSE   | Mean Square Error                               | 均方误差        |
| AWGN  | Additive White Gaussian Noise                   | 加性高斯白噪声     |
| CSI   | Channel State Information                       | 信道状态信息      |
| AO    | Alternating Optimization                        | 交替优化        |
| LMI   | Linear Matrix Inequality                        | 线性矩阵不等式     |
| SOC   | Second Order Cone                               | 二阶锥         |
| DC    | Difference-of-Convex                            | 凸差          |
| IID   | Independent Identically Distribution            | 独立同分布       |
| ReLU  | Rectified Linear Unit                           | 线性整流单元      |
| TRM   | Target Response Matrix                          | 目标响应矩阵      |
| SINR  | Signal to Interference plus Noise Ratio         | 信干噪比        |
| MOOP  | Multi-Objective Optimization Problem            | 多目标优化问题     |
| SOOP  | Single Objective Optimization Problem           | 单目标优化问题     |
| SVD   | Singular Value Decomposition                    | 奇异值分解       |
| EVD   | Eigenvalue Decomposition                        | 特征值分解       |
| TCM   | Transmit Correlation Matrix                     | 发射相关矩阵      |
| PDNN  | Proposed Deep Neural Network                    | 所提的深度神经网络   |
| I/Q   | In-Phase/Quadrature                             | 同相/正交       |
| FCNN  | Fully-Connected Neural Network                  | 全连接神经网络     |
| CRB   | Cramer-Rao Bound                                | 克拉美罗界       |
| FIM   | Fisher Information Matrix                       | 费舍尔信息矩阵     |
| MINLP | Mixed Integer Nonlinear Programming             | 混合整数非线性规划   |
| BnB   | Branch and Bound                                | 分支定界        |
| SCA   | Successive Convex Approximation                 | 连续凸逼近       |
| ROI   | Range of Interest                               | 兴趣范围        |
| RMS   | Root Mean Square                                | 均方根         |
| AirFL | Over-the-Air Federated Learning                 | 基于空中计算的联邦学习 |
| MMSE  | Minimum Mean Square Error                       | 最小均方误差      |
| QCQP  | Quadratically Constrained Quadratic Programming | 二次约束二次规划    |
| SDP   | Semidefinite Programming                        | 半正定规划       |



缩写词列表

---

|          |                            |               |
|----------|----------------------------|---------------|
| SDR      | Semidefinite Relaxation    | 半正定松弛         |
| F-MMSEBF | Fixed MMSE Beamforming     | 固定的 MMSE 波束成形 |
| MFBF     | Matched Filter Beamforming | 匹配滤波波束成形      |
| ZFBF     | Zero Forcing Beamforming   | 迫零波束成形        |

---



# 目 录

|                       |      |
|-----------------------|------|
| 致谢 .....              | I    |
| 摘要 .....              | III  |
| <b>Abstract</b> ..... | V    |
| 插图 .....              | VII  |
| 表格 .....              | XI   |
| 符号对照表.....            | XIII |
| 缩写词列表.....            | XV   |
| 目录                    |      |
| <b>1 绪论</b> .....     | 1    |
| 1.1 研究背景 .....        | 1    |
| 1.2 边缘智能概述 .....      | 3    |
| 1.2.1 移动边缘计算.....     | 3    |
| 1.2.2 人工智能.....       | 5    |
| 1.2.3 边缘智能.....       | 7    |
| 1.3 边缘智能网络 .....      | 9    |
| 1.3.1 网络架构.....       | 9    |
| 1.3.2 网络能力 .....      | 11   |
| 1.3.3 网络需求.....       | 12   |
| 1.4 通感算一体化研究现状 .....  | 13   |
| 1.4.1 通信和计算一体化.....   | 13   |
| 1.4.2 通信和感知一体化.....   | 15   |
| 1.4.3 感知和计算一体化.....   | 17   |
| 1.5 研究挑战 .....        | 18   |
| 1.6 研究内容和章节安排 .....   | 20   |

|                                |    |
|--------------------------------|----|
| <b>2 通感算一体化基础理论和关键技术</b> ..... | 23 |
| 2.1 引言 .....                   | 23 |
| 2.2 基础理论 .....                 | 24 |
| 2.2.1 基本要素 .....               | 24 |
| 2.2.2 性能指标 .....               | 25 |
| 2.2.3 一体化程度 .....              | 27 |
| 2.2.4 系统架构 .....               | 29 |
| 2.3 关键技术 .....                 | 30 |
| 2.3.1 无线感知技术 .....             | 30 |
| 2.3.2 一体化波形技术 .....            | 31 |
| 2.3.3 大规模MIMO和波束成形技术 .....     | 32 |
| 2.3.4 干扰管理技术 .....             | 33 |
| 2.3.5 空中计算技术 .....             | 34 |
| 2.3.6 边缘联邦学习技术 .....           | 36 |
| 2.4 本章小结 .....                 | 37 |
| <b>3 面向感知的通信和计算一体化</b> .....   | 39 |
| 3.1 引言 .....                   | 39 |
| 3.2 系统模型 .....                 | 40 |
| 3.2.1 计算模型 .....               | 40 |
| 3.2.2 通信模型 .....               | 43 |
| 3.3 联合设备选择和收发机设计 .....         | 45 |
| 3.3.1 问题构造 .....               | 45 |
| 3.3.2 算法设计 .....               | 46 |
| 3.3.3 算法分析 .....               | 50 |
| 3.4 仿真结果 .....                 | 52 |
| 3.4.1 通信性能 .....               | 53 |
| 3.4.2 计算性能 .....               | 56 |
| 3.5 本章小结 .....                 | 58 |
| <b>4 计算辅助的通信和感知一体化</b> .....   | 59 |
| 4.1 引言 .....                   | 59 |
| 4.2 系统模型 .....                 | 60 |

|          |                               |           |
|----------|-------------------------------|-----------|
| 4.2.1    | 感知信息提取 .....                  | 61        |
| 4.2.2    | 通信信号解码 .....                  | 62        |
| 4.3      | 联合感知波形和通信波束设计 .....           | 63        |
| 4.3.1    | 感知速率最大化问题 .....               | 63        |
| 4.3.2    | 通信速率最大化问题 .....               | 65        |
| 4.3.3    | 问题转化 .....                    | 67        |
| 4.4      | 基于 DL 的算法设计 .....             | 67        |
| 4.4.1    | 预处理模块 .....                   | 68        |
| 4.4.2    | 后处理模块 .....                   | 68        |
| 4.4.3    | 网络骨架 .....                    | 69        |
| 4.4.4    | 恢复模块 .....                    | 70        |
| 4.4.5    | 复杂度分析 .....                   | 70        |
| 4.5      | 仿真结果 .....                    | 71        |
| 4.6      | 本章小结 .....                    | 76        |
| <b>5</b> | <b>通信支撑的感知和计算一体化 .....</b>    | <b>77</b> |
| 5.1      | 引言 .....                      | 77        |
| 5.2      | 系统模型 .....                    | 78        |
| 5.2.1    | 感知任务 .....                    | 80        |
| 5.2.2    | 计算任务 .....                    | 81        |
| 5.3      | 联合卸载选择和资源分配设计 .....           | 82        |
| 5.3.1    | 问题构造 .....                    | 83        |
| 5.3.2    | 算法设计 .....                    | 84        |
| 5.3.3    | 算法分析 .....                    | 88        |
| 5.4      | 仿真结果 .....                    | 89        |
| 5.5      | 本章小结 .....                    | 94        |
| <b>6</b> | <b>功能复用的通信、感知和计算一体化 .....</b> | <b>97</b> |
| 6.1      | 引言 .....                      | 97        |
| 6.2      | 系统模型 .....                    | 98        |
| 6.2.1    | 信息通信 .....                    | 100       |
| 6.2.2    | 目标感知 .....                    | 101       |
| 6.2.3    | 模型计算 .....                    | 101       |

|                               |            |
|-------------------------------|------------|
| 6.3 联合发射和接收波束成形设计 .....       | 103        |
| 6.3.1 加权总性能最大化设计 .....        | 104        |
| 6.3.2 发射总功率最小化设计 .....        | 109        |
| 6.3.3 算法分析 .....              | 111        |
| 6.4 仿真结果 .....                | 112        |
| 6.5 本章小结 .....                | 118        |
| <b>7 总结与展望</b> .....          | <b>119</b> |
| 7.1 研究总结 .....                | 119        |
| 7.2 未来展望 .....                | 120        |
| 参考文献 .....                    | 123        |
| 附录 .....                      | 133        |
| A.1 第 3 章相关证明 .....           | 133        |
| A.1.1 引理 3.2 的证明 .....        | 133        |
| A.2 第 4 章相关证明 .....           | 134        |
| A.2.1 引理 4.1 的证明 .....        | 134        |
| A.2.2 引理 4.2 的证明 .....        | 134        |
| A.2.3 $M_s$ 和 $M_c$ 的推导 ..... | 135        |
| A.3 第 5 章相关证明 .....           | 136        |
| A.3.1 FIM 的推导 .....           | 136        |
| A.3.2 命题 5.1 的证明 .....        | 137        |
| A.3.3 问题 (5.23) 的凸性证明 .....   | 138        |
| A.4 第 6 章相关证明 .....           | 139        |
| A.4.1 定理 6.1 的证明 .....        | 139        |
| A.4.2 定理 6.2 的证明 .....        | 140        |
| 攻读博士学位期间参与项目及研究成果 .....       | 143        |

# 1 绪论

## 1.1 研究背景

在过去几十年，移动通信实现了从 1G 到 5G 的飞跃式发展<sup>1</sup>。如图 1.1 所示，1G 实现了语音业务，2G 开启了数字通信时代，3G 开始支持多媒体业务，4G 迎来了移动互联网，彻底改变了人们的生活方式。与前四代网络相比，5G 首次将应用场景扩展到了物联网（Internet of Things, IoT）领域，通过与垂直行业的深度融合，致力于构建“万物互联”的信息社会<sup>[1-3]</sup>。根据移动通信十年一更迭的规律，在 5G 商用之际，业界已将重点转向 6G 网络的研究，以克服当前 5G 网络的局限性和瓶颈，旨在实现“万物智联”的美好愿景<sup>[4-6]</sup>。

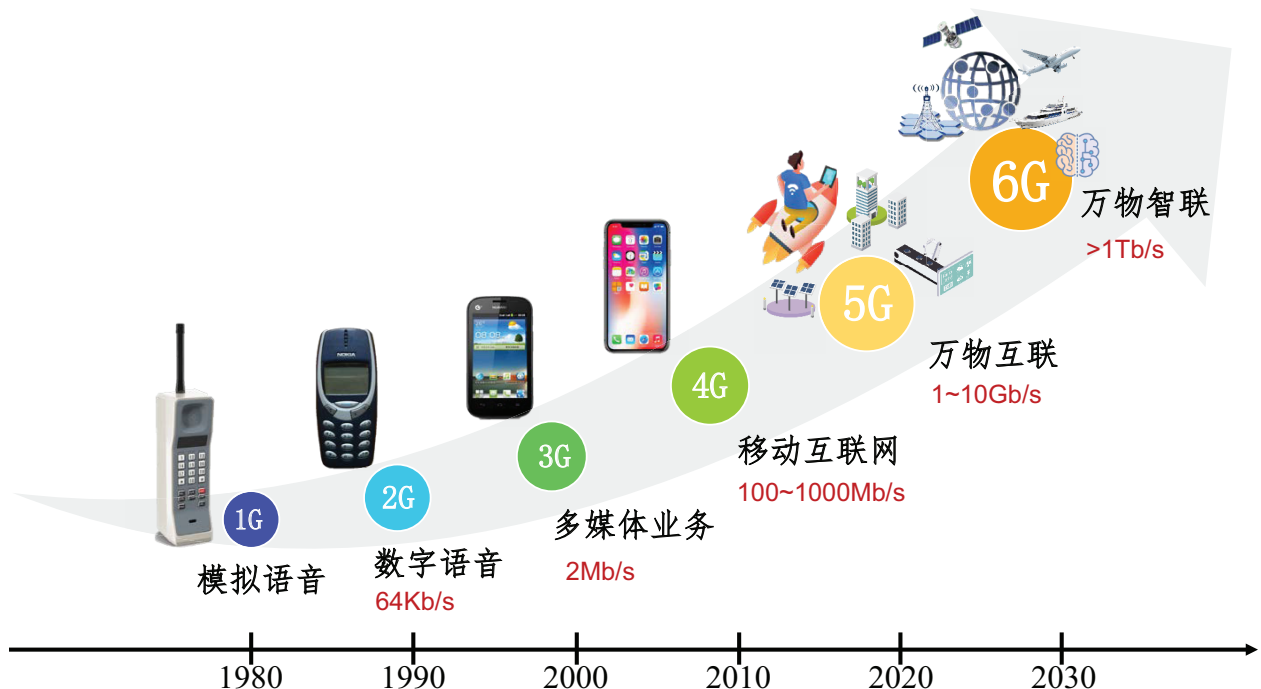


图 1.1 移动通信发展史

5G 网络的大规模部署为 IoT 的数字化和自动化开辟了新道路，使得网络连接数量和

<sup>1</sup> 第一/二/三/四/五/六代移动通信 (1st/ 2nd/ 3rd/ 4th/ 5th/ 6th Generation Mobile Communication, 1/2/3/4/5/6G)



图 1.2 全球物联网设备连接数量和数据流量预测

数据流量空前激增。如图 1.2 所示，到 2030 年预计有近 300 亿台 IoT 设备同时在线<sup>[7]</sup>，全球移动数据流量将超过 5000 EB/月<sup>[8]</sup>。与此同时，人工智能（Artificial Intelligence, AI）技术正蓬勃发展，渗透到了各行各业，促使了通信网络与 AI 的不断融合。在当前的 5G 网络中，AI 已被用作附加模块来提升网络性能<sup>[9-11]</sup>。然而其应用方式还是以集中式智能为主，通过在云计算中心汇聚大量的数据，利用集中的算力对数据进行分析、处理和推演<sup>[12, 13]</sup>。在未来大规模节点和海量数据的场景下，这种中心智能的方式显然给通信传输带宽、数据处理实时性以及安全隐私性都带来了很大的压力和挑战。因此，无线网络正逐渐从中心智能向边缘智能演进<sup>[14-16]</sup>，催生出具有“边缘智能”能力的新型网络范式，即边缘智能网络，旨在提供无处不在的实时智能服务。

如图 1.3 所示，边缘智能网络作为下一代移动通信系统的核心组成部分，通过通信、感知和计算的协同与交互，拓展了智慧交通、智慧农业、智慧工业和智慧医疗等垂直应用场景，可以提供全息通信、数字孪生、扩展现实等全新业务，推动了真实物理世界与虚拟数字世界的深度融合。值得注意的是，这些数据驱动的新兴智能业务不仅对通信和计算提出了更加极致的性能要求，还对感知提出了高精度、高分辨率的需求<sup>[17, 18]</sup>。在这种情况下，边缘智能网络需要支持海量节点的随遇接入，并提供超高精度、超高可靠和超低时延的信息感知、传输和处理服务。然而，传统无线网络的通信和、感知和计算功能是独立存在的。例如，传感器网络只负责采集数据<sup>[19]</sup>，通信网络只负责传输数据<sup>[20]</sup>，云计算中心只负责处理数据<sup>[21]</sup>。这种分离化的设计不仅会造成本就稀缺的频谱资源与昂贵的硬件资源的严重浪费，还难以满足边缘智能网络中新兴智能应用的实时性要求。

因此，迫切需要研究面向边缘智能网络的通信、感知和计算一体化设计。一般而言，通信、感知和计算一体化的关键是根据边缘智能网络的特点，进行多维资源的协同与共



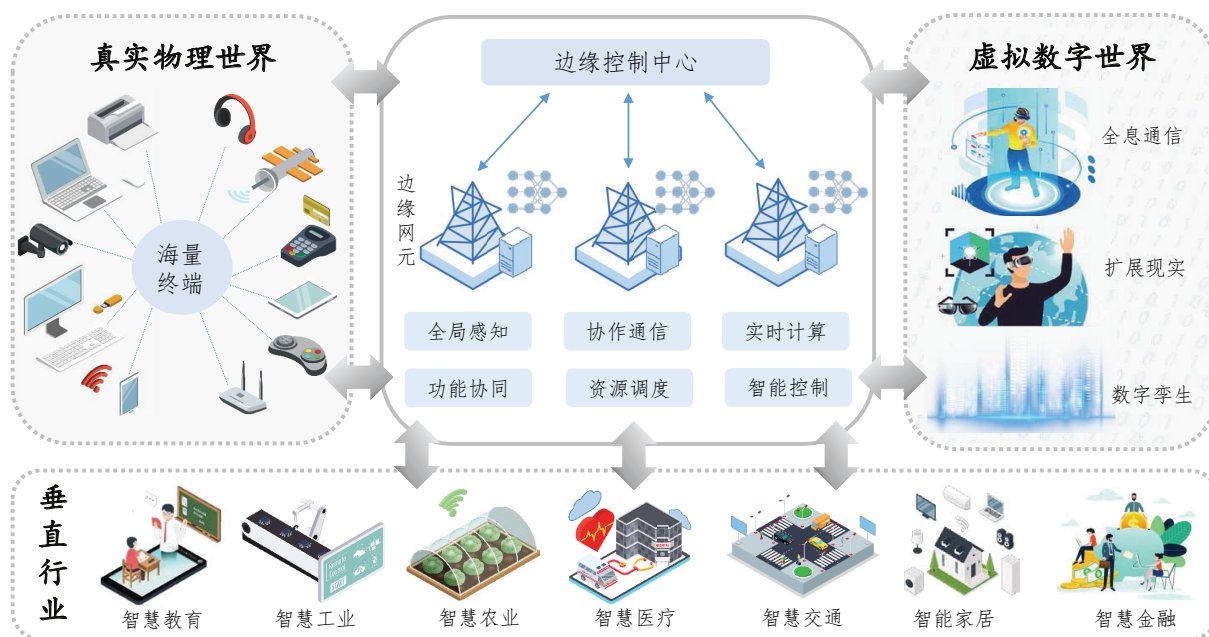


图 1.3 边缘智能网络示意图

享，实现通信、感知和计算功能的深度融合和互惠增强，从而有效提升网络资源利用率和系统整体性能。基于此，本文深入研究了面向边缘智能网络的通信、感知和计算一体化基础理论、关键技术、典型架构和核心算法，为下一代移动通信网络和未来物联网的发展提供重要的理论基础和技术支撑。

## 1.2 边缘智能概述

由移动边缘计算（Mobile Edge Computing, MEC）技术和 AI 技术驱动的边缘智能被认为是当前 5G 网络缺失的关键元素之一，由此也成为了下一代移动通信网络中非常有前景的解决方案，以支持其新功能和 new 服务。本小节通过简要回顾 MEC 技术和 AI 技术，阐述边缘智能的必要性和可行性，并详细介绍其当前的研究进展。

### 1.2.1 移动边缘计算

MEC 技术的发展历史可以追溯到 1998 年，阿卡迈公司推出了第一个商业化的内容分发网络，通过将网站内容缓存在全球各地的服务器上，显著提高了用户访问网站的响应速度<sup>[22]</sup>。2008 年，卡内基梅隆大学的研究人员提出了 Cloudlet 的概念，利用本地服务器来提供更快速、更可靠的云服务<sup>[23]</sup>。随着物联网和智能设备的普及，在网络边缘进行实时计算的需求越来越大，雾计算<sup>[24]</sup>和边缘计算<sup>[25]</sup>的概念相继被提出。随后，欧洲电信标准

协会（European Telecommunications Standards Institute, ESTI）于 2014 年正式启动对 MEC 的标准化研究<sup>[26]</sup>，并于 2016 年将“MEC”中的“M”扩展为多址接入的含义<sup>[27]</sup>。2018 年起，MEC 开始在 5G 网络中得到广泛应用，并成为 5G 网络的重要组成部分。

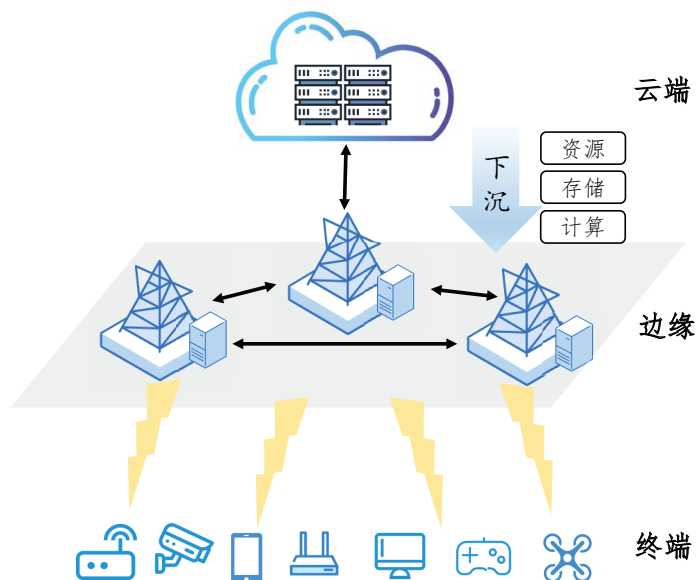


图 1.4 MEC 架构

如图 1.4 所示，MEC 的原理是将云计算中心的资源、存储和计算等能力下沉到网络边缘，通过部署具有通信、计算和存储等功能的边缘节点和服务器，在靠近数据产生侧实时处理数据，实现网络业务本地化<sup>[28, 29]</sup>。与传统集中式处理的移动云计算相比，MEC 具有如下优点：

- **减小时延：** MEC 利用部署在网络边缘的终端或服务器提供数据处理服务，从而大大降低计算所需的通信时延。特别是对于时延敏感的应用，MEC 的低时延特性显得尤为重要。例如，在传统的视频传输方式下，每个终端请求服务响应首先需要经由基站接入，然后连接核心网，在云端服务器上获取缓存内容后逐层回传。显然，这种传统的集中式处理方式会导致过长的响应时间。而利用 MEC 技术，可以直接在具备通信、计算和存储功能的本地节点或边缘服务器上执行任务，由于无需等待云计算中心的响应，可以大大减少通信时延，提高用户的服务质量（Quality of Service, QoS）。
- **节省带宽：** MEC 服务器部署在网络边缘，能够在本地对流量数据进行卸载，从而有效减缓传输带宽的压力。例如目前流行的网络直播，吸引了大量用户同时访问和请求同一资源，这对网络带宽和链路状态都提出了极高的要求。传统的集中式处理方

式很难承受这种高并发需求带来的带宽压力。利用 MEC 服务器实时缓存内容资源，在本地处理用户请求，能够有效减轻回程链路的带宽压力，大大降低网络拥塞和故障的可能性，提高链路容量。

- **降低能耗：**网络能量消耗主要由数据传输的通信能耗和数据处理的计算能耗组成。由于 MEC 系统的边缘节点和服务器具具备计算和存储能力，一方面可以避免将海量数据传输至遥远的云计算中心，降低通信能耗。另一方面，通过获取终端信道信息和网络状态信息，协调多个边缘节点和服务器的计算资源，以能量有效的方式在本地执行数据处理，减少计算能耗。
- **保护隐私：**由于 MEC 将计算资源和应用程序部署在接近数据源的网络边缘，使得数据处理可以在本地进行，而不需要将数据统一传输到远程云端服务器。这不仅在一定程度上起到了保护用户隐私的作用，同时还大大降低了数据在传输过程中被窃取或篡改的风险。

### 1.2.2 人工智能

AI 是计算机科学的一个重要分支，指机器执行特定任务时，通过从海量数据中挖掘和学习有效信息，模拟人类智能的行为。如图 1.5 所示，AI 的历史可以追溯到 20 世纪 50 年代，艾伦·麦席森·图灵首次提出用机器模拟人类思考的思想<sup>[30]</sup>。经过几十年的发展，AI 已经成为学术界和工业界最热门的领域之一，其研究和应用范围相当广泛，包括机器人<sup>[31]</sup>、专家系统<sup>[32]</sup>、计算机视觉<sup>[33]</sup>和自然语言处理<sup>[34]</sup>等。

机器学习是实现人工智能的典型方法，也是 AI 的一个重要分支，其目的是使机器具有自主地从数据中学习和改进其性能的能力<sup>[35]</sup>。简单来说，机器学习就是通过让计算机自动分析和理解数据，以便根据数据中的模式和趋势做出预测或决策的过程。总体上看，机器学习有三个主要分支：监督学习、无监督学习和强化学习。

- **监督学习<sup>[36]</sup>：**监督学习利用标记的训练数据来构建模型，以便能够正确地分类或做出预测。具体而言，算法学习从输入映射到输出的未知函数，并不断最小化反映输入和输出匹配程度的损失函数。监督学习的常用算法有逻辑回归、支持向量机、朴素贝叶斯、决策树和 K 近邻等。
- **无监督学习<sup>[37]</sup>：**无监督学习利用未标记的训练数据来构建模型，旨在探索数据中的模式和规律。具体而言，算法自主学习数据中的相关性，以及数据之间的复杂关系。无监督学习的常用算法有聚类、降维和关联规则学习等。

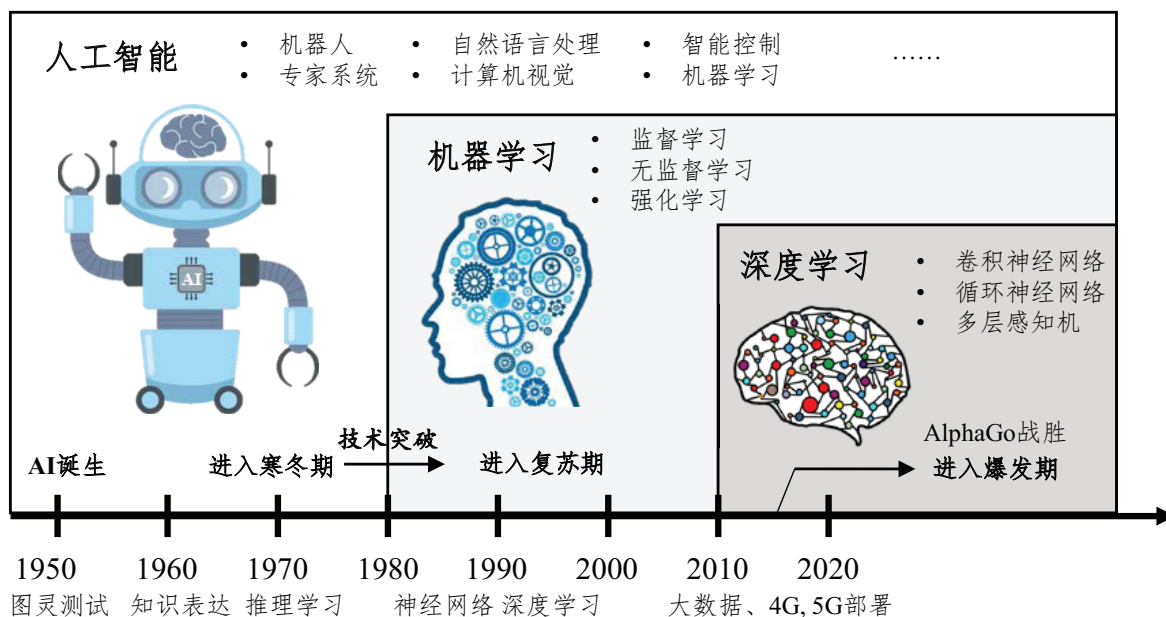


图 1.5 人工智能发展历程

- 强化学习<sup>[38]</sup>：强化学习利用反馈机制来构建模型，以实现实时决策。具体而言，智能代理通过与周围环境互动，不断尝试，学习如何最大限度地改善结果，以达到最佳状态。例如，著名的 AlphaGo 就是基于强化学习设计的。强化学习常用的算法有蒙特卡洛树搜索、Q 学习、策略梯度和深度 Q 学习等。

深度学习（Deep Learning, DL）是一种高效的机器学习算法，其核心是通过构建多层神经网络来实现对数据高度自动化的特征提取和分类，从而达到比传统机器学习算法更高的准确率和性能<sup>[39]</sup>。深度学习主要有三大类神经网络：卷积神经网络（Convolutional Neural Network, CNN）、循环神经网络（Recurrent Neural Network, RNN）和多层感知机（Multi-Layer Perceptron, MLP）。

- CNN<sup>[40]</sup>：CNN 具有高效的特征提取机制和强大的学习能力，尤其适合用来处理具有类似网格拓扑结构的数据，在图像处理和计算机视觉等方面应用非常广泛。CNN 主要由输入层、卷积层、池化层、全连接层和输出层组成。其中卷积层用于提取输入的特征，池化层用来减少特征的维度，而全连接层的作用则是将提取到的特征映射转换为预测结果。
- RNN<sup>[41]</sup>：RNN 具有递归结构，能够保留前后信息，并且可以根据之前的输入来影响后续的输出，尤其适合用来处理具有序列结构的数据，如时间序列数据、语音信号、

自然语言文本等。在 RNN 中，每个时间步都有一个输入和一个输出，同时还有一个隐藏状态，用于存储之前所有时间步的信息。

- **MLP<sup>[42]</sup>**: MLP 具有强大的拟合能力，可以学习非线性判别式模型，尤其适合用来解决复杂的分类和回归问题。MLP 是一种基于前向传播机制的深度前馈网络，由多个感知机组成。其中感知机是一种二分类模型，其原理是将输入向量乘以权重向量并加上偏置项，然后将结果传递给一个激活函数进行变换，最终输出一个二进制值。

### 1.2.3 边缘智能

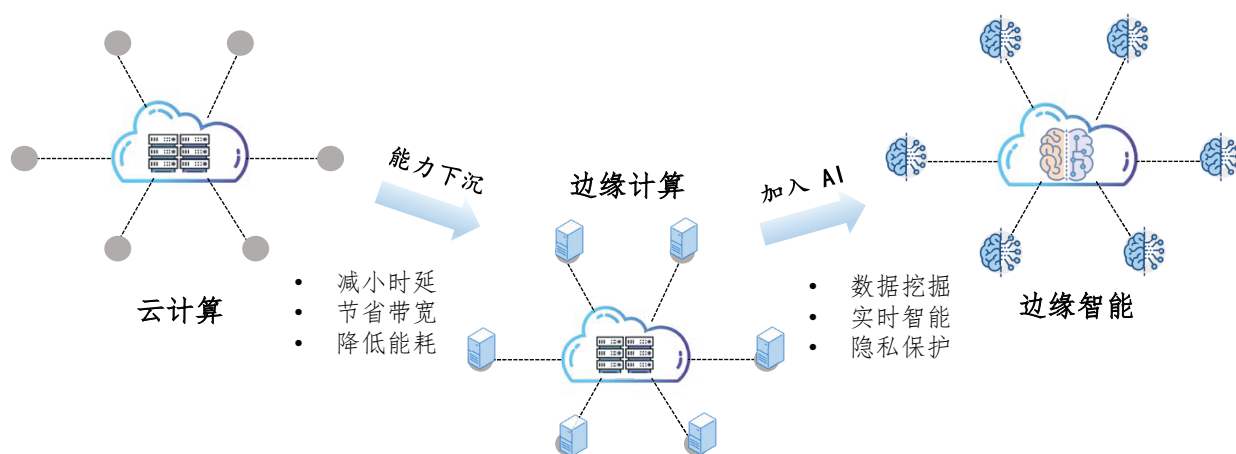


图 1.6 云计算、边缘计算和边缘智能示意图

MEC 和 AI 的结合是必然的，因为它们之间有明显的交集，由此催生出了边缘智能。如图 1.6 所示，边缘智能除了能够享受 MEC 带来的好处之外，MEC 和 AI 之间也能够相互作用、互利互惠：

- 一方面，网络边缘生成的海量数据需要利用 AI 来充分释放其潜力。随着万物互联的蓬勃发展，网络边缘的智能设备，如智能手机、智能家居设备和智能传感器等，会实时产生大量的数据，包括图像、视频、文本和环境数据等。边缘计算可以使边缘终端或服务器在本地进行数据分析和处理，从而避免将大量数据传输到云端再进行计算，能够有效节约带宽、节省能耗以及降低时延。边缘终端和服务器通常只有有限的计算和存储资源，因此可以部署一些轻量级的 AI 算法来充分释放这些数据的潜力，通过快速分析海量数据并从中提取出有效信息，进一步做出实时反馈和预测。
- 另一方面，MEC 能够进一步推动 AI 的发展和應用。数据是 AI 应用的核心，海量高质量的数据可以为 AI 模型提供丰富的训练和验证样本。随着物联网的发展和 MEC

的普及，大量的数据在网络边缘产生和收集，为 AI 应用提供了充足的数据基础。不仅如此，MEC 还能够有效解决 AI 应用中的数据隐私和安全问题。此外，应用场景是 AI 技术得以实现的重要基础。随着信息技术的快速发展，一系列具有计算密集性、数据隐私性和延迟敏感性的边缘应用不断涌现，如智慧医疗、智慧交通、智能家居、智能制造等，为 AI 技术的应用提供了广阔的发展空间和实践平台，进一步推动了 AI 的发展。

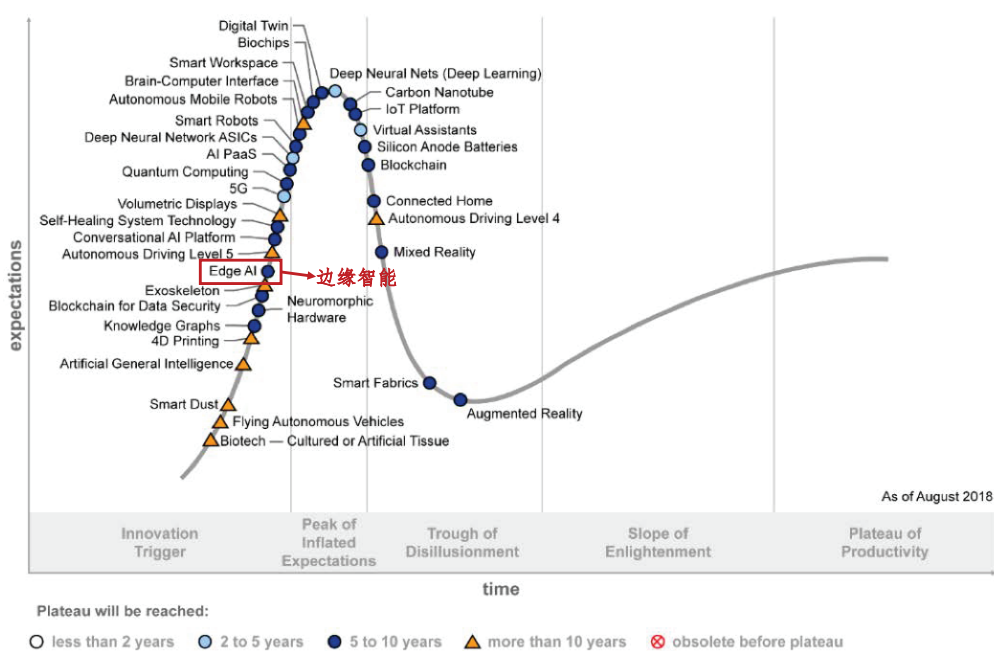


图 1.7 物联网技术成熟曲线（来源：Gartner）

鉴于 MEC 和 AI 结合的需求和优势，边缘智能应运而生，通过将计算资源和 AI 算法部署在网络边缘，利用本地节点和边缘服务器完成海量数据的实时智能处理。基于这一特性，边缘智能受到了广泛的关注。2018 年 8 月，边缘智能的概念首次出现在 Gartner 发布的物联网技术成熟曲线中<sup>[43]</sup>。如图 1.7 所示，Gartner 预测边缘智能仍处于创新触发阶段，并将在未来五到十年内达到生产力的平稳期。随后，涌现出一系列关于边缘智能的研究。例如，文献[14]对边缘智能进行了全面的调研和总结，并给出了边缘智能的四个基本组成部分，即边缘缓存、边缘训练、边缘推理和边缘卸载。文献[15]明确指出了 AI 运行在网络边缘的动机，并给出了边缘智能的总体框架、关键技术以及未来研究方向。文献[16]从更广泛的角度定义了边缘智能，将其分为“AI for edge”和“AI on edge”两个方面。前者侧重于借助 AI 技术为 MEC 系统的关键问题提供更优的解决方案，后者则研究如何在网络边缘部署 AI 算法和模型。实际上，早已有大量工作尝试将边缘智能应用于现实生活中，初步证明了边缘智能的可行性和有效性。例如，文献[44]通过智能手机和边缘服务器实现

了人脸识别应用程序。结果表明，与云端智能相比，延迟从 900 ms 降低至 169 ms。文献[45]证明了使用 Cloudlet 在可穿戴认知辅助设备上执行识别任务可节省 30% 至 40% 的能量消耗。文献[46]的研究人员在智能手机上成功实现了基于 DL 的行为识别，并取得了比浅层神经网络更好的性能，这表明普通的智能终端也可以部署轻量级的 DL 模型来执行复杂的计算任务。此外，由谷歌研究团队开发的智能输入法 Gboard 是最著名的边缘智能用例之一，它利用联邦学习（Federated Learning, FL）支持智能手机基于用户的输入历史和使用习惯在本地训练和运行 AI 模型，以实时预测用户键入的下一个单词、短语或句子，并提供自动纠正和自动完成功能<sup>[47]</sup>。

为了将智能带到网络边缘，工业界也对边缘智能进行了诸多尝试。在软件方面，传统云供应商纷纷推出了边缘智能平台，如谷歌的 Cloud IoT Edge、亚马逊的 AWS Greengrass 和微软的 Azure IoT Edge，支持边缘设备在本地运行预训练的 AI 模型进行实时决策和推演。在硬件方面，用于网络边缘运行 AI 模型的芯片和图形处理器（Graphics Processing Unit, GPU）已经在市场上商业化。例如，以英伟达 Jetson 系列为代表的嵌入式处理器，适合运行 DL 模型，支持统一计算设备架构（Compute Unified Device Architecture, CUDA）和 Tensorflow，可用于智慧城市、智慧医疗等场景中的终端设备；以英特尔 Movidius 系列和谷歌 Edge TPU 为代表的 GPU，适合部署在边缘服务器或边缘节点上，实现高效推理；以赛灵思 Zynq 系列为代表的现场可编程门阵列（Field Programmable Gate Array, FPGA）芯片，具有高度的可编程性和灵活性，可支持多种 DL 框架和模型。

### 1.3 边缘智能网络

本小节详细介绍具有“边缘智能”的新型无线网络，即边缘智能网络，包括其网络架构、网络能力和网络需求。

#### 1.3.1 网络架构

如图 1.8 所示，边缘智能网络主要由大量的多功能边缘节点、多个一体化网元和一个边缘控制中心组成，它们之间通过数据总线和控制总线进行交互。与以往侧重于单一业务功能与服务的节点和网元不同，边缘智能网络的节点和网元对通信、感知和计算的资源与能力进行了全面整合与深度融合。

- 边缘节点：通信节点、感知节点和计算节点是传统无线网络的普通节点，如传感器、IoT 设备和用户终端等，而一体化节点是集通信节点、感知节点和计算节点的部分功

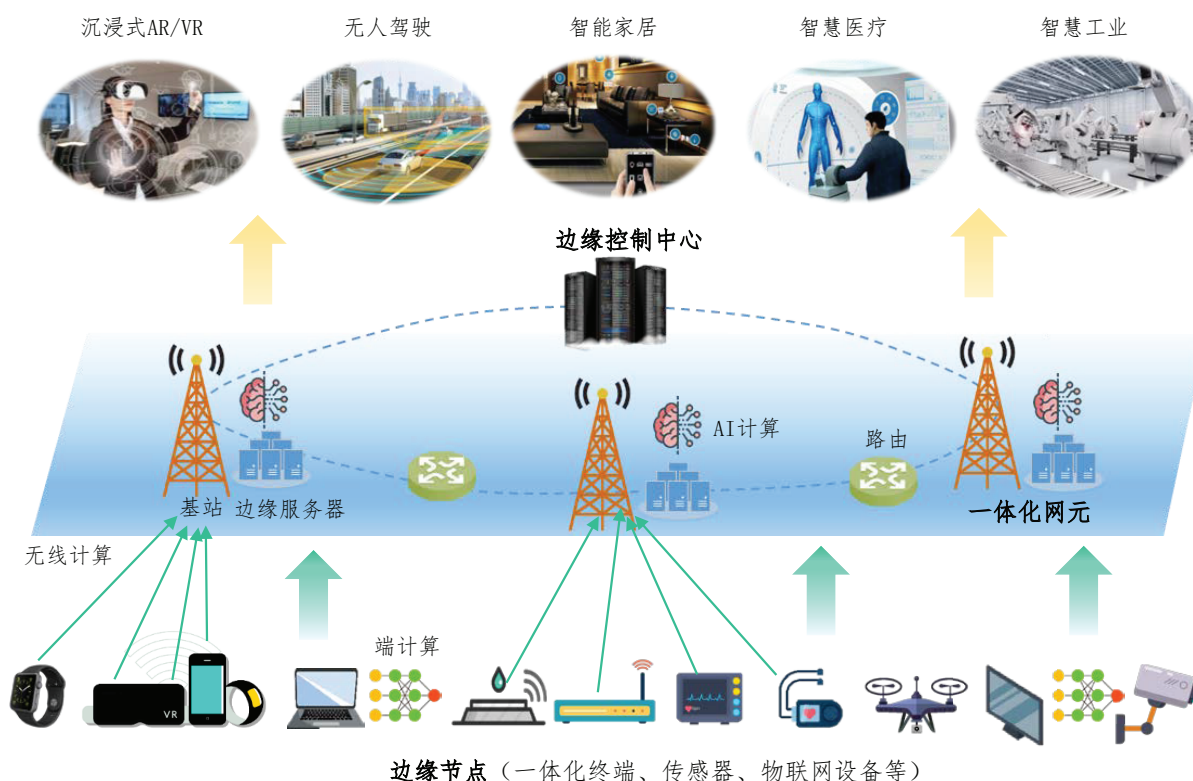


图 1.8 边缘智能网络架构

能或者全部功能于一体的多功能节点。这些多功能节点通常部署在网络边缘，它们同时具有用于发送和接收数据的通信能力，用于捕获目标状态和采集环境数据的感知能力，以及用于处理数据和制定决策的计算能力。例如，在智能家居场景中，多功能一体化终端可以实时监测并控制室内温湿度，同时远程操控智能家电系统。在工业自动化场景中，多功能一体化终端可以实时监测机器状态和车间安全，并控制生产过程。在智慧医疗场景中，多功能一体化终端可以实时监测、跟踪和记录患者健康情况以及提供远程诊疗服务。

- 一体化网元：一体化网元由基站和边缘服务器组成，其中基站为通信和感知一体化的双功能基站，不仅具有通信功能，还被赋予了感知能力。由于双功能基站配备了边缘计算服务器，因此一体化网元是集通信、感知和计算功能于一体的多功能网元。特别是，一体化网元可以利用基站的大规模天线阵列和服务器的强大算力资源在网络边缘侧同时实现更高速率的传输、更高精度的感知和更高效率的计算。
- 边缘控制中心：边缘控制中心同样部署在网络边缘侧，能够进行数据交换、算力调度和业务编排，是边缘智能网络的大脑。多个一体化网元可通过与边缘控制中心的



交互来进行协同通信、联合感知和合作计算。边缘控制中心基于网络全局感知，充分利用并调度通信、感知和计算多维资源，实现面向任务需求的业务编排，保障用户服务质量以及网络可靠运行。

### 1.3.2 网络能力

为了给不断涌现的新兴智能应用场景，如无人驾驶、全息通信和感官互联等，提供更加智能化、高效化的服务和解决方案，边缘智能网络不仅需要拥有“边缘智能”能力，同时还需具备全局感知能力，如图 1.9 所示。

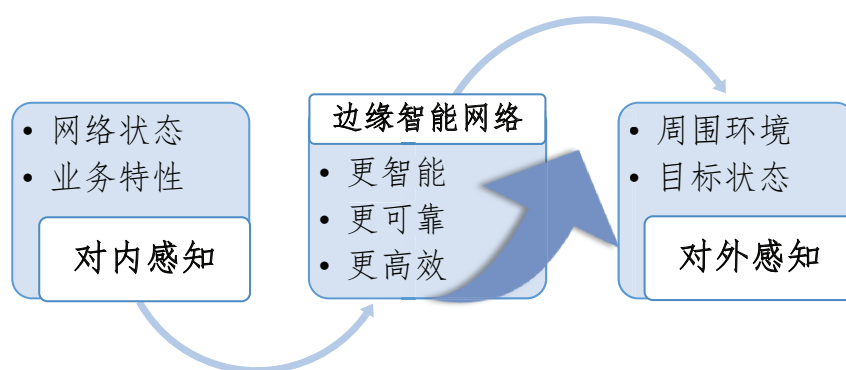


图 1.9 边缘智能网络的感知能力

- **对外感知：**通过感知和理解周围环境的信息，可以使边缘智能网络更加智能化，能够自适应满足业务需求，提高处理效率和提升服务质量。具体而言，通过感知能力，边缘智能网络可以汇聚来自各种边缘节点和网元的感知数据，并利用强大的计算能力从中提取出有用的信息，从而进行高效的推演、决策和控制。例如，在智能家居领域，通过感知能力，边缘智能网络可以获取家庭成员的活动状态和室内环境信息，从而自动控制和调节室内温湿度、照明系统和智能窗帘等，提供更加舒适的居住环境。在工业互联网领域，边缘智能网络通过感知能力实时监测生产线的运行状态、设备状况和工厂环境的变化，实现智能化的生产调度和故障预测，提高了生产效率和质量。
- **对内感知：**通过感知网络状态和业务特性，实现自适应资源管理和优化策略，提升网络的性能及其可靠性，从而更好地满足不同应用的需求。具体而言，在边缘智能网络中，由于终端数量庞大且服务需求复杂，网络状态和业务特性会不断变化。通过感知网络状态，可以及时发现网络中的问题，如网络拥塞、带宽瓶颈等，从而采

取相应的措施进行优化和调整，提高网络的性能和可靠性。同时，感知业务特性可以帮助边缘智能网络更好地管理网络资源，如带宽、存储和算力等，从而根据业务需求进行自适应调度，提高网络资源的利用率和系统性能。例如，在智能交通领域，通过感知网络状态和业务特性，边缘智能网络可以及时发现交通拥堵的情况，并采取相应的措施进行调整，如采取特殊区域限行、调整红绿灯的时长等，从而改善交通状况。在视频监控领域，边缘智能网络可以感知视频流的特性，如分辨率、码率等，从而根据业务需求进行优先级调度，保障重要视频流的传输质量。

### 1.3.3 网络需求

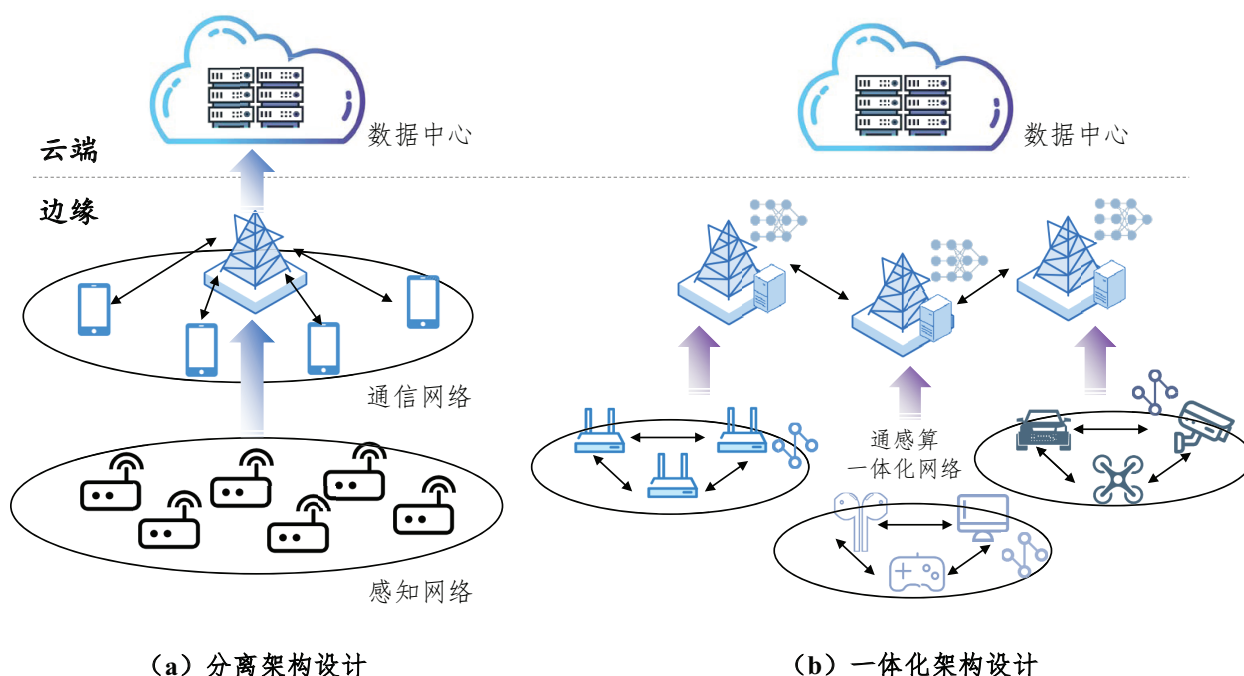


图 1.10 功能分离和功能一体化架构比较

随着物联网的快速发展，越来越多的传感器和智能设备部署在网络边缘。与此同时，信息通信技术的不断突破催生出一系列新兴业务与智能应用，对网络提出了多维极致性能要求。在这种情况下，边缘智能网络需要在有限的资源下实现高精度感知、高速率通信和高效率计算。如图 1.10 (a) 所示，在过去传统的系统中，通信、感知和计算功能都是单独设计的。例如，雷达系统或传感器网络只专注于测速、测距或采集环境数据，无线网络只是传输信息的管道，而云数据中心只负责大数据计算。这不仅浪费了现有资源，还制约了网络性能的提升，更无法满足未来数据驱动的智能服务的高性能需求。因此，面向边缘智能网络的通信、感知和计算一体化研究迫在眉睫。如图 1.10 (b) 所示，通过在网络边缘部

署集通信、感知和计算部分功能或全部功能于一体的多功能节点和网元，利用资源共享、多点协作和智能交互的方式，实现对海量数据更快速、更准确、更智能的获取、传输和处理，从而提供按需定制的通信、感知和计算服务。

## 1.4 通感算一体化研究现状

近年来，通信、感知和计算一体化引起了学界和业界的广泛关注。在 2020 年的世界 5G 大会上，中国移动首次提出了面向 6G 网络的通感算一体化框架，并合作发布了相关白皮书<sup>[48]</sup>。随后，北京邮电大学研究团队在 2021 年提出了面向 6G 网络的通感算融合理论和架构，并指出通感算融合技术是支撑 6G 网络中人-机-物智慧互联、智能体高效互通的重要手段<sup>[49]</sup>。2022 年，中国通信学会发布了《通感算一体化网络前沿报告》，报告强调通感算一体化是实现物理世界与数字世界相互作用和深度融合的重要基石<sup>[50]</sup>。同年，在中国算力大会的通感算一体化创新发展分论坛上，学术界和产业界代表以“通感算一体化创新赋能产业高质量发展”为主题纷纷发表了对通感算一体化的演进趋势、技术创新、应用实践和生态建设的看法，并成立了通感算一体化创新实验室。在 2023 年的全球 6G 技术大会上，业界专家共同探讨了通感算融合的关键技术，以及面向产业实现的问题与挑战。此外，还有一些研究综述针对特定场景下的通感算一体化展开了全面的调研，如面向扩展现实<sup>[51]</sup>、车联网<sup>[52]</sup>、无人机网络<sup>[53]</sup>、卫星通信网络<sup>[54]</sup>和智能超表面辅助的通信网络<sup>[55]</sup>等。

然而，由于通感算一体化尚在愿景和框架构设的起步阶段，现有的相关研究工作主要还是集中于通信、感知和计算的部分功能融合或业务共存设计。下面分别介绍通信和计算一体化、通信和感知一体化以及感知和计算一体化的研究现状。

### 1.4.1 通信和计算一体化

通信和计算一体化是指将通信技术和计算技术相结合，使得通信和计算功能可以在同一系统中实现，促进海量数据的传输、处理和分析效率，同时提高系统的整体性能，为实现数字化转型和智能化发展提供支持。一般地，通信和计算一体化可以分为以下三类：

- 通信计算协同：通信计算协同是指将通信和计算紧密结合，实现通信和计算的资源共享和功能协同以提高系统的性能、效率和可靠性。无线网络往往需要对大规模分布式节点产生的海量数据进行处理，但先通信后计算的传统方式不仅会给通信带宽带来极大的挑战，还会造成服务响应时间过长的的问题。因此，可以采用边传输边计算的方式，共同提高通信效率和计算效率。基于此，文献<sup>[56]</sup>研究了多址接入信道上

的计算问题，提出了一种利用同步传输产生的干扰进行可靠分布式计算的编码技术，这正是空中计算技术的起源。接着，文献[57]发现传输数据若是独立同分布的高斯随机变量，即使不采用编码技术，一个简单的模拟传输也可以实现失真最小化。由此激发了一系列关于空中计算的工作，旨在利用无线多址信道的天然叠加特性，实现通信和计算一体化。例如，文献[58]提出在物联网海量数据的场景下使用空中计算技术缓解通信压力并提高计算准确度。文献[59]为任务关键型物联网设计了一种基于多入多出（Multiple-Input Multiple-Output, MIMO）技术的空中计算框架，并提出一种低复杂度的波束成形算法，实现低时延和高精度的数据聚合。

- 通信辅助计算：通信辅助计算是指利用通信技术为计算应用和服务提供支持和优化。在通信网络中，往往分布着大量计算资源，这些资源可以通过通信节点的交互、管理和调度，为计算应用提供支持。具体而言，可以将一个计算密集型的任务分解至网络中的节点，通过分布式计算的方式实现对复杂任务的数据处理，从而提高计算效率和资源利用率，降低计算成本。例如，针对多用户 MEC 系统，文献[60]提出了一种联合部分卸载与资源分配的方案，其中计算任务可以一部分在本地计算，一部分卸载到边缘服务器上计算。文献[61]研究了多 MEC 服务器支持的计算任务卸载问题，提出了一种联合任务卸载和资源分配算法以减少任务完成的时间和能耗。文献[62]针对异构网络中多任务计算设计了一种两层卸载框架，并提出了一种联合计算卸载和用户关联优化算法，以最小化系统能耗。
- 计算辅助通信：计算辅助通信是指利用计算技术为通信系统提供支持和优化。通信系统往往面临复杂的信号处理或资源分配问题，如高维度矩阵计算或复杂优化算法，这些操作需要计算功能的辅助或者计算资源的支持。因此，可以利用先进的 AI 算法来提取数据特征，实现智能信号处理和资源分配。例如，文献[63]针对正交频分复用（Orthogonal Frequency Division Multiplexing, OFDM）系统，借助 DL 实现了在无线信道衰落和干扰情况下的信道估计和信号检测，结果表明基于 DL 的方案相比于传统方案具有更高的鲁棒性，并且其性能可与最小均方误差估计器相媲美。文献[64]提出了基于 Q 学习和基于深度强化学习的优化算法来解决多用户 MEC 系统的资源分配问题，结果表明，与传统基线算法相比，基于强化学习的优化算法能够显著降低系统成本。文献[65]研究了多用户下行多入单出（Multiple-Input Single-Output, MISO）系统的波束成形设计问题，提出了一种基于 DL 的最优波束成形通用设计框架，相比于依赖迭代的传统算法，有效降低了计算复杂度，实现了快速波束成形设计。

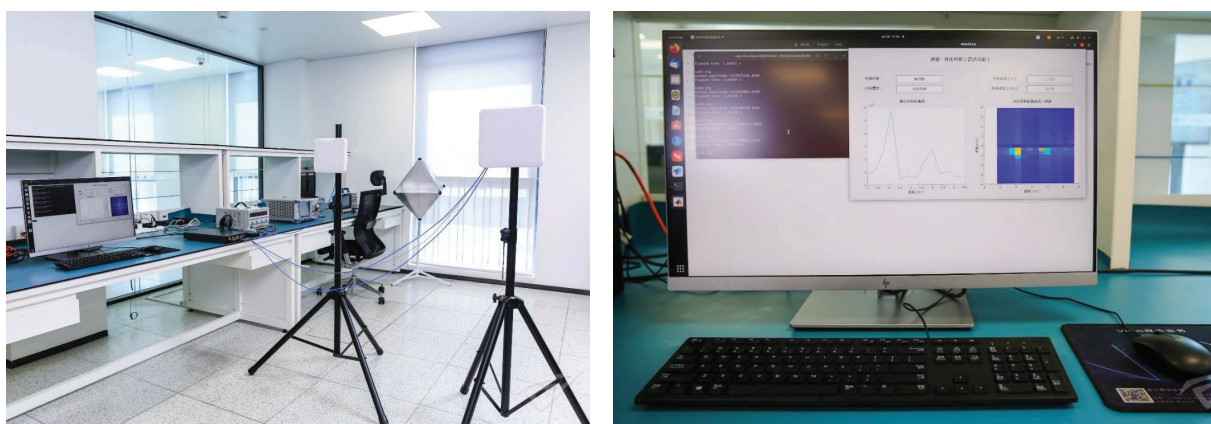
未来将迎来泛在智能的信息化社会，通信和计算一体化是必然的发展趋势。因此，通信和计算一体化也逐渐引起了业界的关注。例如，在 2021 年和 2022 年，华为作为主办方分别举办了第一届和第二届通算一体网络理论研讨会，学术界和工业界相关领域的专家们共同探讨了通信和计算一体化的应用场景、基础理论、关键技术和现存挑战。在 2023 年的全球 6G 技术大会中，亚信科技提出了算力内生网络及其关键技术，旨在通过通信和计算一体化来推动网络的智能化。

#### 1.4.2 通信和感知一体化

通信和感知一体化是实现通感算一体化的核心技术，起源于雷达和通信技术的发展和融合。早期通信和雷达系统由于应用场景和业务需求不同，一直被独立研究。特别是在 1990 年之前，雷达技术一直领先于通信技术。随着移动通信的广泛普及，雷达技术和通信技术开始相互借鉴和影响。德国在二战期间研发的相位阵列雷达“Mammut”可以看作是第一个多天线系统，它启发了 MIMO 通信技术的诞生<sup>[66]</sup>。随后，MIMO 通信技术的发展反过来也推动了 MIMO 雷达技术的发明<sup>[67]</sup>。在这个阶段，出现了一些将雷达和通信功能结合在一起的方案。例如，文献[68]首次尝试以脉冲间隔调制的方式将通信信息嵌入到雷达波形之中，实现雷达通信双功能信号。文献[69]将相反斜率的线性调频波形应用在雷达脉冲和通信脉冲上，实现在相同的天线上同时发送和接收雷达数据和通信数据。文献[70]提出了一种软件定义的雷达和通信数据融合系统，其中雷达功能和通信功能分别在不同的时隙中工作。为了提高频谱利用率，基于频谱共享的雷达和通信共存研究吸引了众多学者的关注。例如，文献[71]使用同一频谱将雷达系统和通信系统集成在同一物理空间内，但雷达和通信系统分别使用两套不同的硬件和波形。文献[72]也研究了雷达与通信系统共存于同一频段的场景，提出了一种有效的干扰消除方法来减少由频谱共享引起的干扰。进一步，文献[73]研究了 MIMO 雷达和 MIMO 蜂窝系统的共存问题，设计了相应的预编码算法来最小化雷达系统和蜂窝系统之间的干扰。

随着信息技术的快速发展，感知技术（包括无线感知与雷达感知）和通信技术都向着高频段、大规模天线和混合模拟数字结构等方向发展，以满足更高的性能需求。例如，大规模 MIMO 技术为通信系统提供了极高的空间自由度，从而大幅提升了通信系统的性能，成为了 5G 乃至 6G 的关键技术之一<sup>[74]</sup>。同时，大规模 MIMO 技术也带来了极高的空间分辨率，可以有效支持高精度和高分辨率的无线感知功能<sup>[75]</sup>。毫米波技术具有丰富的频谱资源，与大规模 MIMO 技术相结合可以获得更大的通信容量<sup>[76]</sup>。同时，基于毫米波的感知技术也被广泛应用于高精度成像和行为检测等领域<sup>[77]</sup>。此外，具有混合模拟和数字结构的

天线阵列广泛应用于大规模 MIMO 无线通信系统和雷达感知系统中，以降低硬件成本和能耗<sup>[78, 79]</sup>。不难看出，感知技术和通信技术在使用频段和硬件架构方面有着越来越多的相似性。此外，感知功能和通信功能都是基于射频信号实现的，它们在信道特性和信号处理方面有着天然的相似性。因此，自然而然地期望在同一无线资源和软硬件资源下实现通信和感知的功能协同与性能提升，以更好地支持多样化智能应用的需求。鉴于其优势特点和发展趋势，通感一体化在学术界引起了一波研究热潮。例如，文献[80]率先定义了通信和感知一体化的概念，讨论了实现通信和感知一体化的主要挑战、机遇和未来研究方向，并指出通信和感知一体化将在物联网时代发挥至关重要的作用。随后，文献[81]对通信和感知一体化进行了全面的调研，特别是从理论和技术的角度对一体化波形设计和接收信号处理的研究现状和关键挑战进行了详细的分析和总结，并从协同增益和融合增益两方面给出了通感一体化的应用场景和实际用例。此外，文献[82]从基本界限的角度对通信和感知一体化进行了全面的研究，总结了通感一体化系统的主要性能指标和性能边界，并讨论了待解决的问题和未来研究方向，为一体化系统设计提供了有益的见解和指导。



(a) 基于通感一体化目标测速原型样机

(b) 室内目标测速结果

图 1.11 VIVO 通信研究院展示的通感一体化原型样机<sup>[83]</sup>

作为 6G 潜在关键技术之一，通感一体化也引起了业界的广泛关注和研究，全球许多公司和机构都已开展相关的项目和试验。特别是，我国在通感一体化技术的研究上处于世界领先的水平。例如，华为率先在 2021 年完成了全球首个面向 5G-Advanced 的通信和感知一体化技术验证，其中双功能基站探测距离超过 500 米，且对车辆和行人的感知准确度高达 100%。如图 1.11 所示，VIVO 通信研究院搭建了基于通感一体化的目标测速原型样机，其中用于无线感知的资源开销仅为 7%。当发射功率和天线数增加时，该样机还可以支持室外场景的目标感知。此外，中国移动研究院构建了通感一体化测试解决方案，并

联合 R&S 公司使用其研发的动态目标模拟器 AREG800A 实现了高精度定位功能的模拟验证。该测试解决方案具有高度的灵活性，通过配置不同的变频模块，还可以实现基于毫米波和太赫兹频段的通感一体化验证。与此同时，一些国际组织和联盟也正在推动通感一体化的标准化进程。例如，ETSI 将通感一体化列为信息通信技术的技术趋势之一；电气与电子工程师协会（Institute of Electrical and Electronics Engineers, IEEE）成立了通信感知一体化新兴技术倡议委员会；第三代合作伙伴计划（The 3rd Generation Partnership Project, 3GPP）开始在 R17 及以后版本中探索通感一体化技术，并将其定义为 5G-Advanced 以及 6G 时代的潜在关键技术；国际电信联盟（International Telecommunication Union, ITU）在国际移动通信（International Mobile Telecommunications, IMT）-2020 推进组和 IMT-2030 推进组中均对通感一体化展开了详细研究，并发布了《通信感知一体化技术研究报告》<sup>[83]</sup>。

#### 1.4.3 感知和计算一体化

IMT-2030 推进组发布的《6G 总体愿景及潜在关键技术》白皮书指出感知功能和计算功能都是未来无线网络不可或缺的重要成分<sup>[6]</sup>。感知和计算一体化是指利用感知技术和计算技术相互协作，为智能系统提供更加准确和高效的数据处理和决策能力。在感算一体化系统中，感知功能可以理解为系统的感官，为系统获取丰富的环境数据，而计算则是系统的大脑，通过对系统中获取的感知数据进行分析和处理，从中提取有用的信息和特征，以实现更精准和更智能的决策和控制，从而为各个领域的应用和服务提供更加优质和高效的支持。感算一体化的应用范围非常广泛，包括智能制造、智慧交通和智能医疗等领域。

- 在智能制造领域，可以通过传感器、摄像头等感知设备，为机器提供实时的环境信息。与此同时，通过对感知数据进行智能分析和处理，可以实现更加精准和智能高效的生产和制造。例如，文献[84]研究了基于传感器网络的智慧工厂系统，提出了一种基于强化学习的动态生产调度算法。结果表明，该智能算法不仅提升了生产调度问题的多目标性能，而且能有效地应对制造系统的突发事件。文献[85]设计了一种基于智能机器人的自动化生产系统架构，并提出了基于认知制造和边缘计算的解决方案。结果表明，与传统生产系统相比，所提的自动化生产架构不仅有效降低了系统指令数量，还显著提高了产品组装效率和生产效率。
- 在智慧交通领域，可以通过激光雷达、视觉传感器等感知设备，为驾驶车辆提供实时的交通状况和道路情况。与此同时，通过汇集环境数据和车辆信息进行智能分析、处理、决策和反馈，可以实现更安全和更智能的交通管理和控制。例如，文献[86]

研究了车联网中感知辅助的交通信息采集和传播问题，建立了一个用于车联网感知网络的加权无向图模型，并提出了一种最优的智能车辆信息交互和节点选择方案。文献[87]研究了车联网的安全问题，通过 CNN 提取并分析路侧感知单元的链路负载行为特征，设计了一个数据驱动的路侧感知单元入侵检测系统，以应对各种导致交通流不规则波动的恶意入侵与攻击，从而确保用户信息和基础设施的安全。文献[88]针对 MEC 使能的车联网中道路交通状况和车辆机动性的不可预测性，提出了一种具有不稳定感知能力的分布式在线计算卸载算法，以提高服务效率和服务质量。结果表明，与现有算法相比，所提算法可将服务匹配率提高约 25% 并将车辆功耗降低约 54%。

- 在智能医疗领域，可以通过生物传感器、医学影像等感知技术，为医疗设备和医生提供实时的生理数据和影像数据。与此同时，通过对医学感知数据进行实时的分析和有效的处理，提取病人疾病的特征和趋势，可以为病人提供个性化的智能诊断和治疗服务。例如，文献[89]指出传感器网络、物联网和机器学习技术的结合可以辅助医生更准确地识别疾病，有效减轻医生的工作，给患者、医生、医院以及医疗设备制造商带来了新的希望和机遇。文献[90]提出了一种新颖的基于物联网感知的智能医院系统架构，用于自动监测和跟踪医院和护理机构内的患者、人员和生物医学设备。文献[91]设计了一个基于 MEC 的智能医疗架构来实现网络内和上下文感知数据的处理。基于这个架构，不仅可以实现多模态数据的低失真压缩，还可以完成基于特征提取的事件检测，以确保在紧急情况下的高可靠和低时延响应。

## 1.5 研究挑战

由以上研究现状可知，通信和计算一体化、通信和感知一体化以及感知和计算一体化都在各自领域中取得了不少进展，这也进一步证明了通信、感知和计算一体化的可行性。尽管通信、感知和计算一体化已经引起了学术界和工业界的极大兴趣，但其研究尚处于理论愿景和框架设想的初步探索阶段。为了更好地支持边缘智能网络的各种新兴智能业务和应用，亟需对通信、感知和计算一体化开展进一步的研究，实现更高效、更可靠、更智能的数据获取、传输和处理，以满足不同应用场景的差异化性能需求。面向边缘智能网络的通信、感知和计算一体化设计，关键是如何高效地利用有限的无线资源和软硬件资源实现通信、感知和计算功能之间的相互作用和互利互惠，从而达到基于业务特点和需求的自适应调度和复用。然而，在边缘智能网络中实现通信、感知和计算一体化并不是一个简单的



任务，其存在如下一些挑战性的问题：

1. 海量数据处理：在边缘智能网络中，大量分布式节点采集到的数据规模极大，且具有多样的结构和语义，需要对获得的数据进行实时处理。然而，边缘节点和网元的计算和存储能力存在差异化。例如，部分能力受限的终端无法支持本地实时数据处理，部分资源有限的边缘服务器也无法同时处理海量数据。为了有效提高通信、感知和计算一体化的整体性能，可以将原本“分布获取、集中融合”的模式转变为“分布获取、协同融合”的模式，以减少传输的数据量，在缓解通信压力的同时提高信息感知的效率。因此，关键是如何设计高效的协同数据传输和处理方案来提高信息感知效率。
2. 复杂干扰协调：在边缘智能网络中，联合信号处理的基本思想是在同一接收机上同时完成通信信息解码、感知目标估计和数据信号计算。然而，接收机收到的是通信信息、感知回波和计算数据的叠加信号，引起的自干扰和互干扰会严重限制通信、感知和计算的性能。此外，边缘智能网络面临复杂的无线传输环境、多样化的感知场景和海量待处理的数据，给联合信号处理带来了极大的挑战。在这种情况下，传统的信号处理方法无法直接应用到通信、感知和计算一体化场景中。因此，关键是如何有效协调复杂干扰以及引进更先进的信号处理方法来提升通信、感知和计算的整体性能。
3. 联合资源管理：边缘智能网络的不同应用场景具有差异化的需求。因此，实现各类新型智能业务需要通信、感知和计算等多维资源支撑。然而，多维资源与多样化业务的匹配往往存在竞争、协作、制约和耦合等复杂关系。换句话说，无法根据每一类业务自身的特点和需求独立地进行资源分配，这就为资源分配增加了极大的难度。为了有效实现多维资源管理，需要根据网络状态和业务特点进行联合资源管理和自适应调度，以满足差异化的业务需求。因此，关键是如何高效地根据业务需求对通信、感知和计算任务进行建模、分解、分配和优化，从而实现通信、感知和计算多维资源的自适应分配和调度。
4. 协同融合架构：在边缘智能网络中，通信、感知和计算三者是密不可分的整体，它们之间的耦合关系将极大影响边缘智能网络的性能。首先，通信、感知和计算节点的任务相互关联、分布协调。再者，通信、感知和计算的过程相互触发、多重交互，这些都会影响任务的执行时间和完成效率。为了适配通信-感知-计算协同架构，需要探究通信-感知-计算的耦合机理，尤其是三者的性能边界及性能折中。因此，关键是

如何根据边缘智能网络的特点设计通信-感知-计算的协同架构，从而达到通信、感知和计算三者功能的深度融合和互惠增强。

## 1.6 研究内容和章节安排

基于以上研究现状和挑战，本文对面向边缘智能网络的通信、感知和计算一体化展开了深入的研究。针对以上四种不同的研究挑战，分别提出了四种典型的一体化创新架构，即面向感知的通信和计算一体化、计算辅助的通信和感知一体化、通信支撑的感知和计算一体化以及功能复用的通信、感知和计算一体化。对于这四种一体化创新架构，分别根据其业务特点从设计、分析和优化的角度给出了可行的解决方案，解决了资源受限条件下实现边缘智能的关键问题和挑战。

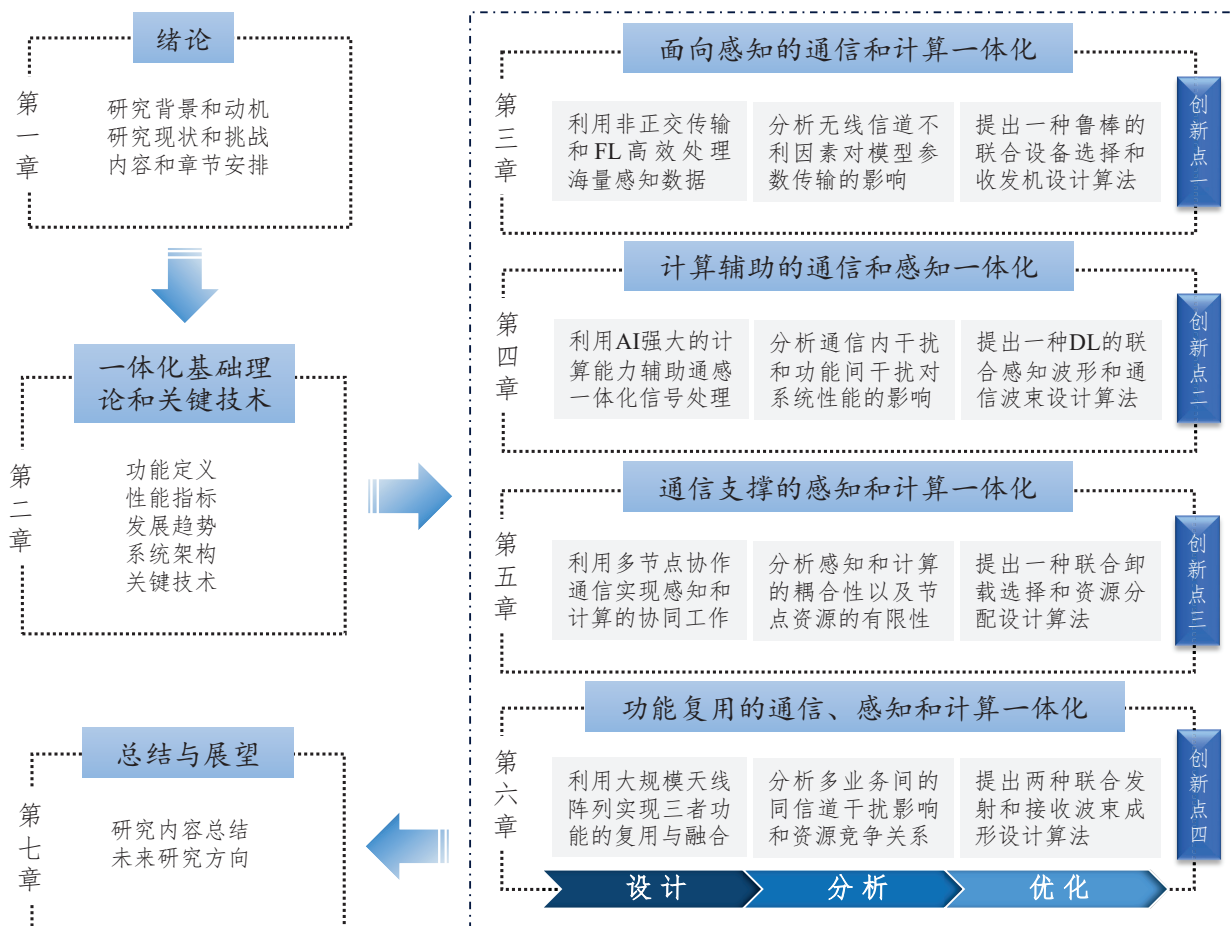


图 1.12 全文内容和章节安排

全文共分为七章，其中第三章到第六章为四种典型的一体化架构及其解决方案，分别对应了本文的四个创新点，均由作者发表或投递的论文整理而成。如图 1.12 所示，本文的

章节内容安排如下：

第一章为绪论，首先介绍了研究背景，接着回顾了边缘智能的由来并介绍了边缘智能网络的特点和需求，然后引出了面向边缘智能网络的通信、感知和计算一体化的研究动机、研究现状和研究挑战，最后概述了全文内容和章节安排。

第二章详细介绍了通信、感知和计算一体化的基础理论和关键技术，给出了通信、感知和计算这三个基本要素在边缘智能网络中的具体定义和性能评价指标，阐述了一体化的发展趋势并提出了一种通用的一体化系统架构，以及分析了边缘智能网络中实现通信、感知和计算一体化的关键技术。

第三章设计了一种面向感知的通信和计算一体化架构，利用非正交传输和边缘 FL 对海量感知数据进行高效处理，将“先感知再通信最后计算”的模式转变为“基于感知数据的模型参数交互”的模式。通过分析信道衰落、干扰和噪声以及信道估计误差等不利因素对 FL 模型参数传递的影响，提出了一种基于空中计算的联合设备选择和收发机设计的鲁棒算法来提高通信可靠性和学习准确性。仿真结果证明了所提算法的有效性和鲁棒性。

第四章设计了一种计算辅助的通信和感知一体化架构，利用多功能基站在同一频谱和相同硬件上实现了通信和感知的协同，有效提高了频谱和硬件的利用效率。通过分析资源共享导致的通信内干扰与通信和感知功能间互干扰对系统性能的影响，借助 AI 强大的计算能力，提出了一种低复杂的基于 DL 的联合感知发射波形和通信接收波束设计算法，提高了通信和感知一体化的整体性能。仿真结果证明了所提 DL 算法的有效性和鲁棒性。

第五章设计了一种通信支撑的感知和计算一体化架构，利用多节点协作通信实现感知信息和计算数据的高效传递和汇聚，显著提升了目标感知精度和数据计算效率。通过分析感知任务和计算任务的耦合关系，在边缘节点资源有限的情况下，提出了一种以最小化加权总能耗为目标的联合卸载选择和资源分配算法，增强了感知和计算一体化的整体性能。仿真结果验证了所提算法的可行性和有效性。

第六章设计了一种功能复用的通信、感知和计算一体化架构，利用配置大规模天线阵列的基站所具有的空间自由度，在同一无线资源块上同时实现了多目标感知、多维计算和多流通信。通过分析功能复用情况下多业务的同信道干扰和资源竞争对系统性能的影响，提出了两种联合发射和接收波束设计算法，分别有效提升了系统加权总性能和降低了系统发射总功率。仿真结果验证了所提的两种算法的可行性和有效性。

第七章总结了全文的研究内容和贡献，并展望了未来重要的研究方向。



## 2 通感算一体化基础理论和关键技术

### 2.1 引言

随着信息技术的不断演进、物联网的快速发展和数据流量的爆发式增长，越来越多的应用场景需要在传统通信基础上，融合无线感知和智能计算等高级功能，以满足更先进、更智能的实时业务需求，如数字孪生、全息通信和感官互联等。对于过去传统的无线网络而言，通信、感知和计算功能是以单一目的独立存在的，感知系统只负责采集数据，通信系统只负责传输数据，计算系统只负责处理数据。显然，这种分离化设计不仅会造成稀缺的无线频谱与昂贵的硬件资源的严重浪费，还难以满足未来无线网络中新兴智能应用的超高性能要求。为此，迫切需要提供一种以任务需求为导向的无线设计新范式，即通信、感知和计算一体化。通信、感知和计算一体化旨在通过资源共享、多点协作和智能交互，实现通信、感知和计算三个原本独立的功能在同一频谱和硬件下的协同工作，显著提升资源利用率和系统性能。鉴于“通信”、“感知”和“计算”都包含了不同的技术范畴，通信、感知和计算一体化具有非常庞大的技术理念与多模式的系统范式。尽管学术界和产业界已经给出了很多关于通信、感知和计算一体化的目标愿景和架构设想，但是目前还没有一个统一而准确的理论来界定其具体的内涵与外延。因此，本章基于边缘智能网络的特点，提出通信、感知和计算一体化的基础理论并给出潜在的关键技术，为面向边缘智能网络的通信、感知和计算一体化架构设计提供重要的理论基础。

内容概述：本章详细介绍了面向边缘智能网络的通信、感知和计算一体化的基础理论和关键技术，为接下来第三章到第六章所提的四个通信、感知和计算一体化典型架构设计奠定了重要的基础。具体结构安排如下：2.2节详细介绍了一体化的基础理论，包括通信、感知和计算这三个基本要素的定义、功能和性能评价指标，以及一体化程度和系统架构；2.3节给出了实现一体化的关键技术，包括无线感知技术、一体化波形技术、大规模MIMO和波束成形技术、干扰管理技术、空中计算技术以及边缘联邦学习技术；2.4节总结了本章内容。

## 2.2 基础理论

### 2.2.1 基本要素

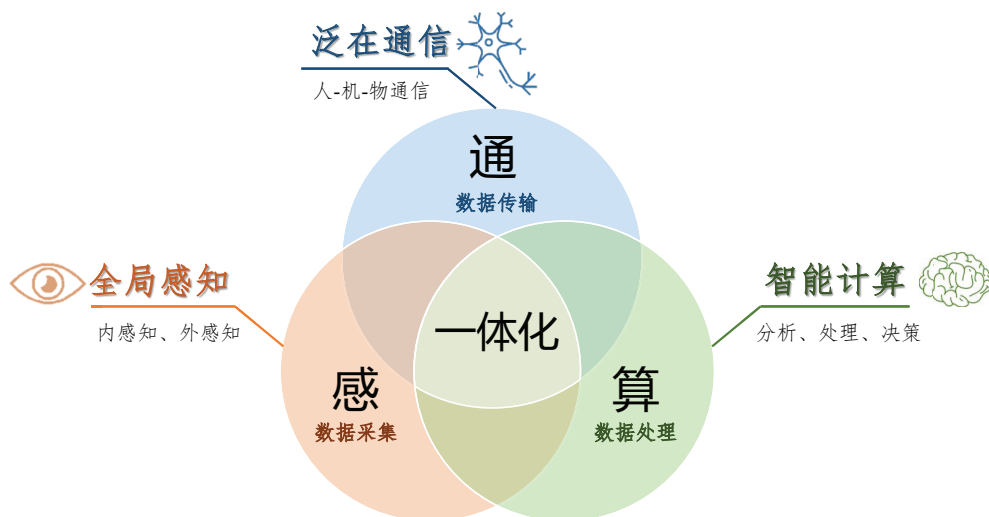


图 2.1 通信、感知和计算功能定义

如图 2.1 所示，在无线通信系统中，通信、感知和计算是最基础也是最重要的功能，它们通过获取、传输和处理信息等方式，为广泛的智能业务提供相应的服务。将上述三个功能集成在一起，能够有效提升网络性能，提高频谱效率，并为新的技术与应用提供支撑。具体地，边缘智能网络中通信、感知和计算的功能可以定义为：

- 通信：通信是指设备通过无线网络进行数据传递和交互的能力，可以看作是网络的神经中枢。这里的“通信”指的是蜂窝移动通信。在过去几十年里，蜂窝移动通信从 1G 发展到 5G，传输速率、系统容量等性能都得到了极大的提升。同时，关注对象也发生了改变。与过去主要针对以人为中心的移动互联网不同，蜂窝物联网进一步拓展了移动通信业务范围，将人与人之间的通信扩展至物与物、人与物的智能互联，推动了蜂窝移动通信向工业、农业、医疗、教育、交通和金融等各行各业的渗透。
- 感知：感知是指利用传感设备、数据接口和数据分析能力从感知对象采集信息的能力，可以看作是网络用于探索物理世界的感官。这里的“感知”指的是利用无线网络进行感知的行为。一般地，感知可以分为内生感知（业务状态、网络属性、终端信息等）和外部感知（目标物体、区域、事件或环境等）。具体而言，业务状态包括业务时延、业务传输速率、业务吞吐量等；网络属性包括频谱效率、网络容量、网

络算力、网络覆盖等；终端信息包括终端功耗、位置信息、接入数量等。外部感知包括对单个目标的位置、距离和速度等状态的感知，也包括对目标区域、事件或者环境进行检测、识别和成像等。

- 计算：计算是指设备和服务器使用算法和 AI 模型处理和分析数据的能力，可以看作是网络的大脑。这里的“计算”指的是边缘节点和服务器执行的数据分析、处理、决策和推演行为。在网络边缘侧处理数据，能够大大缓解网络带宽以及数据中心功耗的压力，降低网络数据泄露的风险，保护用户的数据安全与隐私。此外，由于无需经由网络获得云计算中心的响应，能够显著降低系统时延，增强服务的响应能力。同时，以 DL 为代表的 AI 技术，赋能每个网络边缘节点，使其拥有良好的计算和决策能力，以满足实时业务、敏捷连接、智能应用等服务需求。

### 2.2.2 性能指标

由于边缘智能网络的应用场景需求各异，比如智能家居有连接密度大、感知精度高的需求，但覆盖范围小；无人机网络有低时延、高可靠性的需求，且覆盖范围大；智慧交通有高精度感知、高速率传输和低时延计算的需求。因此，需要从通信、感知和计算三方面考虑性能要求。下面分别列举通信、感知和计算功能的一些公认性能评价指标。

- 通信性能指标：通信性能指标通常用于评价传输信息的速率和质量，主要从有效性和可靠性两方面来评价。有效性是指给定带宽内传输的信息量，即传输信息的速率快慢问题，常用的性能指标有：传输速率，指每秒通过无线信道传输的数据量，传输速率越高，性能越好；传输时延，指信号发射和接收之间的时间延迟，时延越低，性能越好。而可靠性是指接收信息的准确程度，即传输信息的质量好坏问题，常见的性能指标有：误码率，指在信号传输中出现误码的占比，误码率越低，性能越好；信（干）噪比，指接收信号相对于噪声（及干扰）强度的度量，信（干）噪比越高，性能越好；其他常用的性能指标还有：覆盖范围，是指无线信号可以可靠接收的区域，覆盖面积越大，性能越好；接入容量，是指在给定时间内可支持的最大终端设备数量，接入容量越大，性能越好。
- 感知性能指标：感知性能指标通常用于评价获取信息/状态的准确性。边缘智能网络中感知性能的衡量主要考虑网络外部感知的准确性，即周围环境和目标物体的感知。对于周围环境的感知，常用的指标有：空间分辨率，指扫描感知区域能分辨的最小单元，空间分辨率越小，性能越好；感知范围，指感知周围环境所覆盖的范围，

表 2.1 通信、感知和计算性能评价常用指标

| 评价类   | 分类   | 评价标准   |
|-------|------|--------|
| 通信性能  | 可靠性  | 误码率    |
|       |      | 信(干)噪比 |
|       | 有效性  | 传输速率   |
|       |      | 传输时延   |
|       | 其他   | 覆盖范围   |
| 接入容量  |      |        |
| 感知性能  | 周围环境 | 空间分辨率  |
|       |      | 感知范围   |
|       |      | 能量消耗   |
|       | 目标检测 | 虚警概率   |
|       |      | 漏警概率   |
|       | 目标估计 | 均方误差   |
|       |      | 克拉美罗界  |
| 感知互信息 |      |        |
| 感知速率  |      |        |
| 计算性能  | 广义计算 | 计算能耗   |
|       |      | 计算误差   |
|       |      | 计算时延   |
|       | 学习推演 | 平均绝对误差 |
|       |      | 准确度    |
|       |      | 敏感度    |

感知范围越大，性能越好；能量消耗，指单位感知周期内感知周围环境所消耗的能量，能耗越低，性能越好。对于目标感知，涉及到检测目标是否存在，常用的指标有：虚警概率，指实际目标不存在但检测出存在的概率，虚警概率越低，性能越好；漏警概率，指实际目标存在但检测认为不存在的概率，漏警概率越低，性能越好。此外，还涉及到估计目标位置、速度和运动轨迹等状态，常用的指标有：均方误差，指实际参数与估计参数之间的均方误差，均方误差越小，性能越好；克拉美罗界，指无偏估计方差的下界，克拉美罗界越小，性能越好；感知互信息，指感知信号与接收信号之间的熵，感知互信息越大，性能越好；感知速率，指单位时间内感知的信息量，感知速率越大，性能越好。



- **计算性能指标**：计算性能指标通常用于评价处理信息的效率。边缘智能网络中计算性能的衡量主要考虑广义的计算和基于 AI 的学习推演。对于广义的计算，常用的指标有：计算时延，指执行计算任务消耗的时间延迟，计算时延越低，性能越好；计算能耗，指执行计算任务消耗的能量，计算能耗越低，性能越好；计算误差，指实际想要的计算结果与数据通过无线信道后的计算结果之间的均方误差，计算误差越低，性能越好。对于基于 AI 的学习推演，回归任务常用的指标有：平均绝对误差，指预测样本和真实标签之间的平均绝对误差，平均绝对误差越低，性能越好；分类任务常用的指标有：准确度，指预测正确的样本占总样本的比例，准确度越高，性能越好；敏感度，指预测正例的样本数占有所有正例的样本数的比例，敏感度越高，性能越好。

为了便于查看，上述常用的性能指标归纳在表 2.1 中。需要指出的是，通信、感知和计算一体化的性能指标目前尚没有统一的标准。一般地，可以对三个功能的指标进行综合。例如，分别选定三者相关的性能指标，再将它们归一化后的加权和作为通信、感知和计算一体化的性能指标。

### 2.2.3 一体化程度

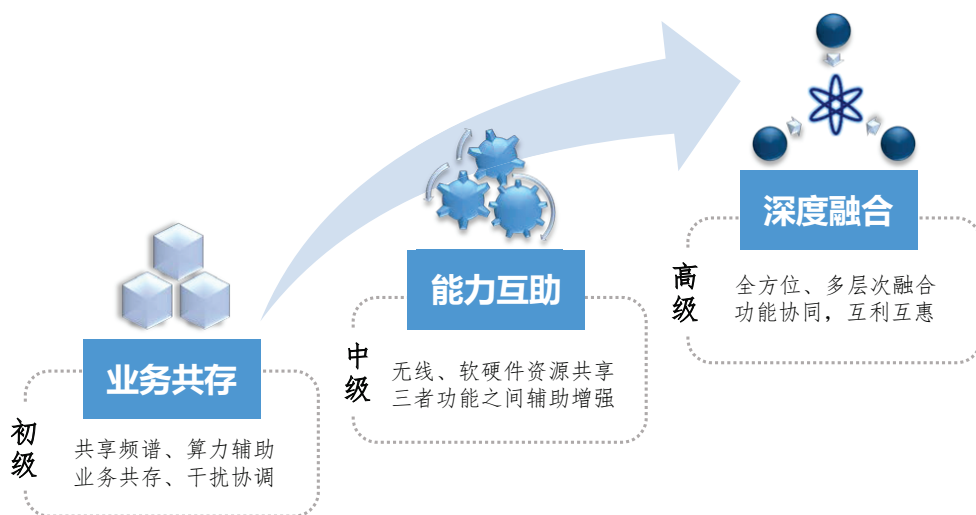


图 2.2 通信、感知和计算一体化发展趋势

如图 2.2 和表 2.2 所示，通信、感知和计算一体化具有分阶段、分层次发展的特点，其发展趋势从初级一体化（业务共存）、中级一体化（能力互助）到高级一体化（深度融合）逐步演进和增强。

表 2.2 通信、感知和计算一体化程度区别

| 一体化程度 | 频谱 | 硬件 | 波形 | 算力 | 存储 | 功能 |
|-------|----|----|----|----|----|----|
| 通感算分离 | 独立 | 独立 | 独立 | 独立 | 独立 | 独立 |
| 初级一体化 | 共享 | 独立 | 传统 | 辅助 | 独立 | 共存 |
| 中级一体化 | 共享 | 共享 | 融合 | 协同 | 共享 | 互助 |
| 高级一体化 | 共享 | 新型 | 新型 | 一体 | 一体 | 协同 |

- 初级一体化：初级一体化是通信、感知和计算一体化技术的初始阶段，通信和感知共享频谱，计算辅助，三者业务共存。在初级一体化阶段，独立的通信系统和感知系统通过共享频谱物理集成在同一空间内。与此同时，诸如 MEC、AI 等计算技术以“外挂式”的形式进行算力辅助，分别为通信系统和感知系统提供数据处理和智能功能。因此，该阶段的设计重点是如何进行高效的干扰协调以提升通信系统和感知系统的性能，以及如何进行有效的资源管理和设计先进的算法来提升计算效率和学习效率。
- 中级一体化：中级一体化是通信、感知和计算一体化技术的发展阶段，通信、感知和计算共享无线资源和软硬件资源，三者能力互助增强。在中级一体化阶段，通信和感知之间的资源共享和协同能力得到了进一步提升，例如，感知可以利用通信网络来实现更大范围、更高精度的环境探测和目标定位，通信也可以借助感知提供的环境信息来进行更可靠的无线传输和资源调度。同时，计算也开始逐步与通信和感知进行深度融合，共同提供更高效、更智能的数据处理和决策能力。因此，该阶段的设计重点是一体化波形技术、一体化波束成形技术和联合数据处理技术等。
- 高级一体化：高级一体化是通信、感知和计算一体化技术的成熟阶段，通信、感知和计算三者相互渗透，实现全方位、多层次的深度融合。在高级一体化阶段，通信、感知和计算系统将变得更加智能和自适应，可以根据不同场景的需求进行自优化和自调整。在系统内部通信、感知和计算功能相互支持、互利互惠，共同促进系统的良性循环和改善。在系统外部可实现按需定制的通信、感知和计算服务，并具有服务质量保障。因此，该阶段的设计重点是多源信息处理技术、多维资源管理技术、新型波形技术、新型硬件架构、协同功能设计和业务智能编排技术等。

## 2.2.4 系统架构

通信、感知和计算一体化系统架构是一种对无线通信、无线感知与智能计算功能、服务和应用进行支持的架构，可以充分利用并调动通信、感知和计算的无线资源及软硬件资源，实现数据、功能、应用和业务之间的有效联动。如图 2.3 所示，通信、感知和计算一体化系统架构可以划分成三层，即资源层、能力层与应用层。

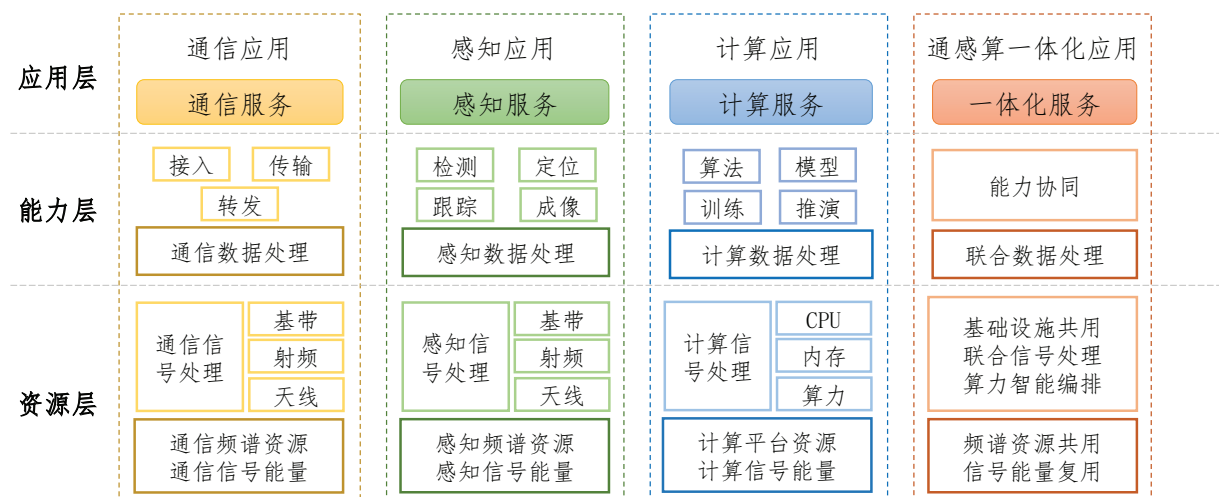


图 2.3 通信、感知和计算一体化系统架构

- 资源层：资源层为系统所需的各种能力提供必要的基础资源，主要包含无线资源、软件资源和硬件资源，具体可以分为通信频谱资源、感知频谱资源以及共用频谱资源；通信射频资源、感知射频资源以及共用射频资源；通信硬件基础设施、感知硬件基础设施、计算硬件平台以及一体化硬件设施平台、分布式的算力资源、内存资源以及智能统一编排算力资源。
- 能力层：能力层是实现服务和应用支撑的数字和物理能力，包含通信功能、感知功能、计算功能以及能力协同，具体可以分为通信数据处理、感知数据处理、计算数据处理和联合数据处理，包括接入、传输和转发的通信功能，包括检测、定位、跟踪和成像的感知功能，包括算法、模型、训练和推演的计算功能，以及包括通信、感知和计算三者的部分或者全部协同功能。
- 应用层：应用层是为用户提供服务与应用的接口，具体可以分为基于高可靠、低时延、大带宽传输的通信服务，基于高精度目标识别和定位、高分辨率环境检测和成像的感知服务，基于高效率数据分析、智能决策和推演的计算服务，以及同时包含

三者部分或全部的一体化服务，可进一步支持人机交互、虚实互动、深度沉浸的通信、感知和计算一体化智能应用。

因此，通信、感知和计算一体化赋能的边缘智能网络是泛在通信网络、无线感知网络和分布式计算网络的有机融合体，网络中的各个网元和节点都有集成通信、感知和计算部分功能或全部功能于一体的能力，这促使网络拥有内生感知和内生智能能力。基于全局感知能力，网络能够对电磁环境、物理环境和数字环境进行全方位的多维感知。同时，结合超强的泛在通信能力和分布式智能计算能力，网络能够最大化共享共用通信、感知和计算相关资源和有效信息，对内实现通信、感知和计算功能相互支持、互利互惠，共同促进网络内部性能的良好循环和提升，对外提供按需定制的高质量通信、感知和计算服务。

## 2.3 关键技术

为了实现通信、感知和计算一体化，并提供高性能的通信、感知和计算服务，边缘智能网络需要强大的技术支持。下面对边缘智能网络的一些关键技术进行介绍。

### 2.3.1 无线感知技术

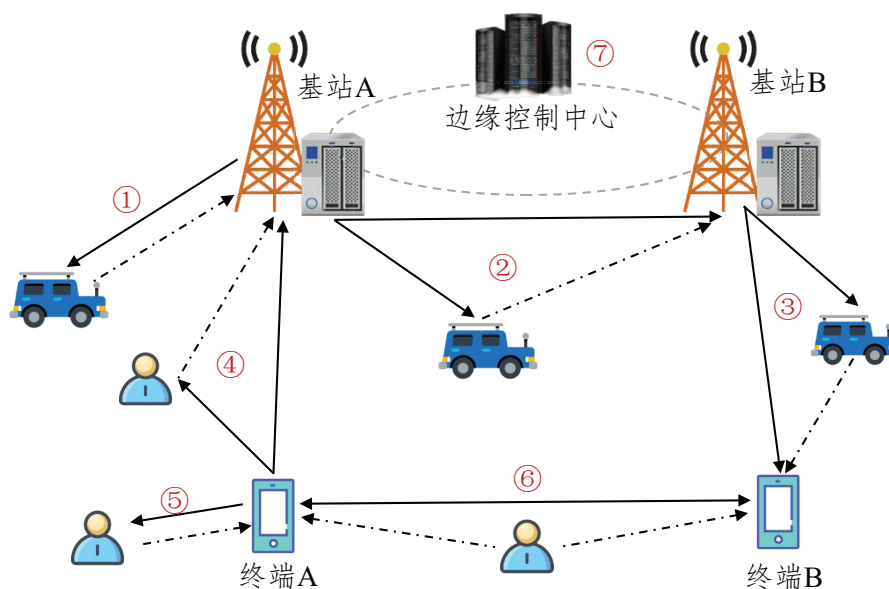


图 2.4 边缘智能网络中无线感知的不同方式

无线感知是通过无线信号获取目标信息的行为<sup>[92]</sup>。由于边缘智能网络中的多功能终端和基站具备通信、感知和计算能力，因此存在多种无线感知方式。图 2.4 给出了 7 种不同

的感知链路：(1) 基站自发自收；(2) 基站 A 发，基站 B 收；(3) 基站发，终端收；(4) 终端发，基站收；(5) 终端自发自收；(6) 终端 A 发，终端 B 收；(7) 多点协作，包括多点发送，多点接收，联合处理感知信息。在实际场景中，目标可以是独立的运动或静止的个体，也可以是一个范围内的空间状态。基于感知任务需求，可以选择合适的感知链路，其中每种感知链路的发射端和接收端可以有多个，也可以同时存在多种感知链路。

通过无线信号感知目标的具体方式为：发射端在  $t$  时刻首先向目标发射无线感知信号  $x(A(t), \theta(t), \nu(t))$ ，该信号包含系列确定参数，如信号幅度  $A$ 、相位  $\theta$  和频率  $\nu$  等。相应地，接收端收到的经目标反射后的回波信号为：

$$y(t) = \alpha(\Delta A(t), \Delta\theta(t), \Delta\nu(t)) x(A(t), \theta(t), \nu(t)) + n(t), t \in [0, T], \quad (2.1)$$

其中  $T$  是信号周期， $\alpha(t)$  是目标响应函数， $\Delta A(t), \Delta\theta(t), \Delta\nu(t)$  分别指信号经目标反射后幅度、相位和频率发生的变化。由于发射信号  $x(A(t), \theta(t), \nu(t))$  在接收端已知，通过信号处理算法恢复出未知响应函数  $\alpha(t)$  中与目标特征相关的参数，即可获得感兴趣的目标状态。

需要注意的是，由于边缘智能网络中终端设备连接数量较大，不可避免会发生感知数据的重叠以及产生通信和感知之间的相互干扰。再者，由于边缘终端的异构性，获取到的数据通常具有不同的功能、结构和语义。因此，如何有效避免重复感知、减缓通信和感知之间的干扰以及高效处理多维感知数据等问题是无线感知技术面临的重要挑战。

### 2.3.2 一体化波形技术

通信和感知一体化波形技术是通信、感知和计算一体化设计的重要基础。针对不同应用场景下的感知和通信功能需求，目前可分为基于传统波形共用的一体化波形和新型一体化波形设计两大类。

基于传统波形共用的一体化波形包括以通信为中心的一体化波形和以感知为中心的一体化波形。以通信为中心的一体化波形设计指在现有的通信波形及帧结构上实现感知功能，如用于 5G 通信的 OFDM 波形<sup>[93]</sup>或基于正交时频空（Orthogonal Time Frequency Space, OTFS）调制的波形<sup>[94]</sup>。原则上，任何通信波形都可以用于主动感知，因为发射机对通信波形是已知的。然而，由于通信数据的随机性以及通信帧结构并非针对感知进行优化，此类方法可能会大大降低感知性能。以感知为中心的一体化波形期望将通信功能融合到现有的感知波形中。然而，由于纯感知波形不包含数据信息，因此不能直接用于通信。基于

表 2.3 基于传统波形共用的一体化波形案例

| 波形分类   | 波形                 | 优点      | 缺点       |
|--------|--------------------|---------|----------|
| 以通信为中心 | 基于扩频序列的一体化波形       | 包络恒定    | 频谱利用率低   |
|        | 基于 OFDM 的一体化波形     | 频谱利用率高  | 多普勒影响大   |
|        | 基于 OTFS 的一体化波形     | 适用高移动场景 | 实现复杂高    |
| 以感知为中心 | 基于幅度调制的 FMCW 波形    | 调制/解调容易 | 包络起伏大    |
|        | 基于频率调制的 LFM 波形     | 频谱扩展小   | 调制/解调复杂  |
|        | 基于相位调制的 LFM 波形     | 包络起伏小   | 感知受相位影响大 |
|        | 基于 OFDM 的 Chirp 波形 | 通信速率高   | 信号处理复杂   |

此，需要将数据信息嵌入到感知波形中，并且不降低感知的性能。例如，将通信信息调制到频率调制连续波（Frequency-Modulated Continuous Wave, FMCM）信号的幅度上<sup>[95]</sup>或将通信信号的调制载波替换成线性频率调制（Linear Frequency Modulation, LFM）信号<sup>[96]</sup>。此类方法虽然能保证感知性能，但由于大部分采用的是脉冲间调制的方法，其应用几乎仅限于低速率通信的场景。基于这两种思想设计的经典波形方案已列在表 2.3 中。

新型一体化波形设计不依赖于现有波形，目前尚处于研究阶段。通过联合优化通信性能和感知性能来设计一体化波形的方法是可行的。例如在发射功率限制下，以最大化通信互信息和感知互信息的加权和或者最小化通信和感知之间的互干扰为目标设计波形。与基于传统波形共用的一体化波形设计相比，通感联合的新型一体化波形设计更具可塑性和灵活性，但其复杂度也更高。

### 2.3.3 大规模 MIMO 和波束成形技术

大规模 MIMO 技术是 MIMO 技术的自然延伸，通过将原有基站天线数（如 4 或 8）提高一个数量级（如 64、128 或 256），充分挖掘空间资源以显著提高系统性能，是当前 5G 乃至未来移动通信系统的关键技术之一<sup>[74]</sup>。得益于大规模的天线阵列，大规模 MIMO 技术能够提供更大的自由度，进一步提高分集和复用增益，从而显著增加通信网络容量、改善通信传输速率。与此同时，由于大规模天线阵列极大地拓展了阵列孔径，能够提供极高的空间分辨率以及带来更强大的抗干扰能力，大幅提高感知的精度。

大规模 MIMO 通常利用波束成形技术来实现通信和感知。波束成形技术是指通过调整每个天线阵元上的信号并进行加权求和，使天线波束指向某个特定方向的技术<sup>[97]</sup>。因此，可以利用大规模阵列天线产生的高增益窄波束来实现通信速率和感知精度的提升。如图 2.5 所示，对于通信而言，大规模天线阵列产生的通信波束能够自适应匹配信道特性以

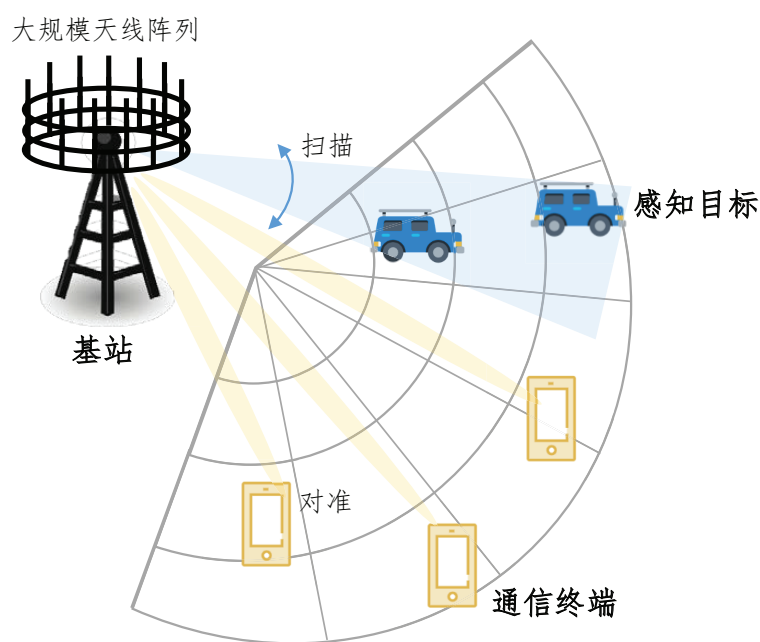


图 2.5 大规模MIMO下通信和感知一体化波束成形示意图

精确对准通信接收机，有效抑制干扰并提升通信质量。对于感知而言，大规模天线阵列产生的感知扫描波束能够对周围环境进行感知，实现更大范围的目标检测并有望提供厘米级的高精度定位服务。在边缘智能网络中，基于全局感知能力，获取电磁环境、网络内部状态和外部物理环境的信息，结合智能计算能力，进行按需定制的自适应波束成形，可以有效满足通信和感知的性能要求。

#### 2.3.4 干扰管理技术

通信、感知和计算一体化旨在利用同一无线资源和软硬件资源在同一系统内实现无线通信功能、无线感知功能和智能计算功能的协同增强、互利互惠。如前面所分析的，无线通信是指利用无线电信号将信息比特从发射机传输到接收机的能力，而无线感知是利用经目标反射的无线电回波分析并获得目标相关参数（如目标的位置、速度、形状等）的能力。然而，由于无线通信和无线感知使用同一频谱资源且共享同一硬件设施，不可避免的会造成通信和感知功能之间的相互干扰。再者，上下行链路使用同一频谱资源同时传输数据，会导致天线间的自干扰问题。此外，当存在多个通信终端时，多用户间产生的干扰也会大大降低系统的总体性能，如图 2.6 所示。因此，如何有效减少天线自干扰、功能间互干扰和多用户间干扰，在通信和感知两者之间取得良好的性能平衡是一项非常具有挑战性的任务。

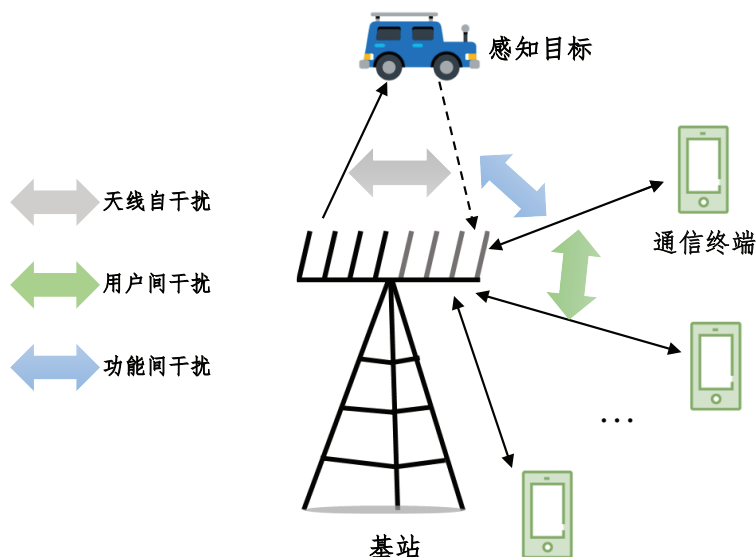


图 2.6 通信和感知一体化存在的干扰情况

对于天线间自干扰问题，可以通过增大收发天线距离、采用全双工技术以及成熟的数字自干扰消除技术来全面抑制天线间自干扰<sup>[98]</sup>。对于通信和感知之间的功能间干扰，可以将最小化感知和通信之间的互干扰作为目标设计通信和感知一体化波形<sup>[99]</sup>。此外，还可以通过合理分配波束资源来抑制通信和感知的功能间干扰及多用户间干扰<sup>[100]</sup>。值得一提的是，通过有效利用智能计算能力和引入智能化干扰管理技术，例如使用 DL 来探索通信和感知的内在关联，可以进行高效的干扰协调，实现所期望的通信性能和感知性能。

### 2.3.5 空中计算技术

在边缘智能网络中，常常需要对大规模边缘设备的海量数据进行分析和处理，这驱使未来无线网络从以数据为中心向以计算为中心转变。换句话说，人们往往只关注从海量数据中分析或处理得出的结果，而不是个体数据本身。例如，基于物联网的温湿度监测系统只对特定区域的平均温湿度感兴趣，而不是每个传感器上收集的观测数据。在此背景下，“先通信后计算”的传统分离架构需要在计算前恢复大量个体数据，无疑会给无线信道带来极大的压力。为了解决这个问题，一种名为“空中计算”的新型通信与计算一体化架构提供了一个实现快速无线数据聚合的有效解决方案<sup>[58]</sup>。如图 2.7 所示，它利用无线多址接入信道的天然叠加特性，可以直接计算各节点并发传输的目标函数，而无需恢复单个节点数据。



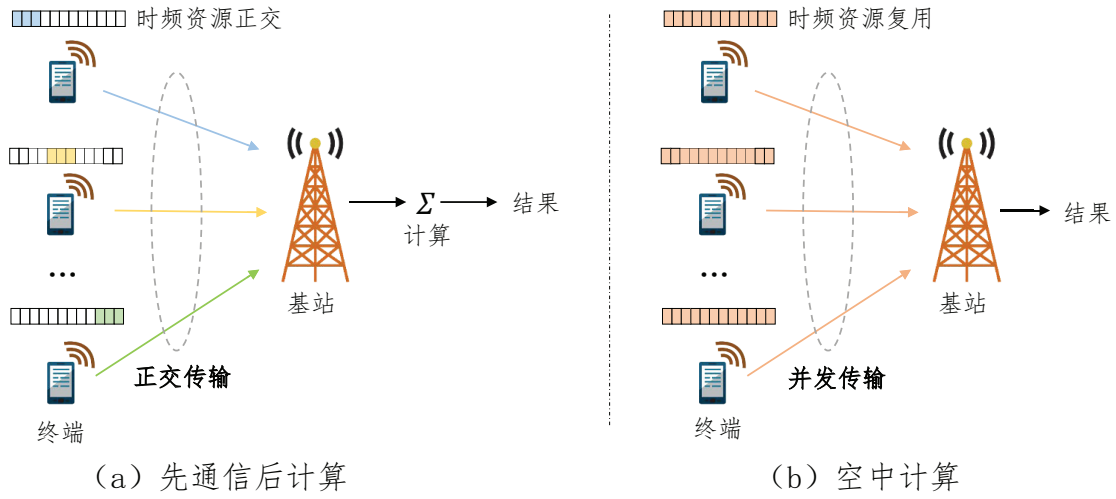


图 2.7 “先通信后计算”与“空中计算”架构比较

表 2.4 适用于空中计算的目标函数例子

| 函数     | $\varphi_k(\cdot)$                      | $\phi(\cdot)$           | $f(\cdot)$                                   |
|--------|---|-------------------------|--|
| 算数平均   | $\varphi_k = s_k$                       | $\phi = 1/K$            | $f = \frac{1}{K} \sum_{k=1}^K s_k$           |
| 加权求和   | $\varphi_k = \vartheta_k s_k$           | $\phi = 1$              | $f = \sum_{k=1}^K \vartheta_k s_k$           |
| 几何平均   | $\varphi_k = \ln(s_k)$                  | $\phi = \exp(\cdot)$    | $f = \left(\prod_{k=1}^K s_k\right)^{1/K}$   |
| 多项式求和  | $\varphi_k = \vartheta_k s_k^{\beta_k}$ | $\phi = 1$              | $f = \sum_{k=1}^K \vartheta_k s_k^{\beta_k}$ |
| 欧几里得范数 | $\varphi_k = s_k^2$                     | $\phi = \sqrt{(\cdot)}$ | $f = \sqrt{\sum_{k=1}^K s_k^2}$              |

考虑由 1 个基站和  $K$  个终端组成的系统，其中每个终端发送的数据信号为  $s_k, k = 1, \dots, K$ ，基站上的目标函数为  $f(s_1, \dots, s_K)$ 。该目标函数具有求和结构，因此称为列线性函数<sup>1</sup>，即

$$f(s_1, \dots, s_K) = \phi\left[\sum_{k=1}^K \varphi_k(s_k)\right], \quad (2.2)$$

其中  $\varphi_k(\cdot)$  是第  $k$  个终端上的预处理函数， $\phi(\cdot)$  是基站上的后处理函数。如表 2.4 所示，给定不同的预处理函数和后处理函数，可以得到一系列不同的目标函数。对于“先通信后计算”的传统分离架构，需要先在基站上逐一恢复出每个终端的数据  $s_k, \forall k$ ，再进行目标函数  $f(s_1, \dots, s_K)$  的计算，这需要使用  $K$  次无线信道。而对于空中计算架构，只需要提前对终端数据进行相应的预处理，通过所有终端的并发传输，再对经空中聚合的数据进行后处理，即可得到期望的目标函数值。可以看出，利用空中计算仅需使用 1 次无线信道就可

<sup>1</sup> 由于任意函数均可分解为若干个列线性函数<sup>[101]</sup>，空中计算理论上可以实现任何函数的计算。

以得到所有数据的计算结果，这大大减少了计算时延，提高了频谱利用率。因此，基于通信和计算一体化的空中计算架构特别适用于通信受限下的海量数据计算场景。

值得注意的是，利用空中计算实现无线信号的完美叠加的前提是分布终端的同步和无线信道的一致性。然而，无线多接入信道的衰落具有随机性，且通过无线链路传输的数据不可避免地会受到噪声和干扰的影响，这些都会降低计算的准确性。因此，需要为空中计算设计可靠的传输方案来实现高精度计算。

### 2.3.6 边缘联邦学习技术

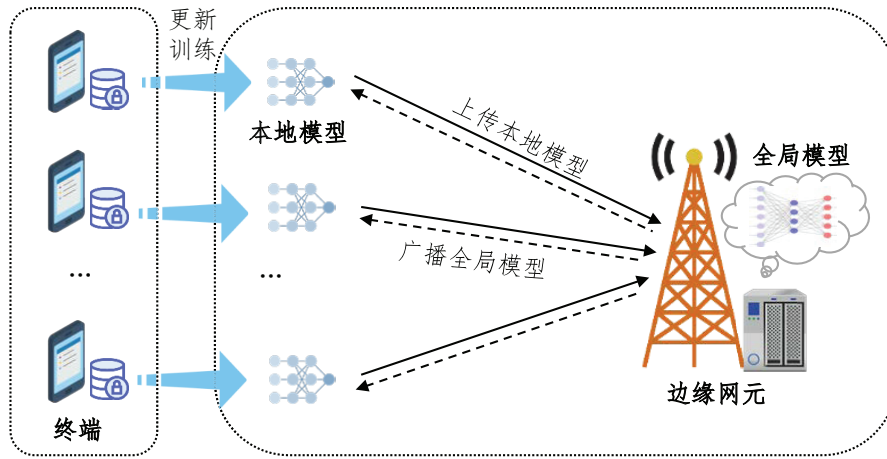


图 2.8 边缘联邦学习模型

在当前移动通信与 AI 协同发展的大环境下，边缘学习是从大量分散于不同终端的海量数据中提取重要信息的有效手段，也是将 AI 从云边智能推向边缘智能的重要使能技术<sup>[102]</sup>。基于数据共享的传统边缘智能需要将大量数据从边缘节点上传至边缘服务器，以训练一个全局模型。然而，由于海量数据传输，不仅会导致网络拥塞，还会造成个人隐私泄露。在这种背景下，基于模型共享的边缘 FL 成为边缘智能的一个新分支，它能够在网络边缘实现具有隐私保护性质的分布式机器学习<sup>[103]</sup>。

如图 2.8 所示，与基于数据共享的传统边缘智能不同，边缘 FL 利用多个智能终端的计算能力，由边缘网元（例如配有边缘服务器的基站）协调，以迭代的方式合作训练全局模型。考虑有  $K$  个终端与 1 个配有边缘服务器的基站利用边缘 FL 共同训练一个全局模型，其中每个终端都有自己的数据集  $D_k, \forall k$ 。在 FL 的第  $t$  轮迭代（也称为通信轮）中，基站首先将全局模型参数  $w^{[t]}, \forall t$  通过下行链路广播给终端，以便所有参与 FL 的终端都可以同

步并初始化它们的本地模型。接着，每个参与 **FL** 的终端分别基于自己的数据集训练并更新其本地模型参数：

$$w_k^{[t+1]} = w_k^{[t]} - \eta_k \frac{\partial \iota_k(w^{[t]}, D_k)}{\partial w}, \forall t, \quad (2.3)$$

其中  $\eta_k$  为更新步长， $\iota_k(\cdot)$  为训练损失函数。然后，终端将更新后的本地模型参数通过上行链路发送至基站的边缘服务器进行全局模型参数聚合：

$$w^{[t+1]} = \psi(w_1^{[t+1]}, w_2^{[t+1]}, \dots, w_K^{[t+1]}), \forall t \quad (2.4)$$

其中  $\psi(\cdot)$  为聚合函数，在实际应用中通常取平均。直至全局模型收敛，迭代停止。由于全局模型的整个训练过程只涉及模型参数的传输，而不是真实的原始数据，因此 **FL** 对边缘智能网络非常有吸引力，可以更好地对具有隐私性的数据进行分析 and 处理。

## 2.4 本章小结

本章详细介绍了面向边缘智能网络的通感算一体化的基础理论和关键技术，包括通信、感知和计算三者功能定义和性能评价指标、通感算一体化程度和发展阶段、通感算一体化系统架构和使能通感算一体化的关键技术。



## 3 面向感知的通信和计算一体化

### 3.1 引言

在边缘智能网络中，多源节点采集到的感知数据具有不同的结构和语义，需要对获得的数据进行实时计算，以减少数据的传输量，提高信息感知的效率。基于此，通过结合非正交传输和无线边缘 FL 对海量数据进行处理，能够将“先感知再通信最后计算”的模式转变为基于“感知数据的模型参数交互”的模式。考虑到数据隐私，基于模型共享的 FL 将模型参数而不是私有数据从边缘设备迁移到服务器，对边缘智能网络特别有吸引力。由于无线资源的有限性和高维模型参数的迭代传输，通信效率成为了边缘智能网络实施 FL 的主要挑战。为此，已有一些工作试图从机器学习的角度解决该问题。一般地，可以减少参与 FL 的终端数量或者通过量化、稀疏化等方式来压缩模型参数，降低通信负荷。然而，这些方法将无线信道视为终端与基站之间独立的无差错传输管道。事实上，模型参数在无线信道上的传输不可避免地会受到信道衰落、干扰和噪声的影响，导致 FL 的性能往往不能令人满意。因此，有必要为边缘智能网络执行 FL 设计一个通信高效的传输方案来提高学习的性能。基于此，本章研究了信道衰落、干扰、噪声和信道估计误差等不利因素对通过无线信道进行模型广播和模型聚合过程产生的影响。为缓解不利因素对边缘智能网络的影响并同时保证网络的通信质量和计算效率，提出了一种基于空中计算的联合设备选择和收发机设计的鲁棒算法来共同提高通信效率和学习准确性。

内容概述：本章提出了一种面向感知的通信和计算一体化架构，利用非正交传输和边缘 FL 实现对分布式节点采集到的海量感知数据的高效处理。针对信道衰落、干扰和噪声以及信道估计误差等不利因素对 FL 模型参数传输的影响，提出了一种基于空中计算的联合设备选择和收发机设计的鲁棒算法来提高通信和计算一体化的性能。具体结构安排如下：3.2 节介绍了面向感知的通信和计算一体化系统模型，3.3 节提出了一种联合设备选择和收发机设计算法，3.4 节利用数值仿真验证了所提算法的有效性和鲁棒性，3.5 节总结了本章内容。

## 3.2 系统模型

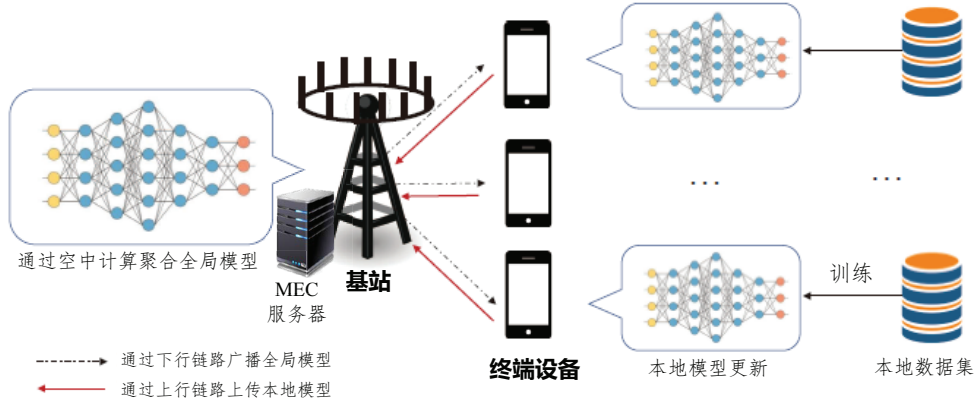


图 3.1 面向感知的通信和计算一体化系统模型

如图 3.1 所示，考虑一个面向感知的通信和计算一体化系统，它由一个配备  $N$  根天线并具有边缘服务器的基站和  $K$  个单天线智能终端设备组成。智能终端设备具有感知能力，可以获得大量的感知信息，并将其作为模型训练的数据集。每个智能终端先利用本地数据集进行模型训练，然后将模型参数传递给基站用于构建全局模型。下面，分别从计算和通信的角度描述这一过程。

### 3.2.1 计算模型

对于 FL 而言，全局模型  $\mathbf{q}$  的共享学习过程是由多个选定的智能终端设备协同训练完成的。每个选中的设备  $k \in \mathcal{K} = \{1, 2, \dots, K\}$  都有自己的本地数据集  $D_k$ ，其标签数据样本为  $\{(\mathbf{g}_j, l_j)\} \in D_k$ ，其中  $(\mathbf{g}_j, l_j)$  表示由训练向量样本  $\mathbf{g}_j$  和真实标签  $l_j$  组成的输入输出数据对。对于样本  $j$ ，定义其损失函数为  $f_j(\mathbf{q}; \mathbf{g}_j, l_j)$ ，则所有数据集的全局损失函数可以表示为

$$\mathcal{F}(\mathbf{q}) = \frac{1}{|D_{\text{total}}|} \sum_{j \in D_{\text{total}}} f_j(\mathbf{q}; \mathbf{g}_j, l_j), \quad (3.1)$$

其中  $D_{\text{total}} = \{D_1, D_2, \dots, D_K\}$ ， $|D_{\text{total}}|$  是数据集  $D_{\text{total}}$  的大小。学习目标是通过最小化全局损失函数  $\mathcal{F}(\mathbf{q})$  来优化全局模型  $\mathbf{q}$ ，即

$$\mathbf{q}^* = \arg \min \mathcal{F}(\mathbf{q}). \quad (3.2)$$

一般地，可以采用基于模型平均的 FL 算法，以迭代的方式来找到最优全局模型  $\mathbf{q}^*$ 。具体而言，该算法通过平均化本地模型  $\mathbf{q}_k$  来获得全局模型  $\mathbf{q}$ ，而不需要在智能设备之间共享

**算法 1:** 基于模型平均的 FL 算法

输入: 本地数据集  $D_i$ , 参与 FL 的设备集  $\mathcal{S}_t$

输出: 全局模型  $\mathbf{q}$

- 1: 初始化 全局模型  $\mathbf{q}^{[0]}$ , 通信轮索引  $t = 1$ ;
- 2: **repeat**
- 3: 基站选择参与 FL 的设备集  $\mathcal{S}_t \in \mathcal{K}$ ;
- 4: 基站通过下行链路向参与 FL 的设备广播当前的全局模型  $\mathbf{q}^{[t-1]}$ ;
- 5: 根据接收到的模型  $\mathbf{q}_i^{[t]} = \mathbf{q}^{[t-1]} + \Delta\mathbf{q}_i^{[t]}$ , 每个设备  $i \in \mathcal{S}_t$  在自己的数据集上  $D_i$  执行随机梯度下降法来更新本地模型;
- 6: 参与 FL 的设备通过上行链路向基站上传其更新的本地模型  $\mathbf{q}_i^{[t]}$ ;
- 7: 基站更新全局模型为  $\mathbf{q}^{[t]} = \frac{1}{\sum_{i \in \mathcal{S}_t} |D_i|} \sum_{i \in \mathcal{S}_t} |D_i| \mathbf{q}_i^{[t]} + \Delta\mathbf{q}^{[t]}$ ;
- 8: 更新通信轮索引  $t = t + 1$ ;
- 9: **until** 全局模型收敛

数据集。由于模型参数在基站和终端设备之间通过无线信道传输, 接收到的模型参数可能会产生失真。例如, 信道估计误差、反馈量化误差或者在衰落信道上的信号采集时延等都可能导致在基站和终端设备上模型参数的不完美恢复。具体地, 在第  $t$  轮迭代中, 设备和基站接收到的模型失真可以被建模为

$$\text{设备: } \mathbf{q}_i^{[t]} = \mathbf{q}^{[t-1]} + \Delta\mathbf{q}_i^{[t]}, i \in \mathcal{S}_t, \quad (3.3a)$$

$$\text{基站: } \mathbf{q}^{[t]} = \frac{1}{\sum_{i \in \mathcal{S}_t} |D_i|} \sum_{i \in \mathcal{S}_t} |D_i| \mathbf{q}_i^{[t]} + \Delta\mathbf{q}^{[t]}, \quad (3.3b)$$

其中  $\mathcal{S}_t$  是第  $t$  轮迭代中选择的设备集,  $\Delta\mathbf{q}_i^{[t]}$  和  $\Delta\mathbf{q}^{[t]}$  分别是第  $t$  轮迭代中第  $i$  个设备的模型广播误差和基站的聚合误差。综上, 基于模型平均的 FL 算法被描述为算法 1。请注意, 算法 1 中的第 5 步, 本地模型更新采用随机梯度下降法, 其单步的随机梯度下降可以被表示为

$$\mathbf{q}_i^{[t]} = \mathbf{q}_i^{[t-1]} - \eta_i \nabla \mathcal{F}_i(\mathbf{q}_i^{[t-1]}), i \in \mathcal{S}_t, \quad (3.4)$$

其中  $\eta_i$  是学习步长,  $\nabla \mathcal{F}_i(\mathbf{q}_i^{[t-1]})$  是  $\mathcal{F}_i(\mathbf{q}_i^{[t-1]})$  在点  $\mathbf{q}_i^{[t-1]}$  的梯度。根据式 (3.4) 中梯度下降法的更新规则, 通过迭代  $\tau$  步 ( $\tau \geq 1$ ) 直至收敛, 可以得到更新的本地模型  $\mathbf{q}_i^{[t]}$ 。

接下来在常用的 MNIST 数据集上用基于模型平均的 FL 算法训练 CNN, 以展示设备选择、模型聚合误差和模型广播误差对预测精度的影响<sup>1</sup>。从图 3.2 可以看出, 基线情况 (没有聚合误差和广播误差的理想情况) 与没有聚合误差的情况之间的性能差距是可以忽

<sup>1</sup>详细的实验仿真参数设置在 3.4 节中。请注意, 模型误差主要是由模型聚合和无线信道广播过程中的信道衰落和加性噪声引起的。因此, 在实验中, 对基站和设备的接收模型参数中分别加入高斯噪声可以有效模拟由模型聚合和模型广播误差引起的失真。

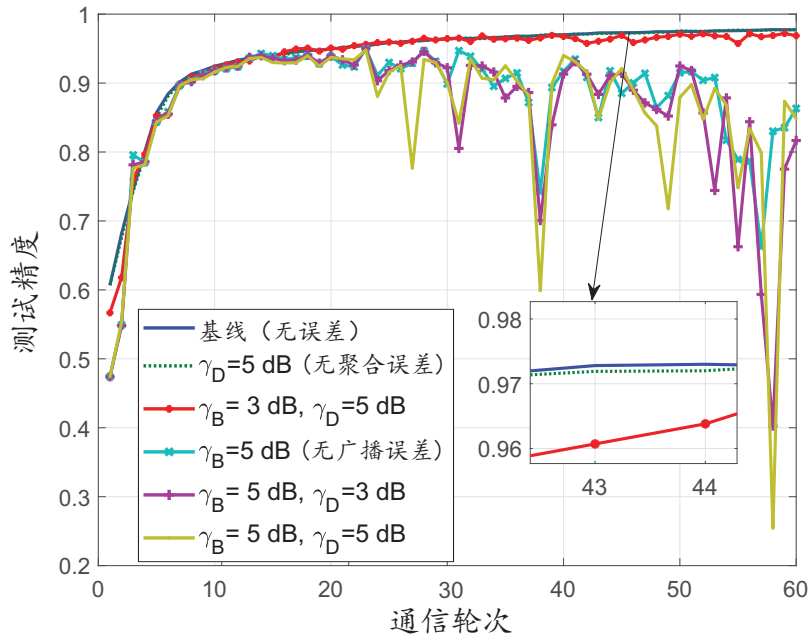


图 3.2 在所有设备参与FL的情况下，不同误差大小对 FL 测试精度的影响

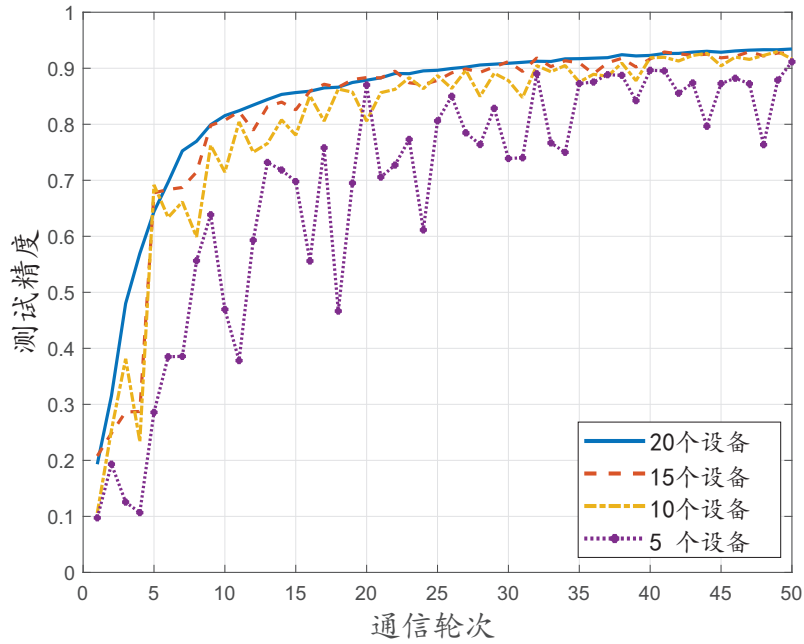


图 3.3 在无传输误差的情况下，参与 FL 的设备数量对 FL 测试精度的影响

略的，其中  $\gamma_D$  和  $\gamma_B$  分别为终端和基站接收到的模型参数的均方误差（Mean Square Error, MSE），代表了广播误差和聚合误差的大小。这是因为模型广播误差虽然作用于初始本地模型，但设备可以调整其神经网络以获得良好的训练性能，减小广播误差的影响。但是，如果存在聚合误差，则广播误差的影响会加剧性能的下降。因此，在 FL 的设计中需要同时考虑模型的聚合误差和广播误差。此外，从图 3.3 中可以发现，参与 FL 的设备数量也



是影响 FL 预测精度的关键因素。在没有误差的情况下，参与 FL 的设备数量越多，FL 全局模型的性能越好。但是，在无线信道存在干扰、噪声和衰落的情况下，参与 FL 的设备数量越多，模型传输误差积累的可能性也越大。因此，一方面需要考虑如何有效减少无线信道不利因素对模型参数传递过程中造成的影响，另一方面需要考虑在误差可控范围内，增加参与 FL 的设备数量。

### 3.2.2 通信模型

如上所述，模型广播和模型聚合都是通过无线信道传递参数来实现的。在本节中，考虑准静态的平坦瑞利衰落信道，其信道系数在一个通信块内保持不变，但在连续块上独立衰落。由于智能终端设备的数据存储和计算能力有限，设置 FL 模型参数的维数为  $d$ ，以保证在一个通信块内传输完整的模型参数。为了简化符号表达，下面将去掉通信轮索引  $t$ ，仅讨论传输矢量信号的一个典型条目，即用  $q$  表示基站上的全局模型，用  $q_i$  表示在第  $i$  个设备上的本地模型。

#### 3.2.2.1 模型广播

在模型广播阶段，基站通过下行链路广播基于全局模型参数构建的发射信号  $\mathbf{x}_B$ ，如下所示

$$\mathbf{x}_B = \mathbf{w}q, \quad (3.5)$$

其中， $\mathbf{w} \in \mathbb{C}^{N \times 1}$  代表基站的发射机向量。那么，第  $k$  个设备的接收信号可以表示为

$$y_k = \mathbf{h}_k^H \mathbf{x}_B + n_k, \quad (3.6)$$

其中  $\mathbf{h}_k$  表示从基站到第  $k$  个设备的信道增益， $n_k$  是方差为  $\sigma_0^2$  的加性高斯白噪声 (Additive White Gaussian Noise, AWGN)。因此，用于衡量第  $k$  个设备的模型失真的 MSE 可以表示为

$$\begin{aligned} \text{MSE}_k^D &= \mathbb{E} \left\{ (v_k y_k - q) (v_k y_k - q)^H \right\} \\ &= v_k \mathbf{h}_k^H \mathbf{w} \mathbf{w}^H \mathbf{h}_k v_k^H - v_k \mathbf{h}_k^H \mathbf{w} + \sigma_0^2 v_k v_k^H - \mathbf{w}^H \mathbf{h}_k v_k^H + 1 \\ &= |v_k \mathbf{h}_k^H \mathbf{w} - 1|^2 + \sigma_0^2 |v_k|^2, \end{aligned} \quad (3.7)$$

其中， $v_k$  表示第  $k$  个设备的接收机标量。

## 3.2.2.2 模型聚合

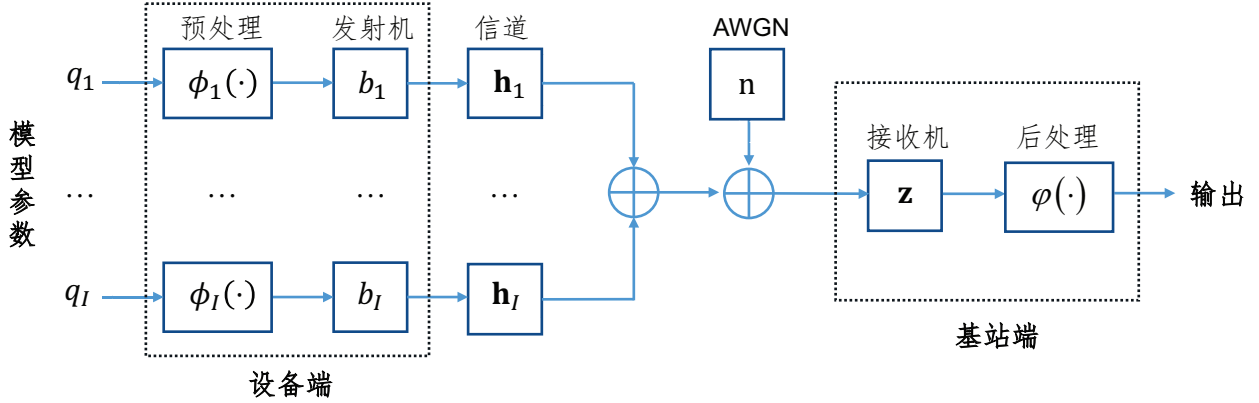


图 3.4 基于空中计算的模型聚合框图

由于模型聚合的加和平均属于列线性函数的范畴<sup>[104]</sup>，可以采用空中计算来提高分布式设备模型聚合的通信效率，这一模型聚合流程如图 3.4 所示。基于式 (3.1) 中的全局损失函数，定义  $\phi_i(x) = |D_i|x$  和  $\varphi(x) = \frac{1}{\sum_{i \in \mathcal{S}} |D_i|}x$  分别为第  $i$  个设备的预处理函数和基站的后处理函数。因此，在基站上聚合本地模型的目标为

$$\tilde{q} = \frac{1}{\sum_{i \in \mathcal{S}} |D_i|} \sum_{i \in \mathcal{S}} |D_i|q_i, \quad (3.8)$$

其中  $\mathcal{S}$  是参与 FL 的设备集。为了简化分析，假设所有本地数据集具有统一的大小，即  $|D_k| = D, \forall k \in \mathcal{K}$ 。在这种情况下，式 (3.8) 中模型聚合的目标可以简化为

$$\tilde{q} = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} q_i. \quad (3.9)$$

为了最小化由信道衰落和噪声引起的目标函数信号的失真，需要在基站端进行接收波束成形。因此，基站接收到的信号可以表示为

$$\hat{q} = \mathbf{z}^H \sum_{i \in \mathcal{S}} \mathbf{h}_i b_i q_i + \mathbf{z}^H \mathbf{n}, \quad (3.10)$$

其中  $\mathbf{z} \in \mathbb{C}^{N \times 1}$  表示基站的接收机向量， $b_i$  表示第  $i$  个设备的发射机标量， $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \sigma_1^2 \mathbf{I})$  是基站端的 AWGN。一般地，模型聚合的精度可以采用  $\sum_{i \in \mathcal{S}} q_i$  和  $\hat{q}$  之间的 MSE 来表示，即

$$\text{MSE}_B = \mathbb{E} \left\{ \left( \hat{q} - \sum_{i \in \mathcal{S}} q_i \right) \left( \hat{q} - \sum_{i \in \mathcal{S}} q_i \right)^H \right\} = \sum_{i \in \mathcal{S}} |\mathbf{z}^H \mathbf{h}_i b_i - 1|^2 + \sigma_1^2 \|\mathbf{z}\|^2. \quad (3.11)$$

检查式 (3.7) 和式 (3.11)，可以发现模型广播误差和聚合误差受参与 FL 的设备数量、基站收发机和设备收发机的共同影响，从而决定 FL 的性能。因此，根据瞬时信道状态信息（Channel State Information, CSI）进行联合设备选择和收发机设计是有意义的。此外，在实际的边缘智能网络中，由于不完美的信道估计或反馈，有时只能获得部分 CSI。因此，在联合设备选择和收发机设计时，有必要考虑信道的不确定性，即鲁棒传输方案设计。为了表征信道的不确定性，采用常见的确定性 CSI 误差模型<sup>[105, 106]</sup>。具体而言，从第  $i$  个设备到基站的之间的真实 CSI  $\mathbf{h}_i$  可以被建模为

$$\mathcal{H}_i \triangleq \left\{ \mathbf{h}_i = \hat{\mathbf{h}}_i + \mathbf{e}_i \mid \|\mathbf{e}_i\| \leq \varepsilon_i \right\}, \quad (3.12)$$

其中  $\hat{\mathbf{h}}_i$  是获取的 CSI， $\mathbf{e}_i$  是 CSI 误差，其范数在一个给定的范围  $\varepsilon_i$  内。

### 3.3 联合设备选择和收发机设计

受到以上关键观察结果的启发，即全局模型通过下行链路产生的广播误差和本地模型通过上行链路产生的聚合误差的组合会导致 FL 的性能显著下降，而模型训练的收敛可以通过选择更多的设备参与 FL 来加速。因此，本节的目标是在边缘智能网络中进行联合设备选择和收发机设计来提升通信和计算一体化的系统性能。

#### 3.3.1 问题构造

联合优化方案考虑式 (3.12) 中信道不确定性的特点，在分别满足模型广播的 MSE 和模型聚合的 MSE 要求的情况下，最大化参与 FL 设备的数量。其等价于以下优化问题：

$$\max_{\mathcal{S}, v_i, b_i, \mathbf{w}, \mathbf{z}} |\mathcal{S}| \quad (3.13a)$$

$$\text{s.t.} \quad \max_{\mathbf{h}_i \in \mathcal{H}_i} \sum_{i \in \mathcal{S}} \left| \mathbf{z}^H \mathbf{h}_i b_i - 1 \right|^2 + \sigma_1^2 \|\mathbf{z}\|^2 \leq \gamma_B, \quad (3.13b)$$

$$\max_{\mathbf{h}_i \in \mathcal{H}_i} \left| v_i \mathbf{h}_i^H \mathbf{w} - 1 \right|^2 + \sigma_0^2 |v_i|^2 \leq \gamma_{D,i}, i \in \mathcal{S}, \quad (3.13c)$$

$$b_i \leq \sqrt{P_i}, i \in \mathcal{S}, \quad (3.13d)$$

$$\|\mathbf{w}\|^2 \leq P_{\max}, \quad (3.13e)$$

其中目标函数 (3.13a) 中的  $|\mathcal{S}|$  是所选设备集的数量。 $\gamma_B$  和  $\gamma_{D,i}$  分别是基站上模型聚合和第  $i$  个设备上模型广播的 MSE 要求。约束 (3.13d) 和 (3.13e) 分别是第  $i$  个设备和基站的发

射功率限制条件，其中  $P_i$  是第  $i$  个设备的最大发射功率预算， $P_{\max}$  是基站的最大发射功率预算。由于复杂的目标函数  $|\mathcal{S}|$ ，优化问题 (3.13) 是非凸的。为了解决这个问题，需要引入一个辅助变量  $\boldsymbol{\chi} = [\chi_1, \chi_2, \dots, \chi_K]^T$  来表示设备被选择的优先级，其中  $\chi_k$  表征了第  $k$  个设备所要求的 MSE 和可实现的 MSE 之间的差距。换句话说， $\chi_k$  的值越小，第  $k$  个设备被选中参与 FL 的优先级就越高。通过将  $\chi_k$  添加到相应 MSE 约束条件的右侧，最大化所选设备数量就等价于最小化非零  $\chi_k$  的数量<sup>[107]</sup>，其中  $\chi_k = 0$  表示第  $k$  个设备被选择参与 FL。因此，可以将问题 (3.13) 重新表示为以下稀疏最小化问题：

$$\min_{v_k, b_k, \mathbf{w}, \mathbf{z}, \boldsymbol{\chi}, \rho_k, \vartheta_k} \|\boldsymbol{\chi}\|_0 \quad (3.14a)$$

$$\text{s.t.} \quad \max_{\|\mathbf{e}\| \leq \varepsilon_k} \left| \mathbf{z}^H (\hat{\mathbf{h}}_k + \mathbf{e}_k) b_k - 1 \right|^2 \leq \rho_k + \chi_k, \quad (3.14b)$$

$$\sum_{k \in \mathcal{K}} \rho_k + \sigma_1^2 \|\mathbf{z}\|^2 \leq \gamma_B, \quad (3.14c)$$

$$\max_{\|\mathbf{e}\| \leq \varepsilon_k} \left| v_k (\hat{\mathbf{h}}_k + \mathbf{e}_k)^H \mathbf{w} - 1 \right|^2 \leq \vartheta_k, \quad (3.14d)$$

$$\vartheta_k + \sigma_0^2 |v_k|^2 \leq \gamma_{D,k} + \chi_k, \quad (3.14e)$$

$$\|\mathbf{w}\|^2 \leq P_{\max}, \quad (3.14f)$$

$$b_k \leq \sqrt{P_k}, \quad (3.14g)$$

$$\rho_k \geq 0, \vartheta_k \geq 0, \quad (3.14h)$$

$$\chi_k \geq 0, \quad (3.14i)$$

其中， $\rho_k$  和  $\vartheta_k$  是用于分割原始 MSE 约束条件的辅助变量。然而，由于信道不确定性和耦合变量，即发射机  $\mathbf{w}, b_i$  和接收机  $\mathbf{z}, v_i$ ，引起了非凸稀疏目标函数 (3.14a) 和非凸约束 (3.14b) 和 (3.14d)，稀疏最小化问题 (3.14) 仍然是非凸的。

### 3.3.2 算法设计

为了解决上述非凸优化问题 (3.14)，本节提出了一个两步算法框架，包括确定设备优先级顺序的稀疏性诱导和在保证 MSE 要求的情况下找到最多同时参与 FL 的设备数量的可行性检测。

## 3.3.2.1 稀疏性诱导

对于以  $l_0$  范数形式存在的非凸稀疏目标函数，可以采用一种基于最小化不可行性之和的  $l_1$  松弛法<sup>[107]</sup>，即用  $l_1$  范数代替  $l_0$  范数。因此，问题 (3.14) 可以被重构为

$$\begin{aligned} \min_{v_k, b_k, \mathbf{w}, \mathbf{z}, \chi_k, \rho_k, \vartheta_k} \quad & \|\boldsymbol{\chi}\|_1 \\ \text{s.t.} \quad & (3.14b) - (3.14i). \end{aligned} \quad (3.15)$$

此外，为了解决涉及信道不确定性的非凸约束 (3.14b) 和 (3.14d)，需要用到以下引理：

**引理3.1.** (舒尔补引理<sup>[107]</sup>) 设  $\mathbf{D}$  是一个定义为  $\mathbf{D} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^H \\ \mathbf{B} & \mathbf{C} \end{bmatrix}$  的厄米矩阵。当且仅当矩阵  $\mathbf{C}$  可逆时满足  $\mathbf{A} - \mathbf{B}^H \mathbf{C}^{-1} \mathbf{B} \succeq \mathbf{0}$  或者当矩阵  $\mathbf{A}$  可逆时满足  $\mathbf{C} - \mathbf{B}^H \mathbf{A}^{-1} \mathbf{B} \succeq \mathbf{0}$ ，矩阵  $\mathbf{D}$  是半正定的。

**引理3.2.** 定义一个矩阵函数为  $\mathbf{F}(\mathbf{x}) = \mathbf{A} - \mathbf{B}^H \mathbf{x} \mathbf{c}^H - \mathbf{c} \mathbf{x}^H \mathbf{B}$ ，其中  $\mathbf{x} \in \mathbb{C}^{m \times 1}$  为变量向量， $\mathbf{B} \in \mathbb{C}^{m \times n}$  为常数矩阵， $\mathbf{c} \in \mathbb{C}^{n \times 1}$  为常数向量， $\mathbf{A} \in \mathbb{C}^{n \times n}$  为常数厄米矩阵。那么，

$$\mathbf{F}(\mathbf{x}) \succeq \mathbf{0}, \forall \mathbf{x} : \|\mathbf{x}\| \leq \varpi.$$

当且仅当存在  $\lambda \geq 0$  时，以下不等式成立：

$$\begin{bmatrix} \mathbf{A} - \lambda \mathbf{c} \mathbf{c}^H & -\varpi \mathbf{B}^H \\ -\varpi \mathbf{B} & \lambda \mathbf{I}_n \end{bmatrix} \succeq \mathbf{0}.$$

*证明:* 请参考附录 A.1.1。 □

现在，先处理约束条件 (3.14b)，即

$$\begin{aligned} & \left| \mathbf{z}^H (\hat{\mathbf{h}}_k + \mathbf{e}_k) b_k - 1 \right|^2 \\ &= \left| \left( \mathbf{z}^H \hat{\mathbf{h}}_k b_k - 1 \right) + \mathbf{z}^H \mathbf{e}_k b_k \right|^2 \\ &\leq \rho_k + \chi_k. \end{aligned} \quad (3.16)$$

根据引理 3.1, 约束条件 (3.16) 可以被转换为

$$\begin{bmatrix} \rho_k + \chi_k & b_k^H \hat{\mathbf{h}}_k^H \mathbf{z} - 1 + b_k^H \mathbf{z} \mathbf{e}_k^H \\ b_k \mathbf{z}^H \hat{\mathbf{h}}_k - 1 + b_k \mathbf{z}^H \mathbf{e}_k & 1 \end{bmatrix} \succeq \mathbf{0}, \forall \mathbf{e}_k : \|\mathbf{e}_k\| \leq \varepsilon_k. \quad (3.17)$$

为了方便表示, 令

$$\mathbf{A}_k \triangleq \begin{bmatrix} \rho_k + \chi_k & b_k^H \hat{\mathbf{h}}_k^H \mathbf{z} - 1 \\ b_k \mathbf{z}^H \hat{\mathbf{h}}_k - 1 & 1 \end{bmatrix}, \mathbf{B}_k \triangleq \begin{bmatrix} \mathbf{0} & -b_k^H \mathbf{z} \end{bmatrix}, \mathbf{c} \triangleq \begin{bmatrix} 1 & 0 \end{bmatrix}^T. \quad (3.18)$$

结合式 (3.17), 可以得到

$$\mathbf{F}_k(\mathbf{e}_k) = \mathbf{A}_k - \mathbf{B}_k^H \mathbf{e}_k \mathbf{c}^H - \mathbf{c} \mathbf{e}_k^H \mathbf{B}_k \succeq \mathbf{0}, \forall \mathbf{e}_k : \|\mathbf{e}_k\| \leq \varepsilon_k. \quad (3.19)$$

接着, 根据引理 3.2, 约束条件 (3.14b) 可以被重构为

$$\begin{bmatrix} \rho_k + \chi_k - \alpha_k & b_k^H \hat{\mathbf{h}}_k^H \mathbf{z} - 1 & \mathbf{0} \\ b_k \mathbf{z}^H \hat{\mathbf{h}}_k - 1 & 1 & \varepsilon_k b_k \mathbf{z}^H \\ \mathbf{0} & \varepsilon_k b_k^H \mathbf{z} & \alpha_k \mathbf{I}_N \end{bmatrix} \succeq \mathbf{0}, \exists \alpha_k \geq 0. \quad (3.20)$$

相似地, 如果存在  $\beta_k \geq 0$ , 约束条件 (3.14d) 可以被重构为

$$\begin{bmatrix} \vartheta_k - \beta_k & v_k \hat{\mathbf{h}}_k^H \mathbf{w} - 1 & \mathbf{0} \\ \mathbf{w}^H \hat{\mathbf{h}}_k v_k^H - 1 & 1 & \varepsilon_k \mathbf{w} v_k^H \\ \mathbf{0} & \varepsilon_k v_k^H \mathbf{w}^H & \beta_k \mathbf{I}_N \end{bmatrix} \succeq \mathbf{0}, \exists \beta_k \geq 0. \quad (3.21)$$

注意, 由于耦合变量的存在, 约束条件 (3.20) 和 (3.21) 仍然是非凸的。因此, 需要利用交替优化 (Alternating Optimization, AO) 方法, 对问题 (3.15) 进行解耦, 然后迭代求解<sup>2</sup>。具体而言, 问题 (3.15) 可以被分解为两个子问题, 一个是基站收发机优化设计, 另一个是设备收发机优化设计。这两个子问题在迭代中依次进行求解, 直到原始问题的目标函数值收敛, 迭代停止, 得到可行解。在这种情况下, 通过给定基站收发机  $\{\mathbf{w}, \mathbf{z}\}$ , 优化设备收发

<sup>2</sup> AO 算法能够在系统性能和计算复杂度之间取得平衡, 因此被现存的工作广泛采用。此外, 根据文献[108]和[109]可得, 当迭代过程中每一个子问题的解都是其全局最优解时, AO 算法的收敛性能够得到保证。

机  $\{v_k, b_k\}$  设计的子问题可以表示为

$$\min_{v_k, b_k, \boldsymbol{\chi}, \rho_k, \vartheta_k, \alpha_k, \beta_k} \|\boldsymbol{\chi}\|_1 \quad (3.22a)$$

$$\text{s.t.} \quad (3.14c), (3.14e), (3.14g) - (3.14i), (3.20), (3.21),$$

$$\alpha_k \geq 0, \quad (3.22b)$$

$$\beta_k \geq 0. \quad (3.22c)$$

可以看出，问题 (3.22) 是一个联合凸问题，因此可以使用一些现成的凸优化工具包来求解，例如 CVX<sup>[110]</sup>。然后，根据从问题 (3.22) 中得到的设备收发机的解  $\{v_k, b_k\}$ ，来优化基站收发机  $\{\mathbf{w}, \mathbf{z}\}$ 。优化基站收发机的子问题可以表示为

$$\min_{\mathbf{w}, \mathbf{z}, \boldsymbol{\chi}, \rho_k, \vartheta_k, \alpha_k, \beta_k} \|\boldsymbol{\chi}\|_1 \quad (3.23)$$

$$\text{s.t.} \quad (3.14c), (3.14e), (3.14f), (3.14h) - (3.14i),$$

$$(3.20), (3.21), (3.22b), (3.22c).$$

不难看出，子问题 (3.23) 也是一个联合凸问题，可以使用凸优化工具包得到最优解。因此，通过迭代优化这两个子问题，可以获得稀疏向量  $\boldsymbol{\chi}^*$ 。接下来，进行第二步可行性检测，即根据求解的稀疏向量，找到可行的设备集  $\mathcal{S}$ 。

### 3.3.2.2 可行性检测

基于上述步骤获得的稀疏向量  $\boldsymbol{\chi}$ ，首先将其元素  $\chi_k, \forall k$  按升序排序，即  $\chi_{\pi(1)} \leq \chi_{\pi(2)} \leq \dots \leq \chi_{\pi(K)}$ ，其中  $\chi_{\pi(i)}$  代表  $\{\chi_1, \dots, \chi_K\}$  中第  $i$  个最小的元素。然后，寻找最大值  $m \in [1, K]$  使得所选设备集  $\mathcal{S}^{[m]} = \{\pi(1), \pi(2), \dots, \pi(m)\}$  中的所有设备是可行的，即满足 MSE 的约束条件。具体而言，对于一个给定的设备集  $\mathcal{S}^{[m]}$ ，通过比较所要求的 MSE 值和实际的 MSE 值，可以检查解的可行性。这一操作可以表述为如下优化问题：

$$\min_{\mathbf{w}, \mathbf{z}, b_i, v_i} \max_{\|\mathbf{e}_i\| \leq \varepsilon_i} \sum_{i \in \mathcal{S}^{[m]}} \left| \mathbf{z}^H (\hat{\mathbf{h}}_i + \mathbf{e}_i) b_i - 1 \right|^2 + \sigma_1^2 \|\mathbf{z}\|^2 \quad (3.24a)$$

$$\text{s.t.} \quad \max_{\|\mathbf{e}_i\| \leq \varepsilon_i} \left| v_k (\hat{\mathbf{h}}_k + \mathbf{e}_k)^H \mathbf{w} - 1 \right|^2 + \sigma_0^2 |v_i|^2 \leq \gamma_{D,i}, \forall i \in \mathcal{S}^{[m]}, \quad (3.24b)$$

$$\|\mathbf{w}\|^2 \leq P_{\max}, \quad (3.24c)$$

$$b_i \leq \sqrt{P_i}, \forall i \in \mathcal{S}^{[m]}. \quad (3.24d)$$

由于问题 (3.24) 的结构与问题 (3.15) 相似，即由耦合变量和信道不确定性导致的非凸性，仍然需要进行优化分解再迭代求解，并利用引理 3.1 和引理 3.2 来处理问题 (3.24) 的非凸性。这里省略问题凸性的转换步骤，直接给出可行性检测分解并转换后的两个凸的子问题。在固定基站收发机  $\{\mathbf{w}, \mathbf{z}\}$  的情况下，优化设备收发机  $\{v_i, b_i\}$  的子问题可以表示为

$$\min_{v_i, b_i, \rho_i, \vartheta_i, \alpha_i, \beta_i} \sum_{i \in \mathcal{S}^{[m]}} \rho_i + \sigma_1^2 \|\mathbf{z}\|^2, \quad (3.25a)$$

$$\text{s.t.} \quad (3.14h), (3.14i), (3.22b), (3.22c), (3.24d),$$

$$\begin{bmatrix} \rho_i - \alpha_i & b_i^H \hat{\mathbf{h}}_i^H \mathbf{z} - 1 & \mathbf{0} \\ b_i \mathbf{z}^H \hat{\mathbf{h}}_i - 1 & 1 & \varepsilon_i b_i \mathbf{z}^H \\ \mathbf{0} & \varepsilon_i b_i^H \mathbf{z} & \alpha_i \mathbf{I}_N \end{bmatrix} \succeq \mathbf{0}, \forall i \in \mathcal{S}^{[m]} \quad (3.25b)$$

$$\begin{bmatrix} \vartheta_i - \beta_i & v_i \hat{\mathbf{h}}_i^H \mathbf{w} - 1 & \mathbf{0} \\ \mathbf{w}^H \hat{\mathbf{h}}_i v_i^H - 1 & 1 & \varepsilon_i \mathbf{w} v_i \\ \mathbf{0} & \varepsilon_i v_i^H \mathbf{w}^H & \beta_i \mathbf{I}_N \end{bmatrix} \succeq \mathbf{0}, \forall i \in \mathcal{S}^{[m]} \quad (3.25c)$$

$$\sum_{i \in \mathcal{S}^{[m]}} \vartheta_i + \sigma_0^2 |v_i|^2 \leq \gamma_{D,i}. \quad (3.25d)$$

相似地，基于从问题 (3.25) 获取的设备收发机  $\{v_i, b_i\}$ ，优化基站收发机  $\{\mathbf{w}, \mathbf{z}\}$  的子问题如下所述：

$$\min_{\mathbf{w}, \mathbf{z}, \rho_i, \vartheta_i, \alpha_i, \beta_i} \sum_{i \in \mathcal{S}^{[k]}} \rho_i + \sigma_1^2 \|\mathbf{z}\|^2, \quad (3.26)$$

$$\text{s.t.} \quad (3.14h), (3.22b), (3.22c), (3.24c), (3.25b) - (3.25d).$$

通过交替求解这两个子问题，可以得到问题 (3.24) 的次优解。注意，若得到的目标值小于或等于基站所要求的 MSE  $\gamma_B$ ，则可以选择集合  $\mathcal{S}^{[m]}$  作为可行设备集，并通过设计的收发机计算相应的聚合误差和广播误差。否则，令  $m = m - 1$  重置所选设备集。综上，联合设备选择和收发机设计的步骤总结为算法 2。

### 3.3.3 算法分析

**收敛性分析：** 对于稀疏性诱导问题 (3.14)，定义其在第  $t$  轮迭代处的目标函数为  $F_1(v_k^{[t]}, b_k^{[t]}, \mathbf{x}^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]})$ 。在算法 2 的步骤 3 中，基于给定的基站收发机，通过求解凸的子



**算法 2**: 联合设备选择和收发机设计算法

---

输入:  $N, K, \sigma_0^2, \sigma_1^2, P_{\max}, \gamma_B, \gamma_{D,k}, P_k, \varepsilon_k, \delta_k, \forall k \in \mathcal{K}$   
 输出:  $\mathcal{S}, \text{MSE}_B, \text{MSE}_i^D, \forall i \in \mathcal{S}$

- 1: 初始化  $\mathbf{z}^{[0]}, \mathbf{w}^{[0]}$ , 收敛精度  $\nu_1$ , 迭代索引  $t = 1$ ;
- 2: **repeat**
- 3: 通过固定  $\{\mathbf{w}^{[t-1]}, \mathbf{z}^{[t-1]}\}$ , 解决问题 (3.22), 得到  $\{v_k^{[t]}, b_k^{[t]}\}$ ;
- 4: 通过固定  $\{v_k^{[t]}, b_k^{[t]}\}$ , 解决问题 (3.23), 得到  $\{\mathbf{w}^{[t]}, \mathbf{z}^{[t]}, \boldsymbol{\chi}^{[t]}\}$ ;
- 5: 计算  $\Delta_1 = |||\boldsymbol{\chi}^{[t]}||_1 - |||\boldsymbol{\chi}^{[t-1]}||_1|$ ;
- 6: 更新  $t = t + 1$ ;
- 7: **until**  $\Delta_1 < \nu_1$
- 8: 得到  $\boldsymbol{\chi}^* = \boldsymbol{\chi}^{[t]}$ , 并按升序排列元素, 即  $\chi_{\pi(1)} \leq \chi_{\pi(2)} \leq \dots \leq \chi_{\pi(K)}$ ;
- 9: 初始化 所选设备集的数量  $m = K$ , 收敛精度  $\nu_2$ ;
- 10: **while**  $m \geq 0$  **do**
- 11: 设置选定的设备集为  $\mathcal{S}^{[m]} = \{\pi(1), \pi(2), \dots, \pi(m)\}$ ;
- 12: **repeat**
- 13: 初始化  $\mathbf{z}^{[0]}, \mathbf{w}^{[0]}$ , 迭代索引  $t = 1$ ;
- 14: 通过固定  $\{\mathbf{w}^{[t-1]}, \mathbf{z}^{[t-1]}\}$ , 解决问题 (3.25), 得到  $\{v_k^{[t]}, b_k^{[t]}\}$ ;
- 15: 通过固定  $\{v_k^{[t]}, b_k^{[t]}\}$ , 解决问题 (3.26), 得到  $\{\mathbf{w}^{[t]}, \mathbf{z}^{[t]}\}$ ;
- 16: 计算  $\Delta_2 = |\text{MSE}_B^{[t]} - \text{MSE}_B^{[t-1]}|$ ;
- 17:  $t = t + 1$ ;
- 18: **until**  $\Delta_2 < \nu_2$
- 19: **if**  $\text{MSE}_B^{[t]} \leq \gamma$  **then**
- 20:  $\mathbf{z}^* = \mathbf{z}^{[t]}, \mathbf{w}^* = \mathbf{w}^{[t]}, v_i^* = v_i^{[t]}, b_i^* = b_i^{[t]}, \forall i \in \mathcal{S}^{[m]}$
- 21: **break**
- 22: **end if**
- 23: 令  $m = m - 1$ ;
- 24: **end while**
- 25: 获得  $\mathcal{S}^{[m]}$ , 并根据  $\mathbf{z}^*, \mathbf{w}^*, v_i^*, b_i^*, \forall i \in \mathcal{S}^{[m]}$  计算  $\text{MSE}_B$  和  $\text{MSE}_i^D$ .

---

问题 (3.22), 可以获得最优的设备收发机。因此, 可以得到

$$F_1(v_k^{[t]}, b_k^{[t]}, \mathbf{x}^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]}) \geq F_1(v_k^{[t+1]}, b_k^{[t+1]}, \mathbf{x}^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]}). \quad (3.27)$$

然后, 在算法 2 的步骤 4 中, 基于给定的设备收发机, 通过求解凸子问题 (3.23), 可以获得最优的基站收发机和设备被选择的优先级。那么, 有

$$F_1(v_k^{[t+1]}, b_k^{[t+1]}, \mathbf{x}^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]}) \geq F_1(v_k^{[t+1]}, b_k^{[t+1]}, \mathbf{x}^{[t+1]}, \mathbf{w}^{[t+1]}, \mathbf{z}^{[t+1]}). \quad (3.28)$$

结合式 (3.27) 和式 (3.28), 可以得出

$$F_1(v_k^{[t]}, b_k^{[t]}, \mathbf{x}^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]}) \geq F_1(v_k^{[t+1]}, b_k^{[t+1]}, \mathbf{x}^{[t+1]}, \mathbf{w}^{[t+1]}, \mathbf{z}^{[t+1]}). \quad (3.29)$$

可以看出，算法 2 求解稀疏性诱导问题的每次迭代后（即步骤 1-8），目标函数  $F_1$  的值是非增的。与此同时，问题 (3.15) 的目标函数值有一个下界，因此算法 2 求解稀疏性诱导问题的收敛性可以得到保证。对于可行性检测问题 (3.24)，基于一个给定的设备集，定义其在第  $t$  轮迭代时的目标函数为  $F_2(v_i^{[t]}, b_i^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]})$ 。相似地，由于子问题 (3.25) 和 (3.26) 都是凸的，可以得到  $F_2(v_k^{[t]}, b_k^{[t]}, \mathbf{w}^{[t]}, \mathbf{z}^{[t]}) \geq F_2(v_k^{[t+1]}, b_k^{[t+1]}, \mathbf{w}^{[t+1]}, \mathbf{z}^{[t+1]})$ 。因此，用于解决可行性检测问题的算法同样是收敛的。此外，图 3.5 中通过仿真验证给出了各种 MSE 要求下所需的迭代次数。可以看出，算法 2 在不同条件下的收敛速度均较快，这也验证了算法 2 的收敛性。

**复杂度分析:** 如上所述，算法 2 主要分为两个阶段，这两个阶段具有相似的迭代结构。因此，仅需要讨论每次迭代的算法复杂度。可以看出，解决稀疏性诱导问题的计算复杂度主要来自步骤 4，即求解联合凸问题 (3.23)。一般地，可以通过使用标准的内点法求解该问题来衡量其算法复杂度<sup>[111]</sup>。具体而言，问题 (3.23) 有  $2K$  个维度为  $N+2$  的线性矩阵不等式 (Linear Matrix Inequality, LMI) 约束， $6K$  个维度为 1 的 LMI 约束，2 个维度为  $N$  的二阶锥 (Second Order Cone, SOC) 约束。因此，对于给定精度  $\Theta > 0$  的解，使用内点法求解问题 (3.23) 每次迭代所需的复杂度为  $\ln \frac{1}{\Theta} \sqrt{10K + 2KN + 4} \cdot n \cdot [6K + 2K(N+2)^3 + 2N^2 + n(n + 6K + 2K(N+2)^2)]$ 。类似地，解决可行性检测问题的计算复杂度主要来自于步骤 9，即求解凸问题 (3.25)。具体地，问题 (3.25) 有  $2K$  个维度为  $N+2$  的 LMI 约束， $5K$  个维度为 1 的 LMI 约束和 1 个维度为  $N$  的 SOC 约束。因此，使用内点法求解问题 (3.25) 每次迭代所需的复杂度为  $\ln \frac{1}{\Theta} \sqrt{9K + 2KN + 2} \cdot n \cdot [5K + 2K(N+2)^3 + N^2 + n(n + 5K + 2K(N+2)^2)]$ ，其中决策变量  $n$  的数量级为  $\mathcal{O}(KN^2)$ 。

### 3.4 仿真结果

本节提供了一些仿真结果来验证所提算法的有效性和鲁棒性。在不失一般性的情况下，假设所有终端设备均随机分布在小区范围内，路径损耗采用 3GPP 标准中的模型，即  $PL_{\text{dB}} = 128.1 + 37.6 \log_{10}(l)$ <sup>[112]</sup>，其中  $l$  (km) 是基站与终端设备之间的距离。为了便于分析，假设所有设备都具有相同的发射功率预算  $P_k = P_0$  和相同的 MSE 要求  $\gamma_{D,k} = \gamma_D$ 。此外，用  $\text{SNR}_B = 10 \log_{10}(P_{\text{max}}/\sigma_0^2)$  和  $\text{SNR}_D = 10 \log_{10}(P_0/\sigma_1^2)$  分别表示基站和设备的发射信噪比（以 dB 为单位）。仿真结果由平均 50 个信道实现，除特别说明外，仿真参数均参照表 3.1 中的设置。下面将从通信性能和计算性能两部分展示数值仿真结果。

表 3.1 算法 2 的仿真参数设置

| 参数          | 数值  |
|-------------|---|
| 基站天线数       | $N = 48$                                  |
| 终端设备数       | $K = 20$                                  |
| 小区半径        | 100 m                                     |
| CSI 不确定性    | $\varepsilon_k = \varepsilon = 0.1$ [105] |
| 基站最大发射信噪比   | $\text{SNR}_B = 10$ dB                    |
| 设备最大发射信噪比   | $\text{SNR}_{D,k} = \text{SNR}_D = 10$ dB |
| 基站上的 MSE 要求 | $\gamma_B = 5$ dB                         |
| 设备上的 MSE 要求 | $\gamma_{D,k} = \gamma_D = 5$ dB          |

## 3.4.1 通信性能

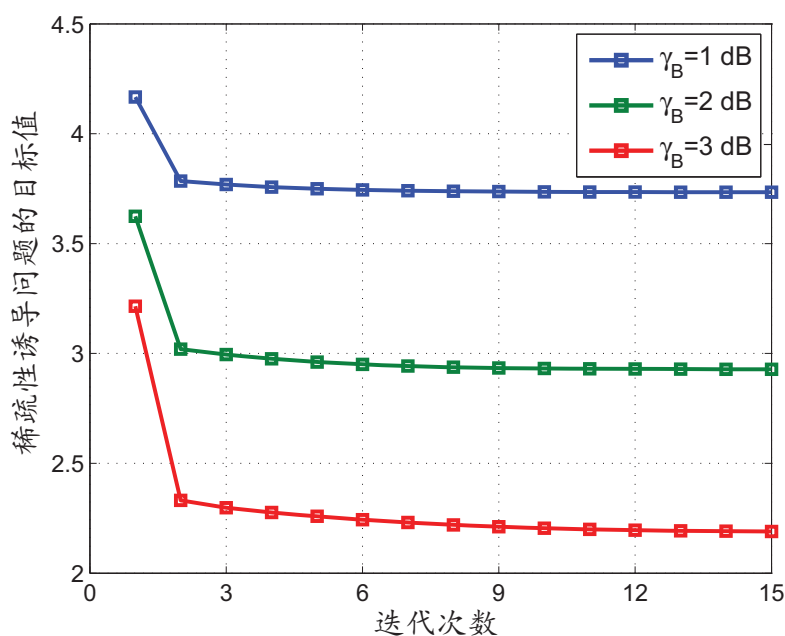


图 3.5 使用算法 2 求解稀疏诱导问题的收敛行为

首先，图 3.5 给出了所提鲁棒算法求解稀疏性诱导问题的收敛行为。可以看出，在不同的基站 MSE 要求  $\gamma_B$  下，目标值  $\|\chi\|_1$  均按预期单调减小，并在 10 次迭代内收敛到一个平稳值，这验证了所提算法具有较快的收敛速度。此外，可以看出  $\gamma_B$  的值越小，收敛速度越快，目标值越大。这是因为较小的  $\gamma_B$  表示更严格的基站 MSE 要求，导致了更低的稀疏性，从而更易于找到最小的目标值。

图 3.6 中展示了基站 MSE 要求  $\gamma_B$  和设备 MSE 要求  $\gamma_D$  对所提算法的设备选择性能的影响。可以发现，随着基站 MSE 要求  $\gamma_B$  的增加，所选设备的平均数量也随之增加。这是

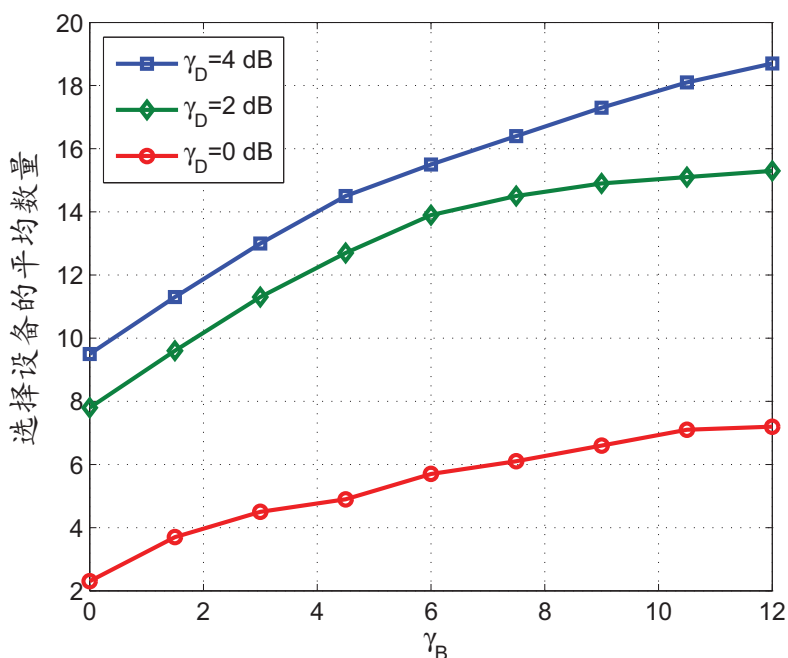


图 3.6 聚合误差和广播误差的要求对平均选择设备数量的影响

因为放松基站上聚合误差约束条件可以使更多的设备被选择去共同参与 FL。此外，通过放宽设备 MSE 要求  $\gamma_D$  也可以增加被选择设备的平均数量。但是，如前文所述，基站和设备的 MSE 约束宽松同时也意味着模型参数失真度的提高，这会导致 FL 的性能下降。因此，需要在合适的 MSE 要求下选择适当的设备数量来平衡 FL 的性能。

图 3.7 给出了不同算法在平均选择设备数量上的性能比较。其中，凸差 (Difference-of-Convex, DC) 非鲁棒算法是基于文献 [113] 和文献 [114] 中的算法组合而来的。具体地，设备端采用文献 [113] 中的迫零接收机用于下行传输和文献 [114] 中的迫零发射机用于上行传输，而基站侧的设备选择和收发机设计则是基于 DC 方法得到的<sup>[114]</sup>。AO 非鲁棒算法与本文提出的算法相似，均采用 AO 方法根据系统要求设计自适应的收发机，不同之处在于它没有考虑信道不确定性。可以看出，对于给定的基站 MSE 要求，随着设备要求 MSE 的放宽，使用所提鲁棒算法所选设备的平均数量也会随之增加，但设备数量的提升是有限的。此外，对于两种非鲁棒算法，其性能趋势似乎没有规律性。这是因为这两种算法在设计收发机时没有考虑信道的不确定性，有时只选择少数几个设备或者甚至不选择设备来确保满足 MSE 的约束条件。注意到，虽然对比算法的复杂度均小于所提算法，但是它们与所提算法的性能差距也较大。比较三种算法，可以得出算法 2 的性能最稳定，而 DC 非鲁棒算法性能最差，这充分证明了本章所提的算法 2 的有效性和鲁棒性。

图 3.8 研究了信道误差范围  $\varepsilon_0$  和基站天线数量  $N$  对所提算法 2 性能的影响。注意  $\varepsilon_0 = 0$  表示完美 CSI 的情况。结果表明，随着信道误差范围的增大，所选设备的平均数量

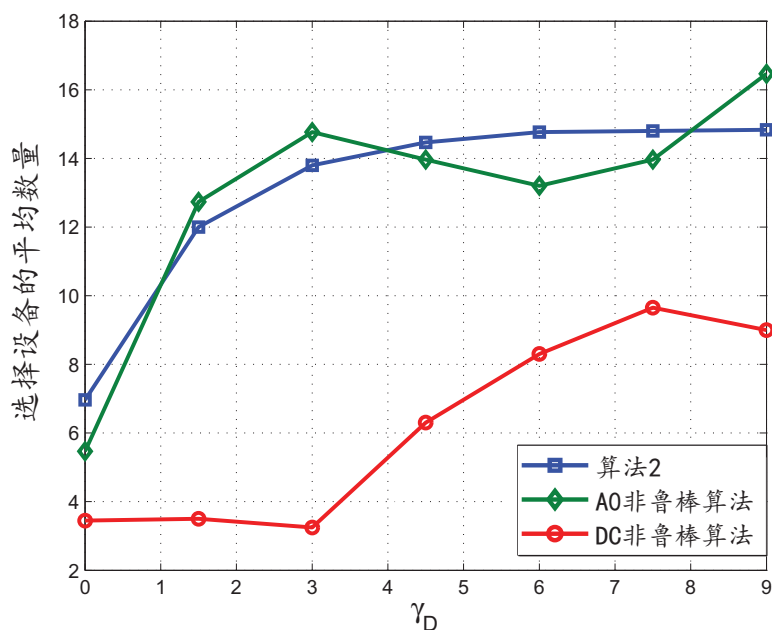
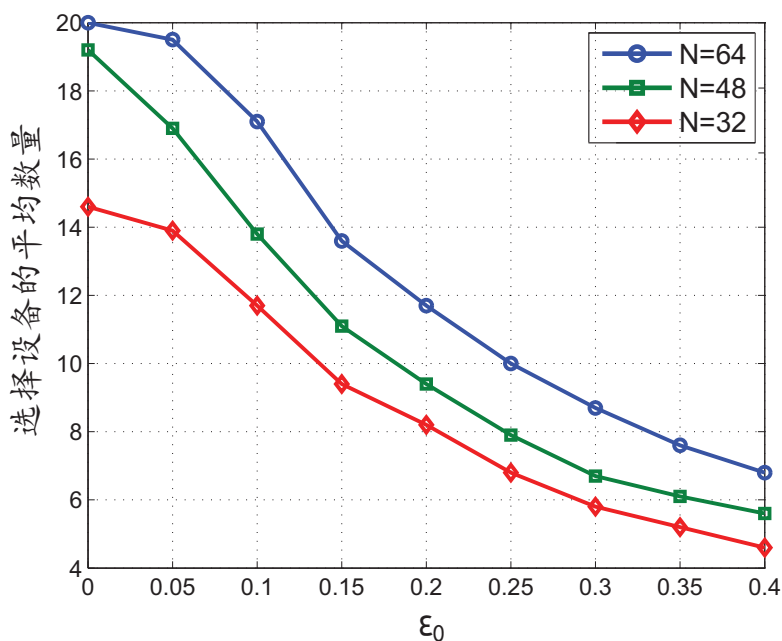
图 3.7 不同算法在不同  $\gamma_D$  要求下的平均选择设备数量对比

图 3.8 不同信道误差范围和不同天线数量下的平均选择设备数量

逐渐减少。这是因为较大的信道误差范围代表信道的不确定性更大，从而需要更多的资源来对抗信道的不确定性并在最差的情况下保证 MSE 的约束条件，因此只有较少的资源用来提高整体系统的性能。此外，可以看到，当增加基站天线数时，所提的鲁棒算法可以在满足同等 MSE 的要求下选择更多的设备参与 FL，从而提高 FL 的性能。尤其是，在基站天线数量为  $N = 64$  的情况下，所提的鲁棒算法在保证 MSE 要求的情况下选择所有设备参

与 FL。因此，通过适当地增加基站天线数量可以减小不利因素对系统的影响并提高系统的整体性能。

### 3.4.2 计算性能

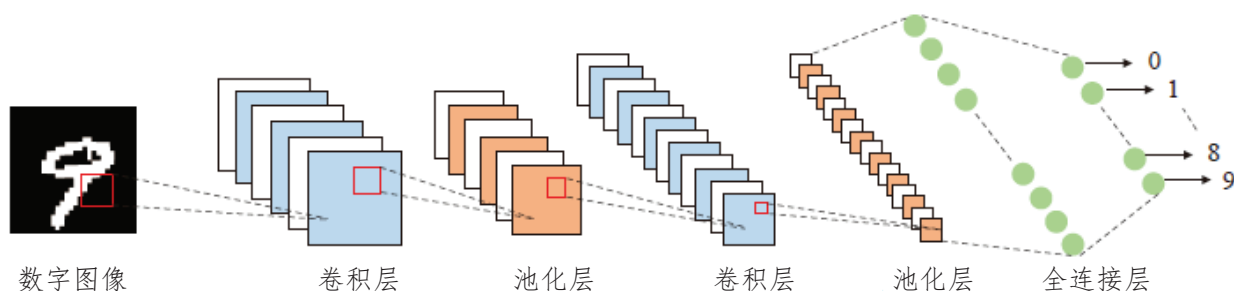


图 3.9 仿真实验中训练 FL 模型所用的 6 层 CNN 架构

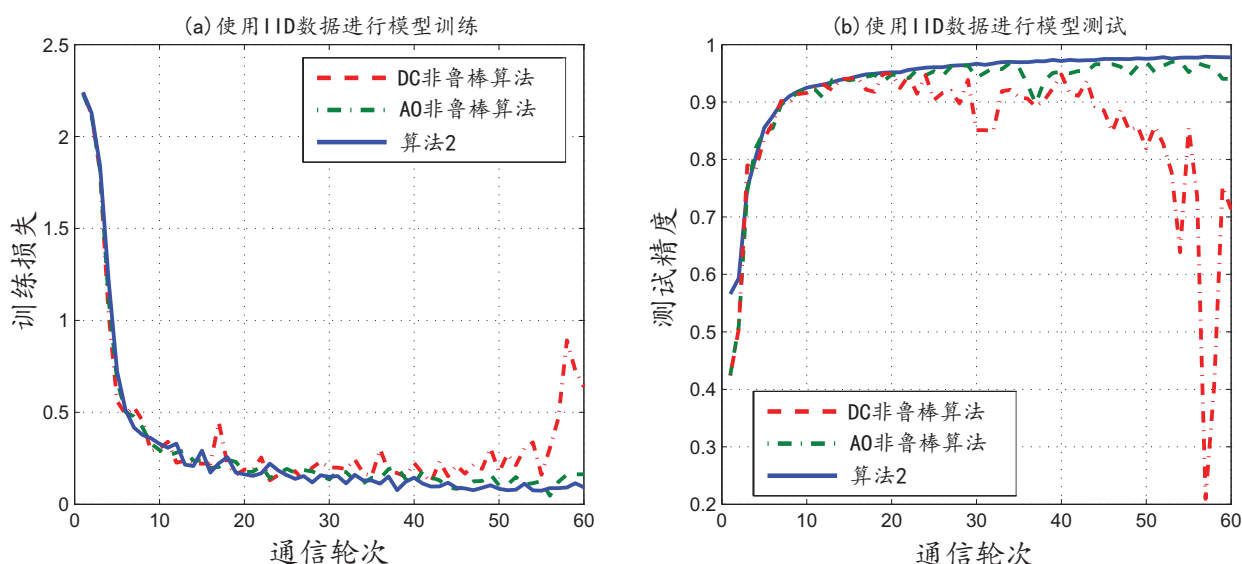


图 3.10 不同算法在IID MNIST数据集上的 FL 性能比较

为了证明算法 2 在实际 FL 应用中的性能，考虑一个常用的手写数字识别 FL 任务，并使用著名的 MNIST 数据集进行训练。该数据集由 10 个类别组成，范围从数字“0”到“9”，总共有 70,000 个标记的训练数据样本，其中 60,000 个样本用于训练，10,000 个样本用于测试性能。为了模拟分布式的移动数据，考虑两种不同类型的数据分布，即独立同分布（Independent Identically Distribution, IID）设置和非 IID 设置。对于 IID 设置，将训练样本随机划分为 20 个相等的子样本集，每个子样本集被分配给一个指定的设备。对于非 IID 设置，首先按照数字标签对数据进行排序，将其分成 40 个子样本集，每个子样本集有 1,500 张图片，并随机分配 2 个子样本集给每个设备。如图 3.9 所示，训练网络使用经典

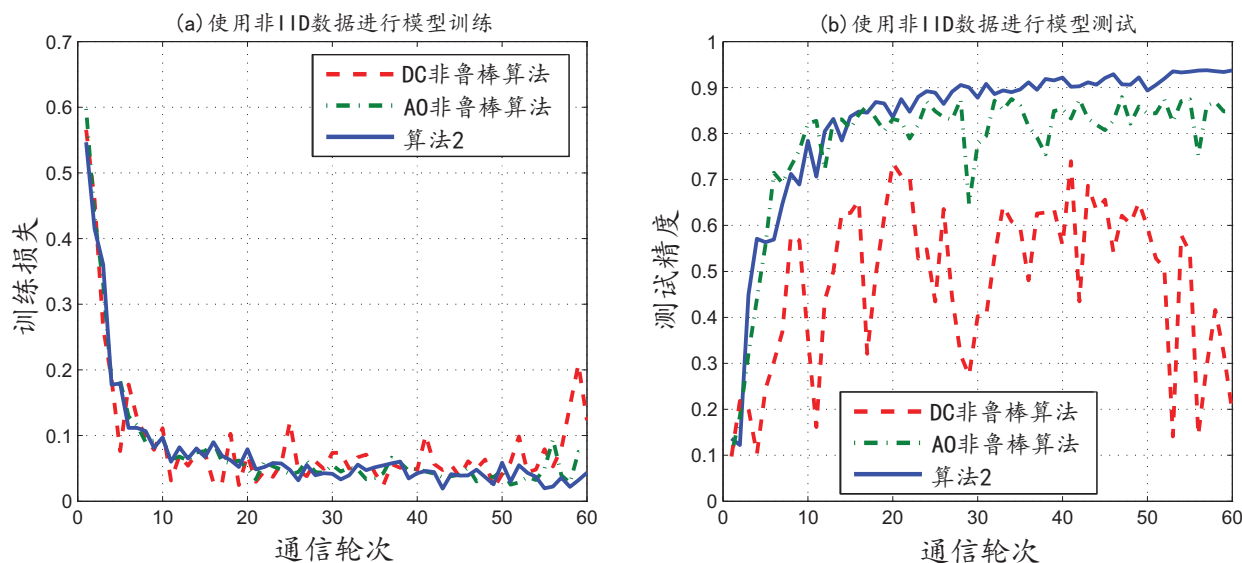
















图 3.11 不同算法在非IID MNIST数据集上的FL性能比较

FL 文献[115]中的 6 层的 CNN，该 CNN 由两个  $5 \times 5$  并使用线性整流单元（Rectified Linear Unit, ReLU）激活函数的卷积层（其中第一个卷积层有 32 个通道，第二个卷积层有 64 个通道），两个  $2 \times 2$  的最大池化层，一个具有 512 个神经元并使用 ReLU 激活函数的全连接层，以及一个使用 Softmax 激活函数作为输出层的全连接层组成，该网络的总参数量为 1,663,370 个。每个设备使用随机梯度下降法更新其本地模型，设置学习率为 0.01。使用交叉熵作为损失函数，批尺寸设为 64，迭代次数设为 1。

**注3.1.** 采取不同的学习任务、数据集和训练方法都将产生不同的预测结果。本节采用的基准 FL 算法由这样一个 6 层 CNN 在 MNIST 数据集上训练用于图像分类，在没有聚合误差和广播误差的理想情况下，测试准确率超过 95%，如图 3.2 所示。因此，使用该网络是足够公平的，可以用于验证算法 2 的性能。

图 3.10 和 3.11 分别在 IID 和非 IID MNIST 数据集比较了上述三种算法在训练损失和测试精度上的性能，其中  $\text{SNR}_B = \text{SNR}_D = 10 \text{ dB}$  和  $\gamma_D = \gamma_B = 5 \text{ dB}$ 。对于 IID 设置，可以看到所提算法 2 和 AO 非鲁棒算法都优于 DC 非鲁棒算法，特别是，算法 2 在降低训练损失和提高测试精度方面最为稳定且表现最佳。对于非 IID 设置，可以发现三种算法在性能上都存在振荡。这是因为 MSE 约束的存在，使得部分设备没有被选择参与 FL，导致一些数据样本没有被训练。例如，第  $k$  个设备被分配两个子样本集“0”和“9”。如果在多个通信轮中均很少选择第  $k$  个设备参与 FL，则数字样本“0”和“9”被训练的次数很少，从而导致数字样本“0”和“9”的识别精度较低。即使如此，可以从仿真结果看出，所

表 3.2 在 IID MNIST 数据集上测试手写数字识别的例子

|          |   |   |   |   |  |   |   |
|----------|---|---|---|---|--|---|---|
| 数字图像     |  |  |  |  |  |  |  |
| 算法 2     | 0   | 5   | 9   | 0   | 2  | 1   | 7   |
| AO 非鲁棒算法 | 0   | 6(×)  | 9   | 0   | 2  | 1   | 7   |
| DC 非鲁棒算法 | 0   | 4(×)  | 4(×)  | 0   | 4(×)   | 4(×)  | 4(×)  |
| 数字图像     |  |  |  |  |  |  |  |
| 算法 2     | 3   | 6   | 2   | 8   | 4  | 3   | 1   |
| AO 非鲁棒算法 | 3   | 6   | 2   | 8   | 4  | 3   | 7(×)  |
| DC 非鲁棒算法 | 4(×)  | 5(×)  | 4(×)  | 4(×)  | 4  | 4(×)  | 4(×)  |

提算法 2 在训练和测试上都表现最好也最稳定，并在非 IID 数据集上也实现了较高的测试精度。为了更生动地展示结果，表 3.2 给出了一些手写数字识别的例子。

### 3.5 本章小结

本章研究了面向感知的通信和计算一体化设计。为了减少通信过程中模型广播和模型聚合的失真，从而提高计算过程中 FL 的性能，提出了一种基于空中计算的模型参数传输方案。该方案构建为一个计算复杂的非凸优化问题，通过优化设备选择和收发机设计来最大限度地选择参与 FL 的设备数量，同时确保模型广播和聚合的 MSE 约束。为了解决这一非凸优化问题，先利用优化分解的方法将原问题分解为稀疏性诱导子问题和可行性检测子问题，然后提出了一种基于迭代优化的鲁棒算法，获得原问题的可行解。仿真结果验证了所提算法对通信和计算一体化系统的有效性和鲁棒性。



## 4 计算辅助的通信和感知一体化

### 4.1 引言

通信和感知一体化的设计原理是将通信和感知功能集成在同一平台。然而，在边缘智能网络中实现通信和感知一体化非常具有挑战性。特别是，由于通信和感知共享部分资源或者使用同一资源，系统可能会遭受严重的功能间干扰，从而导致性能下降。因此，如何有效地协调通信和感知之间的相互干扰已成为通信和感知一体化的关键问题之一。传统上，通常将不同的正交无线电资源分别分配给感知和通信功能以避免功能间的干扰。虽然整个传输过程是无干扰的，但正交使能的通信和感知一体化会导致系统频谱效率低下。因此，需要研究非正交的通信和感知一体化系统来提高频谱利用率。然而，边缘智能网络面临着复杂的无线传输环境，即多功能基站在感知周围多个目标的同时需要服务多个通信终端。换句话说，基站接收到的是来自多个通信终端的数据信号和经多个感知目标反射的回波信号的相互叠加，这给联合信号处理带来了极大的挑战。在这种情况下，传统的信号处理方法可能无法直接应用在通信和感知一体化系统中。为了解决这个问题，可以借助 AI 强大的计算能力，通过提取叠加信号中的有用成分，实现高效的联合信号处理。基于此，针对资源共享导致的通信信号和感知信号的混叠问题，本章提出了一种低复杂的基于 DL 的联合感知发射波形和通信接收波束设计算法来提高通信和感知一体化的整体性能。

内容概述：本章提出了一种计算辅助的通信和感知一体化架构，利用多功能基站在同一频谱和相同硬件上同时实现通信和感知功能，有效提高了频谱和硬件的利用效率。针对资源共享导致的通信信号和感知信号的混叠问题，借助 AI 强大的计算能力，提出了一种基于 DL 的联合感知发射波形和通信接收波束设计算法来提高通信和感知一体化的性能。具体结构安排如下：4.2 节介绍了计算辅助的通信和感知一体化模型，4.3 节提出了联合感知波形和通信波束设计的优化问题，并进行了等效问题转换以降低设计复杂度，4.4 节提出了一种基于 DL 的算法，4.5 节利用数值仿真验证了所提 DL 算法的有效性，4.6 节总结了本章内容。

## 4.2 系统模型

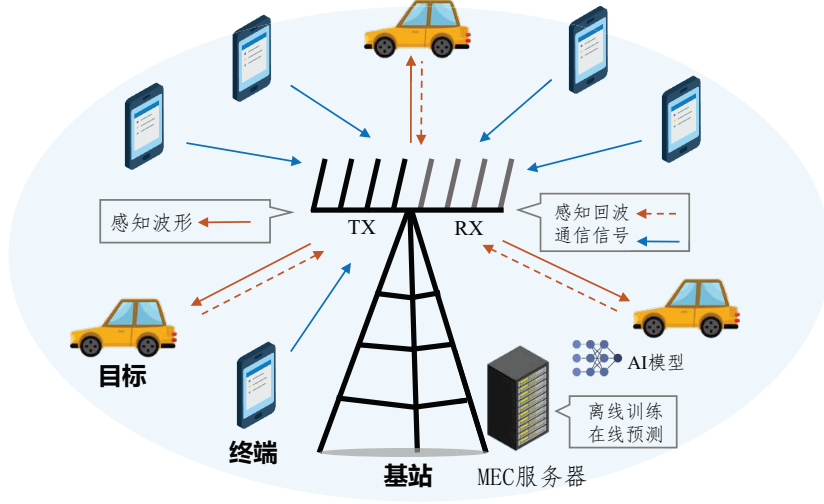


图 4.1 计算辅助的通信和感知一体化系统模型

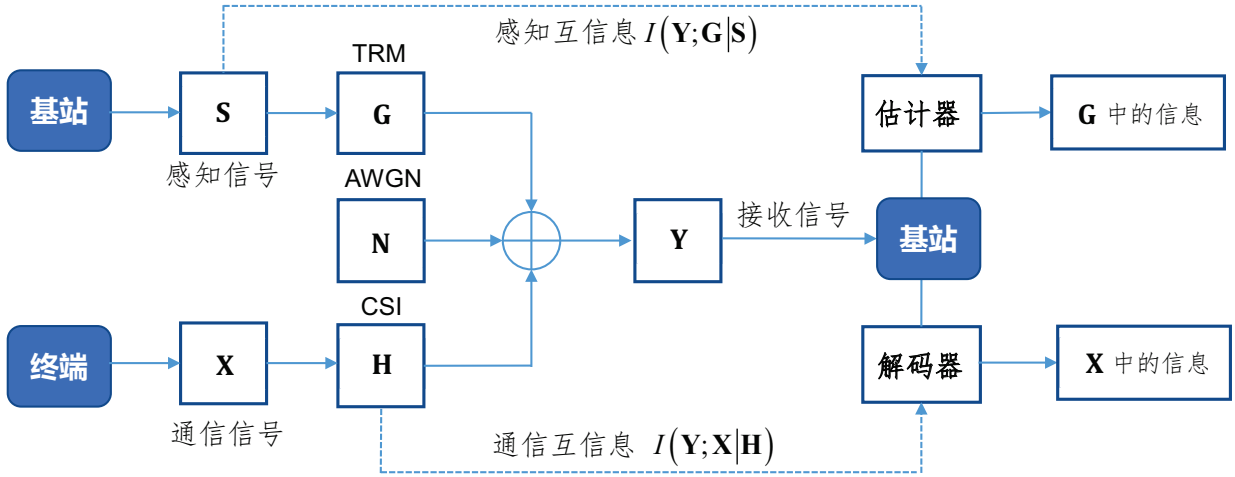


图 4.2 通信和感知一体化信号流图

如图 4.1 所示，考虑在边缘智能网络中部署一个配有  $N_t$  根发射天线和  $N_r$  根接收天线的双功能基站，服务  $K$  个单天线终端的同时感知周围多个目标。具体而言，双功能基站首先广播感知信号  $\mathbf{S} = [\mathbf{s}_1^T, \dots, \mathbf{s}_L^T] \in \mathbb{C}^{L \times N_t}$  用于感知周围环境，其中  $L$  为感知信号波形的长度并满足  $L > N_t$ 。接着，双功能基站同时接收到多个终端发射的通信信号和经多个目标反射的回波信号。具体地，基站接收到的混合信号  $\mathbf{Y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_L^T] \in \mathbb{C}^{L \times N_r}$  可以表示为

$$\mathbf{Y} = \mathbf{S}\mathbf{G} + \mathbf{X}\mathbf{P}\mathbf{H}^T + \mathbf{N}, \quad (4.1)$$

其中  $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_K] \in \mathbb{C}^{N_r \times K}$  是信道矩阵， $\mathbf{h}_k \in \mathbb{C}^{N_r \times 1}$  表示基站和第  $k$  个终端之间的

信道矢量； $\mathbf{P} = \text{diag}[\sqrt{p_1}, \dots, \sqrt{p_K}] \in \mathbb{R}^{K \times K}$  是功率矩阵， $p_k$  表示第  $k$  个终端的发射功率； $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_K] \in \mathbb{C}^{L \times K}$  是通信信号矩阵， $\mathbf{x}_k \in \mathbb{C}^{L \times 1}$  表示第  $k$  个终端发射的通信信号； $\mathbf{N} = [\mathbf{n}_1, \dots, \mathbf{n}_L] \in \mathbb{C}^{N_r \times L}$  是 AWGN 矩阵，其向量元素服从均值为零、方差为  $\sigma_n^2$  的高斯分布，即  $\mathbf{n}_l \sim \mathcal{CN}(0, \sigma_n^2 \mathbf{I})$ 。此外， $\mathbf{G} \in \mathbb{C}^{N_t \times N_r}$  是需要被感知的目标响应矩阵 (Target Response Matrix, TRM)，它可以表示为

$$\mathbf{G} = \sum_i \beta_i \mathbf{a}(\theta_i) \mathbf{b}^T(\theta_i), \quad (4.2)$$

其中  $\beta_i$  和  $\theta_i$  分别是第  $i$  个目标的反射系数和到达角。 $\mathbf{a}(\theta_i)$  和  $\mathbf{b}(\theta_i)$  分别是相应的发射和接收阵列导向矢量。最后，基站将接收到的信号分别送入估计器和解码器进行感知信息提取和通信信号解码。为了更形象地描述这个过程，图 4.2 展示了详细的系统流程。

**注4.1.** 对于基站上分布较广的天线，TRM 的列相关性间的差异是可以忽略的。在这种情况下，为了便于分析，假设矩阵  $\mathbf{G}$  的每一列的向量都具有相同的相关性，即  $\mathbf{R}_T = \mathbb{E} \{ \mathbf{g}_i \mathbf{g}_i^H \}$ ,  $i = 1, \dots, N_r$ ，其中  $\mathbf{g}_i \in \mathbb{C}^{N_t \times 1}$  表示  $\mathbf{G}$  的第  $i$  列向量<sup>[116]</sup>。

#### 4.2.1 感知信息提取

接收到信号后，基站恢复出周围环境中的 TRM  $\mathbf{G}$ 。通过从  $\mathbf{G}$  中提取感知信息，便能得到详细的环境参数，如各个目标的反射系数和到达角。考虑到估计 TRM 的 MSE 最小化等价于感知速率的最大化<sup>[116, 117]</sup>，可以选择感知速率作为感知信息提取的性能指标。为了推导感知速率，即单位时间内感知的信息量，需要用到以下引理：

**引理4.1.** 对于任意矩阵  $\mathbf{A} \in \mathbb{C}^{m \times n}$ ,  $\mathbf{B} \in \mathbb{C}^{n \times p}$  和  $\mathbf{C} \in \mathbb{C}^{p \times q}$ ，可以推出  $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B})$ 。

*证明:* 请参考附录 A.2.1。 □

基于引理 4.1，通过向量化式 (4.1) 中的接收信号  $\mathbf{Y}$ ，可以得到

$$\tilde{\mathbf{y}} = \tilde{\mathbf{S}}\tilde{\mathbf{g}} + (\mathbf{HP}^T \otimes \mathbf{I}_L) \text{vec}(\mathbf{X}) + \tilde{\mathbf{n}}, \quad (4.3)$$

其中  $\tilde{\mathbf{y}} = \text{vec}(\mathbf{Y}) \in \mathbb{C}^{LN_r \times 1}$ ,  $\tilde{\mathbf{S}} = \mathbf{I}_{N_r} \otimes \mathbf{S} \in \mathbb{C}^{N_r L \times N_r N_t}$ ,  $\tilde{\mathbf{g}} = \text{vec}(\mathbf{G}) \in \mathbb{C}^{N_t N_r \times 1}$ ,  $\tilde{\mathbf{n}} = \text{vec}(\mathbf{N}) \in \mathbb{C}^{LN_r \times 1}$ 。基于此，对于已知的感知信号波形  $\tilde{\mathbf{S}}$ ，向量化的 TRM  $\tilde{\mathbf{g}}$  和向量化的接

收信号  $\tilde{\mathbf{y}}$  之间的感知互信息可以表示为

$$\begin{aligned} I(\tilde{\mathbf{y}}; \tilde{\mathbf{g}} | \tilde{\mathbf{S}}) &= h(\tilde{\mathbf{y}} | \tilde{\mathbf{S}}) - h(\tilde{\mathbf{y}} | \tilde{\mathbf{g}}, \tilde{\mathbf{S}}) \\ &= \log_2 \left| \tilde{\mathbf{S}} \mathbf{R}_g \tilde{\mathbf{S}}^H + \mathbf{R}_I \right| - \log_2 |\mathbf{R}_I| \\ &= \log_2 \left| \mathbf{I}_{LN_r} + \mathbf{R}_I^{-1} \tilde{\mathbf{S}} \mathbf{R}_g \tilde{\mathbf{S}}^H \right|, \end{aligned} \quad (4.4)$$

其中  $h(\mathbf{y}|\mathbf{x}) = \int \varrho(\mathbf{y}|\mathbf{x}) \log_2 \varrho(\mathbf{y}|\mathbf{x}) \mathbf{d}\mathbf{y}$  和  $\varrho(\mathbf{y}|\mathbf{x})$  分别是给定  $\mathbf{x}$  情况下  $\mathbf{y}$  的条件微分熵和条件概率密度函数。此外， $\mathbf{R}_I = (\mathbf{H}\tilde{\mathbf{P}}\mathbf{H}^H) \otimes \mathbf{I}_L + \sigma_n^2 \mathbf{I}_{LN_r}$  表示通信干扰加噪声的协方差矩阵， $\mathbf{R}_g = \mathbf{I}_{N_r} \otimes \mathbf{R}_T$  表示  $\tilde{\mathbf{g}}$  的协方差矩阵，其中  $\tilde{\mathbf{P}} = \text{diag}[p_1, \dots, p_K] \in \mathbb{R}^{K \times K}$ 。因此，感知速率可以表示为

$$\mathcal{R}_s(\mathbf{S}) = I(\tilde{\mathbf{y}}; \tilde{\mathbf{g}} | \tilde{\mathbf{S}}) / L. \quad (4.5)$$

#### 4.2.2 通信信号解码

与此同时，基站上的解码器对通信信号进行解码。为了评估通信信号的质量，选择常用的通信速率作为通信信号解码的性能指标。具体地，在第  $l$  个时隙，基站处接收到的信号为

$$\mathbf{y}_l^T = \sum_{k=1}^K \sqrt{p_k} x_{k,l} \mathbf{h}_k^T + \mathbf{s}_l^T \mathbf{G} + \mathbf{n}_l^T, \quad (4.6)$$

其中  $x_{k,l}$  指第  $k$  个终端在第  $l$  个时隙发射的通信信号。为了减少干扰，在基站侧对来自终端的通信信号采用接收波束成形技术。因此， $x_{k,l}$  在基站侧解码后可以表示为

$$\hat{x}_{k,l} = \sqrt{p_k} x_{k,l} \mathbf{h}_k^T \mathbf{w}_k + \sum_{i \neq k} \sqrt{p_i} x_{i,l} \mathbf{h}_i^T \mathbf{w}_k + \mathbf{s}_l^T \mathbf{G} \mathbf{w}_k + \mathbf{n}_l^T \mathbf{w}_k,$$

其中  $\mathbf{w}_k \in \mathbb{R}^{N_r \times 1}$  是基站侧关于第  $k$  个终端的通信信号的接收波束。基于此，接收的通信信号的信干噪比（Signal to Interference plus Noise Ratio, SINR）可以表示为

$$\gamma_k = \frac{p_k |\mathbf{h}_k^T \mathbf{w}_k|^2}{\sum_{i \neq k} p_i |\mathbf{h}_i^T \mathbf{w}_k|^2 + (\|\mathbf{S} \mathbf{R}_T \mathbf{S}^H\| / L + \sigma_n^2) \|\mathbf{w}_k\|^2}, \forall k, \quad (4.7)$$

以及所有终端的平均通信速率可以表示为

$$\mathcal{R}_c(\mathbf{w}_k, \mathbf{S}) = \frac{1}{K} \sum_{k=1}^K \log_2(1 + \gamma_k). \quad (4.8)$$

从式 (4.5) 和 (4.8) 可以看出, 感知速率和通信速率受到感知发射波形  $\mathbf{S}$  和通信接收波束  $\mathbf{w}_k, \forall k$  的共同影响。因此, 有必要联合设计发射感知波形和接收通信波束来提高边缘智能网络中通信和感知一体化的整体性能。

### 4.3 联合感知波形和通信波束设计

在本节中, 考虑到资源有限性和功能公平性, 以归一化的感知速率和归一化的通信速率的加权和最大化为目标进行联合感知波形和通信波束设计。该设计可以表述为以下多目标优化问题 (Multi-Objective Optimization Problem, MOOP) :

$$S\&C\text{-MOOP: } \max_{\mathbf{S}, \mathbf{w}_k} \quad \frac{\alpha}{M_s} \mathcal{R}_s(\mathbf{S}) + \frac{(1-\alpha)}{M_c} \mathcal{R}_c(\mathbf{w}_k, \mathbf{S}) \quad (4.9a)$$

$$\text{s.t.} \quad \text{tr}(\mathbf{S}\mathbf{S}^H) \leq P_s, \quad (4.9b)$$

其中  $0 \leq \alpha \leq 1$  是感知功能的性能权重,  $P_s$  是感知信号发射功率的最大预算,  $M_s$  和  $M_c$  是用于归一化的系数, 分别为最大可达感知速率和最大可达通信速率。显然, 由于目标函数 (4.9a) 是计算困难且非凸的, 很难直接求解问题 (4.9)。因此, 需要对该问题进行转换以降低设计复杂度。具体而言, 首先将这个 MOOP 划分为两个单目标优化问题 (Single Objective Optimization Problem, SOOP), 即感知速率最大化问题和通信速率最大化问题。接着, 分别处理和转换这两个 SOOP。最后, 重新组合两个转换后的 SOOP, 以获得等效但更易于解决的 MOOP。接下来详细描述问题转换的过程。

#### 4.3.1 感知速率最大化问题

忽略通信性能时, 感知速率最大化的 SOOP 可以表示为

$$S\text{-SOOP: } \max_{\mathbf{S}} \quad \mathcal{R}_s(\mathbf{S}), \quad (4.10)$$

$$\text{s.t.} \quad (4.9b).$$

由式 (4.5) 中的  $\mathcal{R}_s(\mathbf{S})$  可知, 感知速率最大化等同于感知互信息最大化。根据克罗内克积的

性质，感知互信息可以重新表示为

$$\begin{aligned}
 I(\tilde{\mathbf{y}}; \tilde{\mathbf{g}} | \tilde{\mathbf{S}}) &= \log_2 \left| \mathbf{I}_{LN_r} + \mathbf{R}_I^{-1} \tilde{\mathbf{S}} \mathbf{R}_g \tilde{\mathbf{S}}^H \right| \\
 &= \log_2 \left| \mathbf{I}_{LN_r} + (\mathbf{R}_H \otimes \mathbf{I}_L) (\mathbf{I}_{N_r} \otimes \mathbf{S} \mathbf{R}_T \mathbf{S}^H) \right| \\
 &= \log_2 \left| \mathbf{I}_{LN_r} + \mathbf{R}_H \otimes \mathbf{S} \mathbf{R}_T \mathbf{S}^H \right|,
 \end{aligned} \tag{4.11}$$

其中  $\mathbf{R}_H = (\mathbf{H} \tilde{\mathbf{P}} \mathbf{H}^H + \sigma_n^2 \mathbf{I}_{N_r})^{-1}$ 。为了减少计算维度，对  $\mathbf{S}$  进行奇异值分解（Singular Value Decomposition, SVD）得到  $\mathbf{S} = \mathbf{U}_s \boldsymbol{\Sigma}_s \mathbf{V}_s^H$ ，其中  $\mathbf{U}_s$  和  $\mathbf{V}_s$  是酉矩阵， $\boldsymbol{\Sigma}_s$  是对角矩阵。 $\boldsymbol{\Sigma}_s$  可以表示为  $\boldsymbol{\Sigma}_s = \left[ (\boldsymbol{\Sigma}_s^\downarrow)^{1/2}, \mathbf{0}_{N_t \times (L-N_t)} \right]^T$ ，其中  $\boldsymbol{\Sigma}_s^\downarrow = \text{diag}([\sigma_{s,1}, \dots, \sigma_{s,N_t}]^T)$ ， $\sqrt{\sigma_{s,k}}$  是  $\mathbf{S}$  的第  $k$  个奇异值，且有  $\sigma_{s,1} \geq \sigma_{s,2} \geq \dots \geq \sigma_{s,N_t}$ 。同时，对  $\mathbf{R}_T$  和  $\mathbf{R}_H$  进行特征值分解（Eigenvalue Decomposition, EVD），可以得到  $\mathbf{R}_T = \mathbf{U}_T \boldsymbol{\Sigma}_T \mathbf{U}_T^H$ ，其中  $\mathbf{U}_T$  是正交矩阵， $\boldsymbol{\Sigma}_T = \text{diag}([\sigma_{t,1}, \dots, \sigma_{t,N_t}]^T)$  是对角矩阵， $\sigma_{t,i}$  是  $\mathbf{R}_T$  的第  $i$  个特征值，且有  $\sigma_{t,1} \geq \sigma_{t,2} \geq \dots \geq \sigma_{t,N_t}$ 。此外， $\mathbf{R}_H$  可以分解为  $\mathbf{R}_H = \mathbf{U}_h \boldsymbol{\Sigma}_h \mathbf{U}_h^{-1}$ ，其中  $\mathbf{U}_h$  和  $\boldsymbol{\Sigma}_h = \text{diag}([\sigma_{h,1}, \dots, \sigma_{h,N_r}]^T)$  是相应的方块矩阵和对角矩阵， $\sigma_{h,j}$  是  $\mathbf{R}_H$  的第  $j$  个特征值，且有  $\sigma_{h,1} \geq \sigma_{h,2} \geq \dots \geq \sigma_{h,N_r}$ 。

因此，感知互信息可以被转化为

$$I(\tilde{\mathbf{y}}; \tilde{\mathbf{g}} | \tilde{\mathbf{S}}) = \log_2 \left| \mathbf{I}_{LN_r} + \mathbf{U}_h \boldsymbol{\Sigma}_h \mathbf{U}_h^{-1} \otimes \mathbf{U}_s \boldsymbol{\Sigma}_s \mathbf{V}_s^H \mathbf{U}_T \boldsymbol{\Sigma}_T \mathbf{U}_T^H \mathbf{V}_s \boldsymbol{\Sigma}_s^H \mathbf{U}_s^H \right|. \tag{4.12}$$

为了进一步处理式 (4.12) 中的感知互信息，提出以下引理：

**引理4.2.** 对于任意方块矩阵  $\mathbf{A} \in \mathbb{C}^{m \times m}$ ,  $\mathbf{B} \in \mathbb{C}^{m \times m}$ ,  $\mathbf{C} \in \mathbb{C}^{n \times n}$  和  $\mathbf{D} \in \mathbb{C}^{n \times n}$ ，等式  $|\mathbf{I}_{mn} + \mathbf{A} \mathbf{B} \otimes \mathbf{C} \mathbf{D}| = |\mathbf{I}_{mn} + \mathbf{B} \mathbf{A} \otimes \mathbf{D} \mathbf{C}|$  成立。

证明：请参阅附录 A.2.2。 □

基于引理 4.2，感知互信息可以进一步转化为

$$\begin{aligned}
 I(\tilde{\mathbf{y}}; \tilde{\mathbf{g}} | \tilde{\mathbf{S}}) &= \log_2 \left| \mathbf{I}_{LN_r} + \boldsymbol{\Sigma}_h \mathbf{U}_h^{-1} \mathbf{U}_h \otimes \boldsymbol{\Sigma}_s \mathbf{V}_s^H \mathbf{U}_T \boldsymbol{\Sigma}_T \mathbf{U}_T^H \mathbf{V}_s \boldsymbol{\Sigma}_s^H \mathbf{U}_s^H \mathbf{U}_s \right| \\
 &= \log_2 \left| \mathbf{I}_{LN_r} + \boldsymbol{\Sigma}_h \otimes \boldsymbol{\Sigma}_s (\mathbf{V}_s^H \mathbf{U}_T) \boldsymbol{\Sigma}_T (\mathbf{U}_T^H \mathbf{V}_s) \boldsymbol{\Sigma}_s^H \right| \\
 &\leq \log_2 \left| \mathbf{I}_{LN_r} + \boldsymbol{\Sigma}_h \otimes \boldsymbol{\Lambda} \right|,
 \end{aligned} \tag{4.13}$$

其中,  $\Lambda = \Sigma_s \Sigma_T \Sigma_s^H = \begin{bmatrix} \Sigma_s^\dagger \Sigma_T & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{R}^{L \times L}$ 。当且仅当  $\mathbf{V}_s = \mathbf{U}_T$  时, 可以达到式 (4.13) 中的上限<sup>[116]</sup>。因此, 令  $\mathbf{V}_s = \mathbf{U}_T$ , 感知速率可以表示为

$$\bar{\mathcal{R}}_s(\boldsymbol{\sigma}_s) = \frac{1}{L} \sum_{i=1}^{N_t} \sum_{j=1}^{N_r} \log_2(1 + \sigma_{t,i} \sigma_{s,i} \sigma_{h,j}), \quad (4.14)$$

其中  $\boldsymbol{\sigma}_s = [\sigma_{s,1}, \dots, \sigma_{s,N_t}]^T$ 。同时, 感知发射功率的约束条件 (4.9b) 等价于

$$\text{tr}(\mathbf{S}\mathbf{S}^H) = \sum_{i=1}^{N_t} \sigma_{s,i} = \|\boldsymbol{\sigma}_s\|_1 \leq P_s. \quad (4.15)$$

因此, 关于感知速率最大化的 SOOP 可以转化为

$$\begin{aligned} T\text{-}SOOP: \quad & \max_{\boldsymbol{\sigma}_s} \quad \bar{\mathcal{R}}_s(\boldsymbol{\sigma}_s), \\ & \text{s.t.} \quad (4.15). \end{aligned} \quad (4.16)$$

#### 4.3.2 通信速率最大化问题

另一方面, 关于通信速率最大化的 SOOP 可以表示为

$$\begin{aligned} C\text{-}SOOP: \quad & \max_{\mathbf{S}, \mathbf{w}_k} \quad \mathcal{R}_c(\mathbf{w}_k, \mathbf{S}) \\ & \text{s.t.} \quad (4.9b). \end{aligned} \quad (4.17)$$

观察式 (4.8) 中  $\mathcal{R}_c(\mathbf{w}_k, \mathbf{S})$  的结构, 不难发现变量  $\mathbf{S}$  和  $\mathbf{w}_k$  是可分离的。换句话说, 可以分别从感知发射波形和通信接收波束两方面考虑通信速率的最大化。因此, 对于一个给定的  $\mathbf{S}$ , 以  $\mathbf{w}_k$  为优化变量的通信速率最大化问题等价于每个通信信号的 SINR 最大化问题, 即

$$\max_{\mathbf{w}_k} \gamma_k. \quad (4.18)$$

利用  $\|\mathbf{S}\mathbf{R}_T\mathbf{S}^H\| = \text{tr}(\Lambda) = \sum_{i=1}^{N_t} \sigma_{s,i} \sigma_{t,i}$ , 式 (4.18) 中的通信信号的 SINR  $\gamma_k$  可以重写为

$$\gamma_k(\boldsymbol{\sigma}_s) = \frac{p_k |\mathbf{h}_k^T \mathbf{w}_k|^2}{\sum_{i \neq k}^K p_i |\mathbf{h}_i^T \mathbf{w}_k|^2 + (\sum_{i=1}^{N_t} \sigma_{s,i} \sigma_{t,i} / L + \sigma_n^2) \|\mathbf{w}_k\|^2}. \quad (4.19)$$

接着，让式 (4.19) 中的分子为 1，SINR 最大化问题 (4.18) 可以等价转换为 SINR 的分母最小化问题，即

$$\min_{\mathbf{w}_k} \quad \mathbf{w}_k^H \tilde{\mathbf{R}}_k \mathbf{w}_k \quad (4.20a)$$

$$\text{s.t.} \quad \sqrt{p_k} \mathbf{w}_k^H \mathbf{h}_k^* = 1, \quad (4.20b)$$

其中  $\tilde{\mathbf{R}}_k = \sum_{i \neq k}^K p_i \mathbf{h}_i^* \mathbf{h}_i^T + (\sum_{i=1}^{N_t} \sigma_{s,i} \sigma_{t,i} / L + \sigma_n^2) \mathbf{I}_{N_r}$ 。为了解决这个问题，可以采用拉格朗日乘子法。具体而言，首先构造拉格朗日函数为

$$\mathcal{L}_c(\mathbf{w}_k) = \mathbf{w}_k^H \tilde{\mathbf{R}}_k \mathbf{w}_k + \lambda (\sqrt{p_k} \mathbf{w}_k^H \mathbf{h}_k^* - 1), \quad (4.21)$$

其中  $\lambda > 0$  是拉格朗日乘子。接着，基于 KKT (Karush - Kuhn-Tucker) 条件，令  $\mathcal{L}_c(\mathbf{w}_k)$  的一阶导数为  $\mathbf{0}$ ，得到

$$\frac{\partial \mathcal{L}_c(\mathbf{w}_k)}{\partial \mathbf{w}_k} = 2\tilde{\mathbf{R}}_k \mathbf{w}_k + \lambda \sqrt{p_k} \mathbf{h}_k^* = \mathbf{0}. \quad (4.22)$$

因此， $\mathbf{w}_k$  可以表示为

$$\mathbf{w}_k = -\frac{\sqrt{p_k}}{2} \lambda \tilde{\mathbf{R}}_k^{-1} \mathbf{h}_k^*. \quad (4.23)$$

然后，在式 (4.23) 上左乘  $\sqrt{p_k} \mathbf{h}_k^T$ ，得到

$$\lambda = -\frac{2}{p_k \mathbf{h}_k^T \tilde{\mathbf{R}}_k^{-1} \mathbf{h}_k^*}. \quad (4.24)$$

最后，将式 (4.24) 代入式 (4.23) 中，可以得到最优的通信接收波束为

$$\mathbf{w}_k^{\text{opt}} = \frac{1}{\sqrt{p_k \mathbf{h}_k^T \tilde{\mathbf{R}}_k^{-1} \mathbf{h}_k^*}} \tilde{\mathbf{R}}_k^{-1} \mathbf{h}_k^*. \quad (4.25)$$

注意到最优的通信波束  $\mathbf{w}_k^{\text{opt}}$  是感知发射波形  $\mathbf{S}$  的函数。基于此，通信速率可以表示为

$$\bar{\mathcal{R}}_c(\boldsymbol{\sigma}_s) = \frac{1}{K} \sum_{k=1}^K \log_2(1 + \gamma_k^{\text{opt}}), \quad (4.26)$$

其中  $\gamma_k^{\text{opt}} = p_k \mathbf{h}_k^T \tilde{\mathbf{R}}_k^{-1} \mathbf{h}_k^*$  是基站使用最优接收波束得到的来自第  $k$  个终端的通信信号的



SINR。因此，关于通信速率最大化的 SOOP 可以转化为

$$\begin{aligned} T-C-SOOP: \quad & \max_{\sigma_s} \quad \bar{\mathcal{R}}_c(\sigma_s), \\ & \text{s.t.} \quad (4.15). \end{aligned} \quad (4.27)$$

### 4.3.3 问题转化

通过重新组合问题 (4.16) 和 (4.27)，原问题 (4.9) 可以等价地转换为

$$\begin{aligned} T-S\&C-SOOP: \quad & \max_{\sigma_s} \quad \frac{\alpha}{M_s} \bar{\mathcal{R}}_s(\sigma_s) + \frac{(1-\alpha)}{M_c} \bar{\mathcal{R}}_c(\sigma_s) \\ & \text{s.t.} \quad (4.15), \end{aligned} \quad (4.28)$$

其中最大可达感知速率  $M_s$  和最大可达通信速率  $M_c$  的详细推导在附录 A.2.3 中。为了更加直观，图 4.3 给出了详细的问题转化过程。

注意到，尽管转换后的问题 (4.28) 的计算维数和设计复杂度与原问题 (4.9) 相比已经大大降低，但由于其目标函数结构极其复杂，仍然很难用传统的优化方法来解决。因此，可以借助 AI 强大的计算能力，开发一种高效的智能优化算法来解决这一问题。

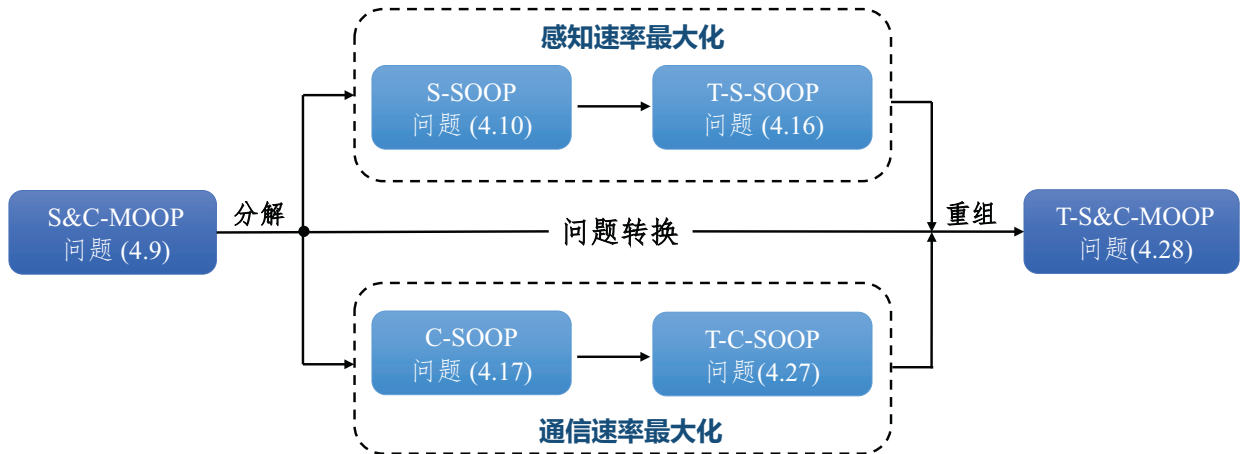


图 4.3 问题 (4.9) 到问题 (4.28) 的转换过程

## 4.4 基于 DL 的算法设计

本节提出了一种基于 DL 的优化算法来获得问题 (4.28) 的可行解。基于前文分析，已知感知性能和通信性能主要受感知发射波形  $\mathbf{S}$  和通信接收波束  $\mathbf{w}_k, \forall k$  的共同影响，而感知

发射波形和通信接收波束的设计又取决于感知发射相关矩阵（Transmit Correlation Matrix, TCM） $\mathbf{R}_T$  和通信 CSI  $\mathbf{H}$ 。因此，可以确定该算法的输入为感知 TCM  $\mathbf{R}_T$  和通信 CSI  $\mathbf{H}$ ，输出为感知发射波形  $\mathbf{S}$  和通信接收波束  $\mathbf{w}_k$ 。如图 4.4 所示，基于 DL 的算法核心是所提的神经网络（Proposed Deep Neural Network, PDNN），主要由预处理模块、网络骨架和后处理模块组成。根据 PDNN 得到的预测结果，利用恢复模块求出问题 (4.28) 的解，即  $\mathbf{S}$  和  $\mathbf{w}_k$ 。接下来详细介绍所提 DL 算法中的每个模块。

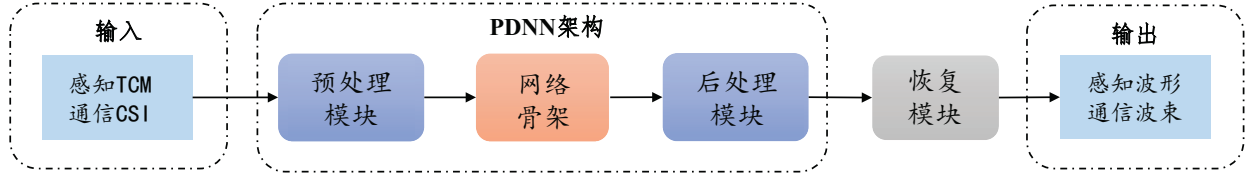


图 4.4 基于DL的优化算法设计框图

#### 4.4.1 预处理模块

由于 PDNN 的输入，即感知 TCM  $\mathbf{R}_T$  和通信 CSI  $\mathbf{H}$ ，是复数矩阵，而 DL 网络目前没有能力处理复数输入，因此需要进行输入预处理。这里采用同相/正交（In-Phase/Quadrature, I/Q）转换法将复数向量  $\mathbf{h}_k$  分为同相分量  $\Re(\mathbf{h}_k^T)$  和正交分量  $\Im(\mathbf{h}_k^T)$ ，其中  $\Re(\cdot)$  和  $\Im(\cdot)$  分别表示实部和虚部。同时，基于转换的问题 (4.28)，输入的感知 TCM 可以由其奇异值  $\sigma_{t,i}, i = 1, \dots, N_t$  代替。因此，将  $\Re(\mathbf{h}_k^T), k = 1, \dots, K$ ， $\Im(\mathbf{h}_k^T), k = 1, \dots, K$  和  $\sigma_{t,i}, i = 1, \dots, N_t$  拼接成一个新的实数向量  $\mathbf{F} \in \mathbb{R}^{(2N_r K + N_t) \times 1}$ ，作为 PDNN 的输入特征向量，即

$$\mathbf{F} = [\Re(\mathbf{h}_1^T), \dots, \Re(\mathbf{h}_K^T), \Im(\mathbf{h}_1^T), \dots, \Im(\mathbf{h}_K^T), \sigma_{1,1}, \dots, \sigma_{t,N_t}]^T. \quad (4.29)$$

#### 4.4.2 后处理模块

根据转化后的优化问题 (4.28)，选择感知发射波形的奇异值  $\sqrt{\sigma_{s,i}}$  的平方，即  $\sigma_{s,i}, i = 1, \dots, N_t$ ，作为 PDNN 的输出。为了保证网络输出的向量满足感知发射功率约束 (4.15)，需要训练两个特征参数，一个是向量  $\boldsymbol{\theta}_s$ ，另一个是标量  $\eta_s$ 。然后，将这两个特征参数拼接起来送到 Lambda 层以满足约束。具体地，Lambda 层的功能可以通过以下步骤实现：

1. 分离出向量  $\boldsymbol{\theta}_s$  和标量  $\eta_s$ ;
2. 归一化向量  $\boldsymbol{\theta}_s$ ，并将其元素从大到小排列，即  $\boldsymbol{\theta}_s = \left[ \frac{\boldsymbol{\theta}_s}{\|\boldsymbol{\theta}_s\|} \right]^\downarrow$ ;

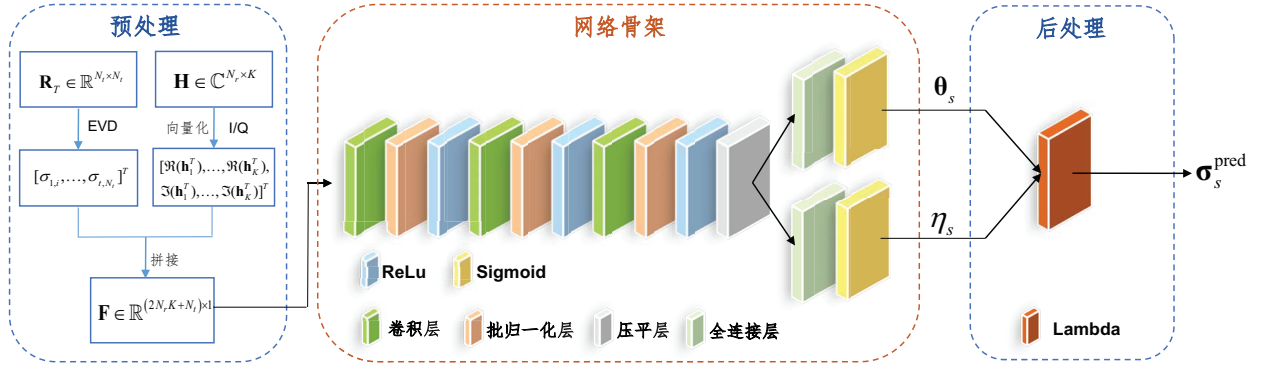


图 4.5 PDNN 架构

3. 用  $P_s$  乘以  $\eta_s$ , 得到  $\eta_s = \eta_s \times P_s$ ;
4. 用  $\eta_s$  乘以  $\theta_s$  得到输出向量  $\sigma_s^{\text{pred}} = \theta_s \times \eta_s$ 。

#### 4.4.3 网络骨架

如图 4.5 所示, PDNN 由输入层、卷积层、批归一化层、压平层、全连接层、使用 ReLU 或 Sigmoid 函数的激活层和 Lambda 层 (输出层) 组成。具体而言, 卷积层用于从输入特征向量中提取特征, 而全连接层可以融合从卷积层提取的特征信息。这是因为全连接层中的每个神经元都与前一层中的所有神经元完全连接。在 PDNN 中, 两个具有  $N_t$  个和 1 个神经元的全连接层分别用于向量  $\theta_s$  和标量  $\eta_s$  的预测。压平层通常用于从卷积层到全连接层的转换, 正如其名, 目的是将输入“压平”, 即将多维输入转换为一维输入。而激活函数的选择主要是基于经验试验确定的, 其中 PDNN 的卷积层均采用 ReLU 函数, 全连接层均采用 Sigmoid 函数。为了加速收敛, 每个卷积层和每个全连接层之前都有一个批归一化层。最后, 输出层由 Lambda 层对输出向量进行缩放和排序, 以满足约束条件 (4.15)。

与传统的监督学习设计不同, PDNN 不需要训练标签。为了有效且直接地提高系统性能, 使用问题 (4.28) 的目标函数值的相反数作为损失函数。在这种情况下, 训练损失的减小恰好对应于归一化感知速率和归一化通信速率加权平均和的增大。具体地, PDNN 的损失函数可以表示为

$$\text{Loss} = -\frac{1}{N_s} \sum_{q=1}^{N_s} \left[ \frac{\alpha}{M_{s,q}} \bar{\mathcal{R}}_s(\sigma_{s,q}^{\text{pred}}) + \frac{(1-\alpha)}{M_{c,q}} \bar{\mathcal{R}}_c(\sigma_{s,q}^{\text{pred}}) \right], \quad (4.30)$$

其中  $N_s$  是训练样本的数量,  $\sigma_{s,q}$  是用第  $q$  个样本经 PDNN 输出的向量,  $M_{s,q}$  和  $M_{c,q}$  分别是对于第  $q$  个样本而言的最大可达感知速率和最大可达通信速率。

#### 4.4.4 恢复模块

最后，利用恢复模块根据输出向量  $\boldsymbol{\sigma}_s^{\text{pred}}$  获得所需的感知发射波形和通信接收波束。具体地，感知发射波形可以表示为：

$$\mathbf{S} = \mathbf{U}_s \boldsymbol{\Sigma}_s^{\text{pred}} \mathbf{U}_T^H, \quad (4.31)$$

其中  $\boldsymbol{\Sigma}_s^{\text{pred}} = [\text{diag}((\boldsymbol{\sigma}_s^{\text{pred}})^{1/2}), \mathbf{0}_{N_t \times (L-N_t)}]^T$ ， $\mathbf{U}_s$  是任意  $L$  维的酉矩阵。此外，通信接收波束可以表示为

$$\mathbf{w}_k = \frac{1}{\sqrt{p_k} \mathbf{h}_k^T (\boldsymbol{\Omega}_k^{\text{pred}})^{-1} \mathbf{h}_k^*} (\boldsymbol{\Omega}_k^{\text{pred}})^{-1} \mathbf{h}_k^*, \quad (4.32)$$

其中  $\boldsymbol{\Omega}_k^{\text{pred}} = \sum_{i \neq k} p_i \mathbf{h}_i^* \mathbf{h}_i^T + (\boldsymbol{\sigma}_t^T \boldsymbol{\sigma}_s^{\text{pred}} / L + \sigma_n^2) \mathbf{I}_{N_r}$ 。

#### 4.4.5 复杂度分析

本小节分析所提 DL 算法的计算复杂度，主要包括预处理、预测和恢复。在预处理阶段，采用了 EVD 操作获取特征值，其复杂度为  $\mathcal{O}(N_t^3)$ 。在预测阶段，前向计算网络的复杂性可以表示为  $\mathcal{O}\left(\sum_{l=1}^{N_L-2} C_l^F C_l^K F_{l-1} F_l + \sum_{l=N_L-1}^{N_L} F_{l-1} F_l\right)$ ，其中  $N_L$  是层数， $C_l^F, C_l^K, F_{l-1}, F_l$  分别是第  $l$  个卷积层特征图映射的区域面积、卷积核面积、输入通道数量和输出通道数量<sup>[65, 118]</sup>。对于恢复阶段，恢复感知发射波形和通信接收波束的复杂度为  $\mathcal{O}(KN_r^3 + L^2 N_t^2)$ 。因此，所提 DL 算法的计算复杂度可以表示为  $\mathcal{O}\left(KN_r^3 + N_t^3 + L^2 N_t^2 + \sum_{l=1}^{N_L-2} C_l^F C_l^K F_{l-1} F_l + \sum_{l=N_L-1}^{N_L} F_{l-1} F_l\right)$ 。为了更直观地表示，表 4.1 中列出了仿真所提 DL 算法所需的平均运行时间，其中仿真参数与 4.5 节的默认设置保持一致。

表 4.1 所提 DL 算法的实际运行时间

|        |        |        |        |        |        |        |        |
|--------|--------|--------|--------|--------|--------|--------|--------|
| $N_t$  | 4      | 6      | 8      | 10     | 12     | 14     | 16     |
| 时间(ms) | 0.2330 | 0.2360 | 0.2404 | 0.2418 | 0.2422 | 0.2447 | 0.2457 |
| $N_r$  | 4      | 6      | 8      | 10     | 12     | 14     | 16     |
| 时间(ms) | 0.1413 | 0.1572 | 0.1732 | 0.1943 | 0.2052 | 0.2206 | 0.2504 |
| $K$    | 2      | 3      | 4      | 5      | 6      | 7      | 8      |
| 时间(ms) | 0.1653 | 0.1897 | 0.2169 | 0.2504 | 0.2687 | 0.2812 | 0.2934 |

## 4.5 仿真结果

表 4.2 所提 DL 算法的仿真参数设置

| 参数        | 仿真值                            |
|-----------|--------------------------------|
| 基站收/发天线数量 | $N_r = 16, N_t = 16$           |
| 终端数量      | $K = 5$                        |
| 小区半径      | 200 m                          |
| 感知权重      | $\alpha = 0.5$                 |
| 感知信号波形长度  | $L = 20$                       |
| 基站最大发射功率  | $\text{SNR}_B = 10 \text{ dB}$ |
| 终端最大发射功率  | $\text{SNR}_C = 0 \text{ dB}$  |
| 训练优化器     | Adam                           |
| 最大训练迭代次数  | 500                            |
| 迷你批大小     | 256                            |
| 初始学习率     | 0.001                          |

本节给出了一些数值仿真结果来验证所提 DL 算法的有效性。假设所有的终端和感知目标都随机分布在小区范围内，路径损耗采用 3GPP 标准中的模型，即  $\text{PL}_{\text{dB}} = 128.1 + 37.6 \log_{10}(d)^{[112]}$ ，其中  $d$  (km) 表示发射机和接收机之间的距离。感知 TCM  $\mathbf{R}_T$  是参照文献[119]中的设置随机生成的。为了便于分析，假设所有终端都具有相同的发射功率  $p_k = P_c, \forall k$ ，用  $\text{SNR}_C = 10 \log_{10}(P_c/\sigma_n^2)$  和  $\text{SNR}_B = 10 \log_{10}(P_s/\sigma_n^2)$  分别表示基站和终端的发射信噪比（以 dB 为单位）。在网络模型的训练过程中，生成 10,000 个训练样本，并将验证分割因子设置为 0.2，即 8,000 个样本用于训练，2,000 个样本用于验证，以评估离线训练阶段的训练效果。同时，产生 2000 个测试样本使用训练好的 PDNN 在线预测结果。设置“PDNN”使用 3 个卷积层，每层分别有 2, 4 和 8 个过滤器，卷积核大小分别为 (5, 1), (3, 1) 和 (3, 1)。此外，使用耐心为 20 的早停机制“EarlyStopping”防止过拟合，以提高训练效率，其中超参数耐心为 20 是指模型在验证集上的表现持续变差超过 20 个迭代次数即停止训练。同时，使用耐心为 10、缩放因子为 0.33 的回调函数“ReduceLRonPlateau”更新学习率，加快收敛速度，其中超参数耐心为 10 是指当经历 10 个迭代轮后模型的性能仍不提升便使用缩放因子来减小学习率。如无特殊说明，仿真参数的默认设置参照表 4.2。

图 4.6 展示了使用不同神经网络训练和测试性能的比较。对比网络为经典的全连接神经网络 (Fully-Connected Neural Network, FCNN)，它主要由四个全连接层组成，分别含有  $8N_t, 4N_t, 2N_t$  和  $N_t$  个神经元。与 PDNN 相似，FCNN 中的每个全连接层之前也采用批归

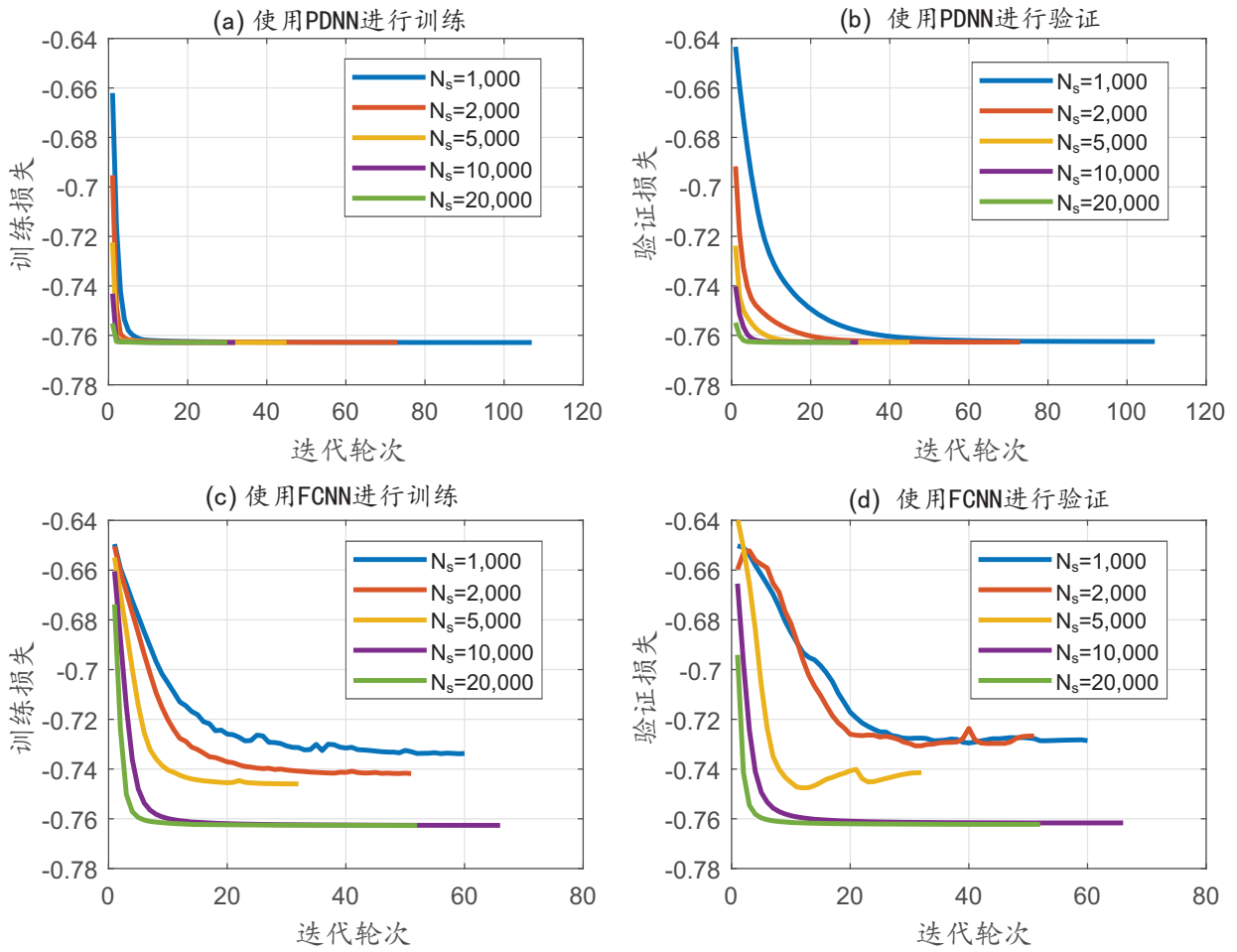


图 4.6 不同 NN 的训练和测试性能比较

一化层，以便更好地训练。可以看出，在没有过拟合的情况下，训练样本越多，收敛所需的训练迭代次数越少，训练效果越好。此外，值得注意的是，即使在小样本训练的情况下，如  $N_s = 1000$ 、 $2000$ 和  $5000$  时，PDNN 在训练集和验证集上都表现良好，而 FCNN 的训练损失曲线和验证损失曲线却剧烈振荡。这表明，所提出的 PDNN 是一种具有较强数据特征提取能力的高性能神经网络，非常适用于通信和感知一体化系统。

图 4.7 给出了所提 DL 算法和两个基线算法的性能比较，即令  $\sigma_{s,i} = \frac{P_s}{N_i}, \forall i$  的平均算法与使用迫零通信接收波束成形和串行干扰消除的迫零算法<sup>[117]</sup>。值得注意的是，本文提出的 DL 算法在  $\alpha = 0$  点上，即仅关注通信时，目标值为 1。也就是说，在感知权重为 0 时得到的加权和速率等于最大通信速率  $M_c$ ，即该解是最优的，证明了所提 DL 算法的有效性。另外，三种算法的加权和速率都随着感知权重  $\alpha$  的增加而减小。这是因为当感知权重  $\alpha$  增加时，系统优化更偏向于感知速率最大化，而通信对感知性能的干扰是不可避免的，导致了总体性能加权和速率的下降。这也可以从点  $\alpha = 1$  上，即仅关注感知时，所提 DL 算法与迫零算法在感知权重为 1 时的性能相同得到验证。可以看出平均算法在低感知权重

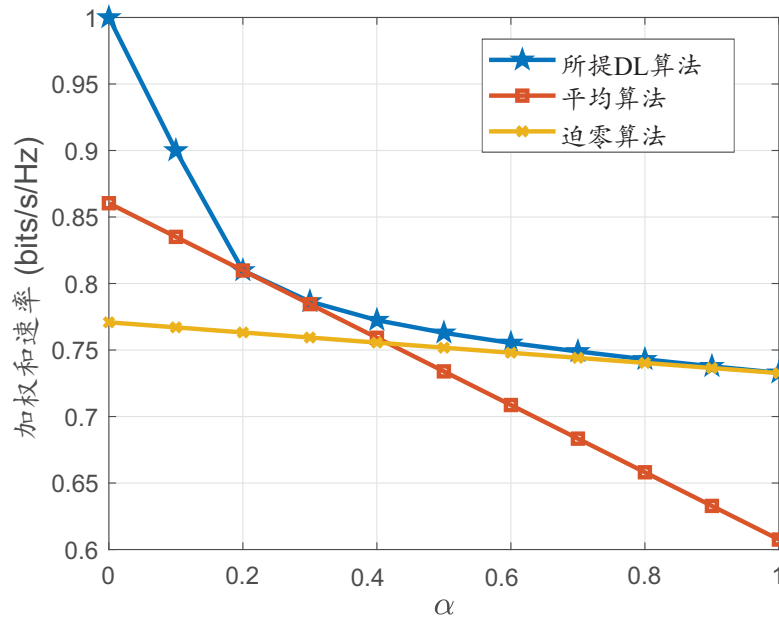


图 4.7 所提的 DL 算法与基线算法的性能比较

区域的性能比迫零算法差，但在高感知权重区域的性能优于迫零算法，而所提的 DL 算法在整个  $\alpha$  区域的性能都表现最好。这是因为平均算法和迫零算法都没有充分考虑感知与通信之间的相互干扰，这也验证了所提 DL 算法的有效性。

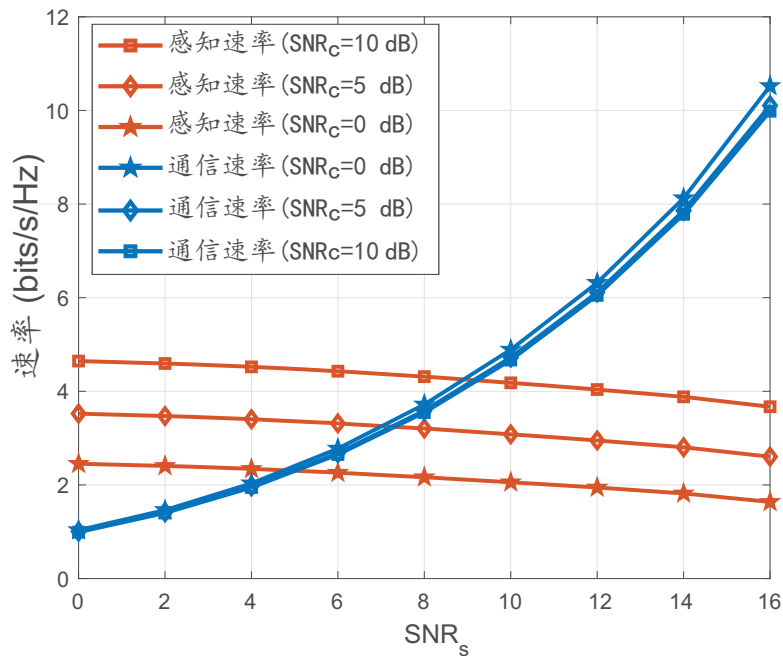


图 4.8 基站和终端的发射功率对性能的影响

接着，图 4.8 探究了基站和终端的发射功率对所提的 DL 算法性能的影响。可以看出，随着基站发射功率的增大，感知速率增大，通信速率减小。这是因为更大的感知信号的发

射功率会带来更多的感知互信息，但会导致更严重的感知对通信的干扰。反之，终端发射功率越大，通信速率越大，感知速率越小。此外，终端发射功率的增加对通信性能的影响较大，而对感知性能的影响较小。同样，基站发射功率的变化对感知性能影响较大，但对通信性能影响较小。因此，可以通过调节基站发射功率和终端发射功率平衡通信性能和感知性能。

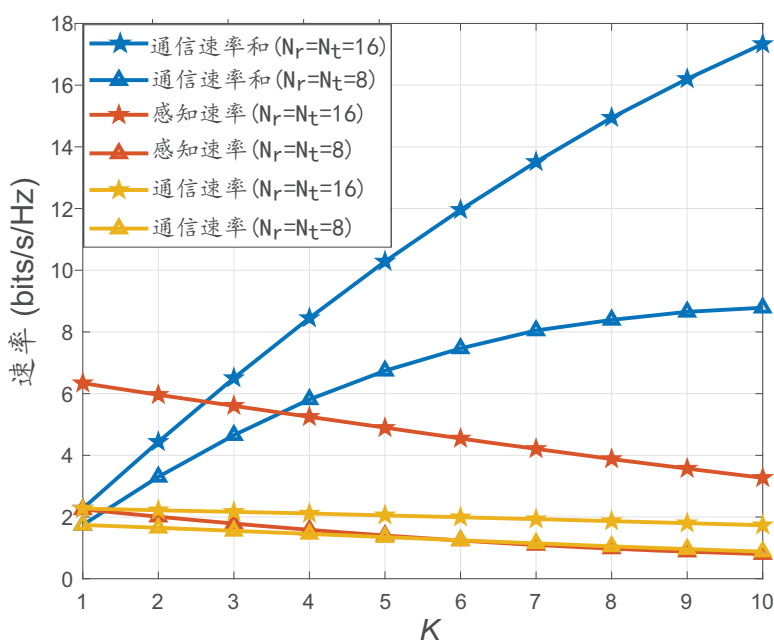


图 4.9 终端数量和基站天线数量对性能的影响

图 4.9 研究了基站天线数量和终端数量对感知速率和通信速率的影响。结果表明，随着终端数目的增加，通信速率总和增大，而感知速率和通信速率均减小。这是因为更多的终端，意味着更多的通信干扰。一方面，通信对感知干扰越大，感知速率越低。另一方面，终端间的干扰过多同样会降低通信性能。此外，随着基站天线数量的增加，得益于更多天线提供的增益，感知速率和通信速率都大幅提高。然而，通过增加基站天线数量获得的增益是有限的，并且增加基站天线会造成较大的开销。因此，通过适当增加系统中基站发射/接收天线的数量来抵抗干扰，提升通信和感知一体化的性能是有意义的。

图 4.10 展示了感知权重和终端数量对系统性能的影响。由图 4.10 (a) 可以看出，当感知权重  $\alpha$  小于等于 0.4 时，加权和速率随  $\alpha$  的增加而减小。当  $\alpha > 0.4$  时，在终端数量较少的情况下 ( $K = 3$ )，加权和速率增长缓慢，而在终端数量较多的情况下 ( $K = 5, 7$ ) 加权和速率开始下降，且终端越多，下降越快。这是因为在有限的资源条件下，感知速率最大化与通信速率最大化是竞争关系，而不是互利关系。当终端数量较小时，可以知道加权和速率曲线的最低点是感知与通信竞争最大的时候。但当终端数量较多时，感知和通信性



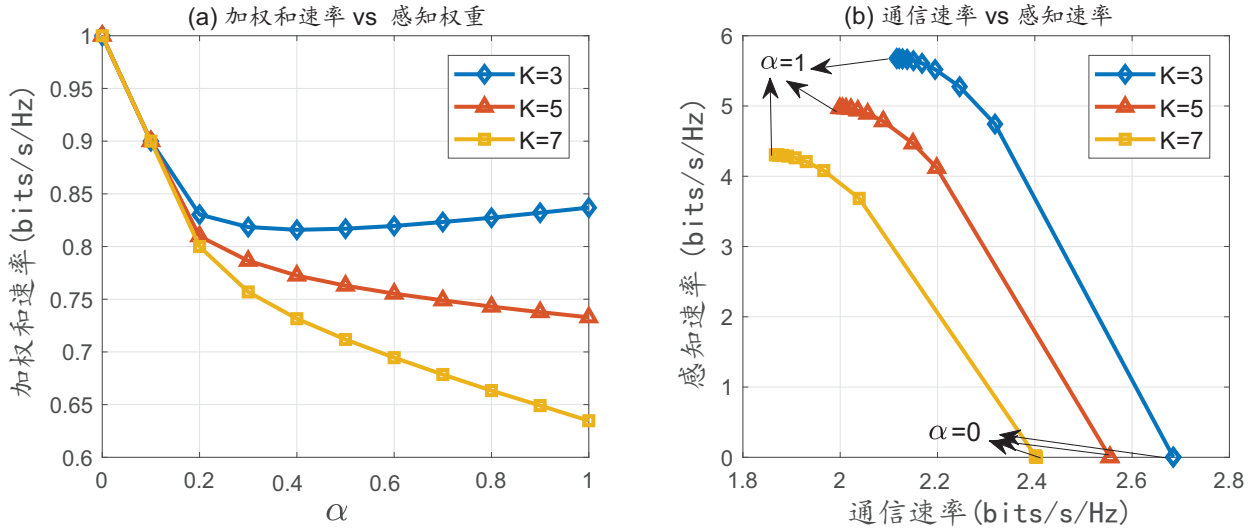


图 4.10 感知权重和终端数量对性能的影响

能都严重下降，导致加权和速率的曲线单调下降。由图 4.10 (b) 可知，感知权重  $\alpha$  在  $[0, 0.5]$  区域时对感知速率和通信速率影响较大，但在  $[0.5, 1]$  区域时对两者影响较小。这是因为当系统优化侧重通信性能时，可以通过调整基站发射功率来提高整体性能，而当系统优化以感知性能为重点时，通信干扰的影响极大地限制了系统性能的提高。因此，选择合适的权重对通信和感知一体化系统性能至关重要。

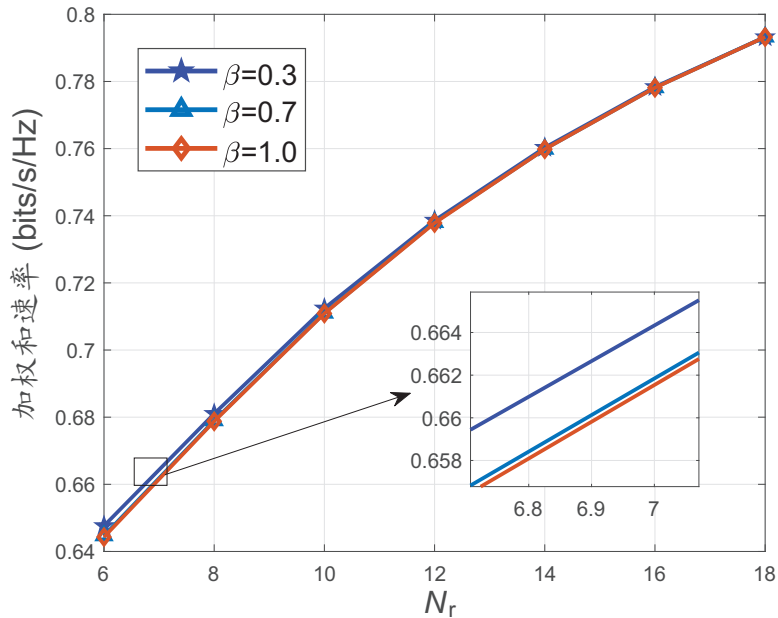


图 4.11 不完美 CSI 和基站接收天线数量对性能的影响

最后，本节考虑了不完美 CSI 的情况，其中通信信道估计存在误差，这适用于大多数实际系统。假设估计的 CSI 和真实的 CSI 可以分别由实际信道估计器和理想信道估计

器获得。为了表示通信 CSI，采用常用的不完美 CSI 模型<sup>[120]</sup>，即  $\mathbf{H} = \sqrt{\beta}\hat{\mathbf{H}} + \mathbf{E}$ ，其中  $\beta \in [0, 1]$  表示信道估计精度， $\mathbf{H}$  为真实 CSI， $\mathbf{E} \sim \mathcal{CN}(0, (1 - \beta)\mathbf{I})$  为信道估计误差矩阵。在离线训练过程中，将估计的 CSI 输入 PDNN，并使用真实的 CSI 计算无监督学习的训练损失。因此，在线预测时，只需要估计的 CSI 就可以得到期望的解。对比图 4.11 中  $\beta = 0.3, 0.7$  的不完美 CSI 情况和  $\beta = 1$  的完美 CSI 情况，可以发现对于所提的 DL 算法，信道估计误差造成的性能损失几乎可以忽略不计。这是因为 PDNN 在训练过程中学习了真实的 CSI 和估计的 CSI 之间的复杂关系，并在预测过程中表现出了鲁棒性。这也体现了基于 DL 的优化算法相对于传统优化算法的优势。此外，随着接收天线数量的增加，性能差距减小，这意味着天线增益可以有效地补偿通信和感知一体化系统中由于信道估计不完善造成的性能损失。

## 4.6 本章小结

本章研究了计算辅助的通信和感知一体化设计。为了减轻资源共享导致的通信内干扰以及通信和感知间互干扰，本章提出了一种以归一化感知速率和归一化通信速率加权和最大化为目标的联合感知发射波形和通信接收波束设计方案。该设计是一个结构复杂的非凸优化问题，难以用传统的优化方法求解。对此，本章先对原问题进行了一系列等价变换以降低设计复杂度，接着给出了基于 DL 的优化算法以获得可行解。最后，数值仿真验证了所提 DL 算法的有效性和鲁棒性。

## 5 通信支撑的感知和计算一体化

### 5.1 引言

为了增强端到端性能，边缘智能网络不仅需要利用感知技术进行实时目标检测和环境感知，还需要利用计算技术进行低延迟数据处理。在传统系统中，感知和计算通常是分开设计的。这是因为这两个功能具有不同的设计目标，前者旨在提高感知精度，而后者试图减少计算延迟。然而，将感知功能与计算功能深度融合以实现各种智能应用，提高边缘智能网络的资源利用效率，是一个必然的趋势。一方面，由于感知系统和通信系统在射频收发机、信道特性和信号处理方面有很多相似之处，利用通感一体化信号同时实现数据传输和环境感知是可能的。另一方面，由于 MEC 技术的发展，通过协调多个边缘节点和网元间的计算资源，可以实现海量数据在网络边缘的实时处理。换句话说，多节点协作通信可以有效传递和汇聚感知信息和计算数据，从而提高感知精度和计算效率。基于此，本章提出了一种通信支撑的感知和计算一体化设计方案，利用多个配有边缘服务器的基站和多个智能终端协作执行感知和计算任务，在边缘智能网络中同时实现高精度的目标感知和低时延的数据计算。考虑到边缘智能网络中部分节点资源受限以及感知和计算任务间存在着耦合关系，需要一种有效的方式来融合感知功能和计算功能。为此，本章在保证感知精度和计算效率的情况下，提出了一种联合卸载选择和资源分配算法来提高感知和计算一体化的整体性能。

内容概述：本章提出了一种通信支撑的感知和计算一体化架构，利用分布式多节点协作通信实现感知信息和计算数据的高效传递和汇聚，有效提升了目标感知精度和数据计算效率。针对边缘智能网络中节点资源的有限性以及感知和计算任务的耦合性，提出了一种联合卸载选择和资源分配算法，增强了感知和计算一体化的整体性能。具体结构安排如下：5.2 节介绍了通信支撑的感知和计算一体化系统模型，5.3 节提出了一种联合卸载选择和资源分配算法，5.4 节通过数值仿真验证了所提算法的有效性，5.5 节总结了本章内容。

## 5.2 系统模型

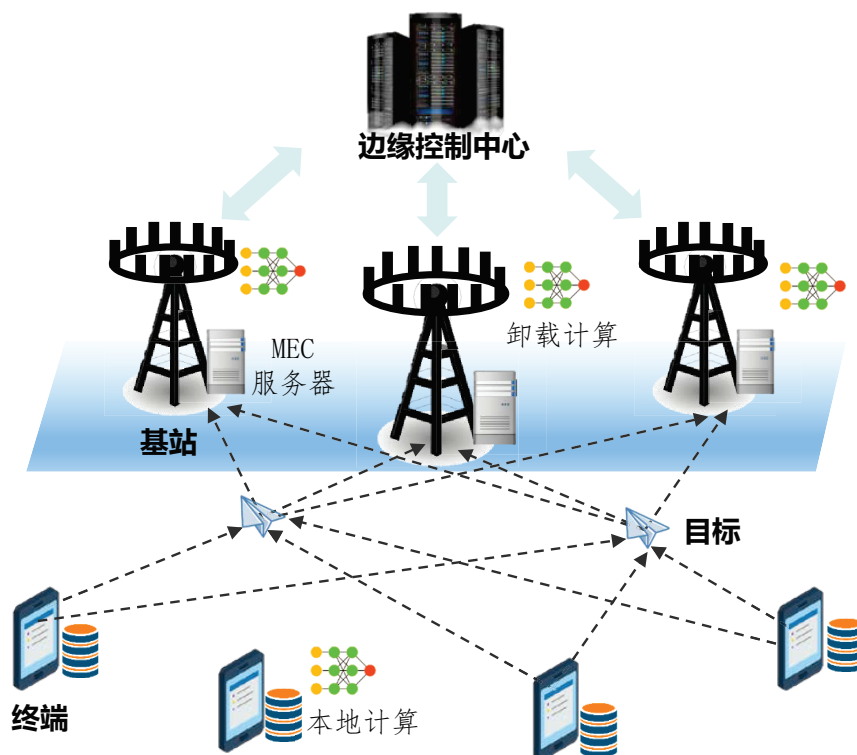


图 5.1 通信支撑的感知和计算一体化系统模型

如图5.1所示，考虑一个通信支撑的感知和通信一体化系统，它由一个边缘控制中心、 $N$  个具有  $K$  根天线并配有 MEC 服务器的基站、 $M$  个单天线的多功能终端和  $L$  个待感知目标组成。系统中的每个终端都有一个需要完成数据处理的计算任务和一个需要合作定位目标的感知任务。具体而言，计算任务是指对终端采集到的数据进行分析、处理或决策，而感知任务是指对每个待感知目标进行定位。其中，计算任务可以由终端在本地执行，也可以卸载到 MEC 服务器上进行处理，而感知任务需要由选择卸载计算的终端、基站和边缘控制中心三者协同完成。通过发射感知和计算双功能信号，每个选择卸载计算的终端将其计算任务卸载到其中一个基站上，并同时执行目标定位。接着，每个基站对接收到的信号进行处理。一方面，基站将目标的位置相关信息上传至边缘控制中心处理感知任务<sup>1</sup>。另一方面，基站对计算任务的数据进行解码，并将解码后的数据送至其 MEC 服务器进行数据处理。计算任务完成后，将计算结果通过下行链路返回给相应的终端。为了更加直观，图 5.2 展示了感知和计算一体化系统流程。接下来详细介绍这一过程。

<sup>1</sup>边缘智能网络的一个优势是低延迟。这是因为数据是在网络边缘而不是云服务器上处理的。由于控制中心也部署在网络边缘，在边缘智能网络的控制中心上处理定位数据仍能保证低延迟。特别是，目标位置信息的数据量通常很小，且基站与控制中心之间通常采用光纤连接，因此在实际系统中传输时延是可以容忍的。再者，多个分布式收发机提供了更丰富的地理信息，这意味着在控制中心上处理感知任务比在基站上处理感知任务能够获得更精确的定位结果。

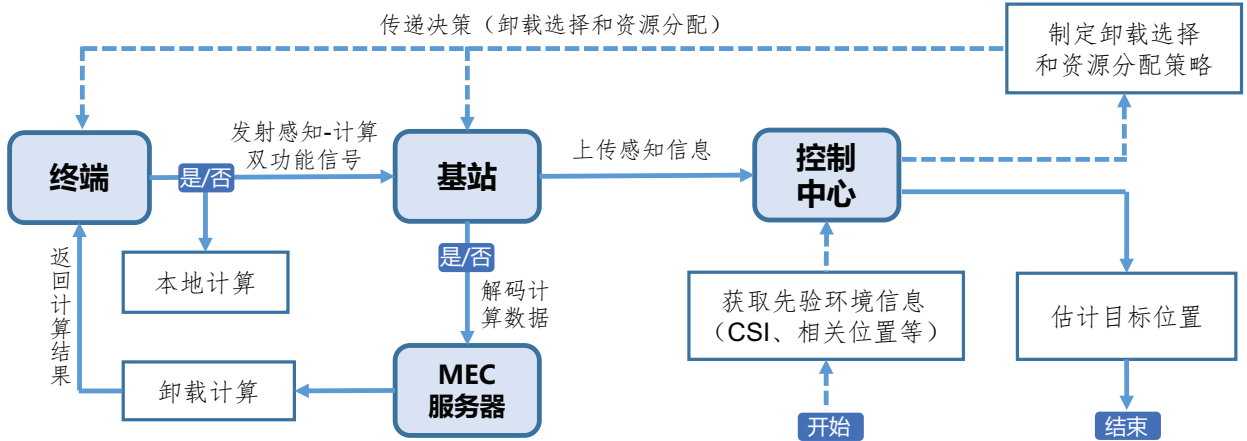


图 5.2 通信支撑的感知和计算一体化系统流程图

不失一般性，假设第  $n$  个基站、第  $m$  个终端和第  $l$  个目标<sup>2</sup>分别位于三维笛卡尔坐标系中的  $(x_n^R, y_n^R, z_n^R)$ ， $(x_m^T, y_m^T, z_m^T)$  和  $(x_l, y_l, z_l)$ 。那么，从  $m$  个终端到第  $l$  个目标的距离、方位角和仰角分别为  $R_{m,l}^T = \sqrt{(x_l - x_m^T)^2 + (y_l - y_m^T)^2 + (z_l - z_m^T)^2}$ ， $\psi_{m,l}^T = \arctan \frac{y_l - y_m^T}{x_l - x_m^T}$  和  $\varphi_{m,l}^T = \arctan \frac{z_l - z_m^T}{\sqrt{(x_l - x_m^T)^2 + (y_l - y_m^T)^2}}$ 。相似地，从第  $l$  个目标到第  $n$  个基站的距离、方位角和仰角分别为  $R_{n,l}^R = \sqrt{(x_l - x_n^R)^2 + (y_l - y_n^R)^2 + (z_l - z_n^R)^2}$ ， $\psi_{n,l}^R = \arctan \frac{y_l - y_n^R}{x_l - x_n^R}$  和  $\varphi_{n,l}^R = \arctan \frac{z_l - z_n^R}{\sqrt{(x_l - x_n^R)^2 + (y_l - y_n^R)^2}}$ 。接着，用二进制卸载系数  $\vartheta_{m,n} \in \{0, 1\}$  表示计算任务的卸载选择，且满足  $\sum_{n=0}^N \vartheta_{m,n} = 1, \forall m$ 。特别地，若第  $m$  个终端的计算任务在本地执行，则  $\vartheta_{m,0} = 1$ ，否则  $\vartheta_{m,0} = 0$ 。若第  $m$  个终端的计算任务被卸载到第  $n$  个基站的 MEC 服务器上，则  $\vartheta_{m,n} = 1$ ，否则  $\vartheta_{m,n} = 0$ 。终端选择本地计算时，即独立完成计算任务，因此不参与协作定位的感知任务。终端选择卸载计算时，通过发送双功能信号，将计算任务卸载到其中一个基站的 MEC 服务器，同时执行目标定位。具体而言，选择卸载计算的终端在第  $t$  个时隙上同时发射一组单位功率的相互正交的双功能信号  $s_m(t), \forall m$ 。对于任意  $m \neq m'$ ，满足  $\int_{\mathcal{T}_m} |s_m(t)|^2 dt = 1, \forall m$  和  $\int_{\mathcal{T}_m} s_m(t) s_{m'}(t)^* dt = 0$ ，其中  $\mathcal{T}_m$  是第  $m$  个终端发射的双功能信号的持续时间<sup>3</sup>。通过合并  $K$  根天线收到的信号，从第  $m$  个终端经第  $l$  个目标散射到第  $n$  个基站上的信号可以表示为

$$r_{m,l,n}(t) = \sum_{k=1}^K \sqrt{(1 - \vartheta_{m,0}) p_m h_{m,l,n,k}} s_m(t - \tau_{m,l,n}) + w_{m,l,n}(t), \quad (5.1)$$

<sup>2</sup>第  $l$  个目标的搜索范围为  $(x_l \pm \frac{\delta c}{B}, y_l \pm \frac{\delta c}{B}, z_l \pm \frac{\delta c}{B})$ ，其中  $\delta$  是给定的搜索系数， $c$  是光速， $B$  是有效信号带宽<sup>[121]</sup>。

<sup>3</sup>对于信号波形，本章采用了目标定位中常用的单位功率正交信号<sup>[122]</sup>，即不存在同频干扰。此外，在感知和计算一体化系统中使用正交的双功能信号既能够提高定位精度，还可以提高计算效率。

其中  $p_m$  是第  $m$  个终端的发射功率,  $w_{m,l,n} \sim \mathcal{CN}(0, \sigma_w^2)$  是方差为  $\sigma_w^2$  的 AWGN,  $\tau_{m,l,n} = (R_{m,l}^T + R_{n,l}^R)/c$  是从第  $m$  个终端经第  $l$  个目标到第  $n$  个基站的传播时延。此外,  $h_{m,l,n,k}$  是从第  $m$  个终端经第  $l$  个目标到第  $n$  个基站上的第  $k$  根天线的信道系数, 包含路径损失、目标散射截面、发射和接收天线间幅度和相位偏差的影响以及滤波器的影响等<sup>[123, 124]</sup>。在感知任务和计算任务执行的过程中, 假设  $h_{m,l,n,k}$  保持不变。接下来, 分别详细讨论感知任务和计算任务。

### 5.2.1 感知任务

基站将接收到的信号发送给控制中心, 控制中心利用无偏估计器确定目标的位置<sup>4</sup>。为了刻画感知任务的性能, 选择常用的克拉美罗界 (Cramer-Rao Bound, CRB) 作为目标定位的性能指标, 它表示了估计位置和实际位置之间 MSE 的下界<sup>[122]</sup>。文献<sup>[121]</sup>中证明了在高信噪比 (超过 10 dB) 的情况下, MSE 近似接近于 CRB。为了方便起见, 定义关于第  $l$  个目标的未知位置参数  $\mathbf{u}_l = [x_l, y_l, z_l]^T$ 。给定向量参数  $\mathbf{u}_l$ , 无偏估计  $\hat{\mathbf{u}}_l$  需满足以下不等式:

$$E \left\{ (\hat{\mathbf{u}}_l - \mathbf{u}_l) (\hat{\mathbf{u}}_l - \mathbf{u}_l)^T \right\} \geq \mathbf{J}_l^{-1}, \forall l, \quad (5.2)$$

其中  $\mathbf{J}_l$  是费舍尔信息矩阵 (Fisher Information Matrix, FIM), 详细推导请见附录 A.3.1。因此, 相应的 CRB 矩阵  $\mathbf{C}_l^{x,y,z}$  可由对矩阵  $\mathbf{J}_l$  进行求逆得到, 即

$$\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}) = \mathbf{J}_l^{-1} = \left( \sum_{m=1}^M (1 - \vartheta_{m,0}) p_m \mathbf{G}_{l,m} \right)^{-1}, \quad (5.3)$$

其中向量  $\bar{\mathbf{p}} = [(1 - \vartheta_{1,0}) p_1, \dots, (1 - \vartheta_{M,0}) p_M]^T$ , 矩阵  $\mathbf{G}_{l,m} = \begin{bmatrix} g_{l,m}^x & g_{l,m}^{xy} & g_{l,m}^{xz} \\ g_{l,m}^{xy} & g_{l,m}^y & g_{l,m}^{yz} \\ g_{l,m}^{xz} & g_{l,m}^{yz} & g_{l,m}^z \end{bmatrix}$ 。具体地,

$$g_{l,m}^x = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{x}_{m,l,n}^2, \quad g_{l,m}^y = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{y}_{m,l,n}^2, \quad g_{l,m}^z = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{z}_{m,l,n}^2.$$

此外,  $g_{l,m}^{xy} = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{x}_{m,l,n} \bar{y}_{m,l,n}$ ,  $g_{l,m}^{yz} = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{y}_{m,l,n} \bar{z}_{m,l,n}$ ,  $g_{l,m}^{xz} = \phi_m \sum_{n=1}^N \sum_{k=1}^K |h_{m,l,k,n}|^2 \bar{x}_{m,l,n} \bar{z}_{m,l,n}$ 。其中,  $\phi_m = \frac{8\pi^2 B^2}{c^2 \sigma_w^2}$ ,  $\bar{x}_{m,l,n} = \sin \psi_{m,l}^T \cos \varphi_{m,l}^T + \sin \psi_{n,l}^R \cos \varphi_{n,l}^R$ ,  $\bar{y}_{m,l,n} = \sin \psi_{m,l}^T \sin \varphi_{m,l}^T + \sin \psi_{n,l}^R \sin \varphi_{n,l}^R$ ,  $\bar{z}_{m,l,n} = \cos \psi_{m,l}^T + \cos \psi_{n,l}^R$ 。

在三维坐标系中, CRB 矩阵的迹与第  $l$  个目标位置估计的 MSE 之间关系可以表示为

<sup>4</sup>基于最大似然估计方法, 控制中心可以根据接收到的信号获得到达时间和到达角度等相关信息, 从而估计出目标的位置<sup>[125, 126]</sup>。

$\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) \leq \sigma_{x,l}^2 + \sigma_{y,l}^2 + \sigma_{z,l}^2$ , 其中  $\sigma_{x,l}^2$ ,  $\sigma_{y,l}^2$  和  $\sigma_{z,l}^2$  分别为第  $l$  个目标关于  $x$ ,  $y$ ,  $z$  轴的位置估计的 MSE。因此, 可以进一步选择 CRB 矩阵的迹作为定位目标的性能指标。通过一系列的矩阵操作, 第  $l$  个目标的 CRB 矩阵的迹可以简化为

$$\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) = \frac{\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}}{\bar{\mathbf{p}}^T (\mathbf{g}_l^x \bar{\mathbf{p}}^T \mathbf{B}_l^x - \mathbf{g}_l^{xy} \bar{\mathbf{p}}^T \mathbf{B}_l^{xy} + \mathbf{g}_l^{xz} \bar{\mathbf{p}}^T \mathbf{B}_l^{xz}) \bar{\mathbf{p}}}, \quad (5.4)$$

其中  $\mathbf{g}_l^x = [g_{1,l}^x, \dots, g_{M,l}^x]^T$ ,  $\mathbf{g}_l^y = [g_{1,l}^y, \dots, g_{M,l}^y]^T$ ,  $\mathbf{g}_l^z = [g_{1,l}^z, \dots, g_{M,l}^z]^T$ ,  $\mathbf{g}_l^{xz} = [g_{1,l}^{xz}, \dots, g_{M,l}^{xz}]^T$ ,  $\mathbf{g}_l^{xy} = [g_{1,l}^{xy}, \dots, g_{M,l}^{xy}]^T$ ,  $\mathbf{g}_l^{yz} = [g_{1,l}^{yz}, \dots, g_{M,l}^{yz}]^T$ 。此外,  $\mathbf{A}_l = \mathbf{g}_l^y (\mathbf{g}_l^z)^T + \mathbf{g}_l^x (\mathbf{g}_l^z)^T + \mathbf{g}_l^x (\mathbf{g}_l^y)^T - \mathbf{g}_l^{xy} (\mathbf{g}_l^{xy})^T - \mathbf{g}_l^{xz} (\mathbf{g}_l^{xz})^T - \mathbf{g}_l^{yz} (\mathbf{g}_l^{yz})^T$ ,  $\mathbf{B}_l^x = \mathbf{g}_l^y (\mathbf{g}_l^z)^T - \mathbf{g}_l^{yz} (\mathbf{g}_l^{yz})^T$ ,  $\mathbf{B}_l^{xy} = \mathbf{g}_l^{xy} (\mathbf{g}_l^z)^T - \mathbf{g}_l^{yz} (\mathbf{g}_l^{xz})^T$ ,  $\mathbf{B}_l^{xz} = \mathbf{g}_l^{xy} (\mathbf{g}_l^{yz})^T - \mathbf{g}_l^y (\mathbf{g}_l^{xz})^T$ 。对于式 (5.4) 中的 CRB 矩阵的迹, 有如下命题:

**命题5.1.** CRB 矩阵的迹  $\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}))$  是  $\bar{\mathbf{p}}$  的单调递减函数。换句话说,  $\bar{\mathbf{p}}$  越大,  $\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}))$  越小。

证明: 请参阅附录 A.3.2。 □

由命题 5.1 可知, 参与协作定位的终端数量越多, 发射功率越高, 相应 CRB 矩阵的迹越小, 即定位精度越高。因此, 选择更多数量的终端进行卸载计算和增加终端的发射功率可以提高目标定位的精度。

### 5.2.2 计算任务

同时, 基站将接收到的信号送至 MEC 服务器进行数据计算。为了便于表示, 将第  $m$  个终端的计算任务定义为  $\Omega_m \triangleq (D_m, C_m)$ , 其中  $D_m$  和  $C_m$  分别为输入数据的大小 (单位为比特, 即 bit) 和计算强度 (单位为计算每比特数据所需的中央处理器运行的周期数, 即 CPU cycles/bit)。为了刻画计算任务的性能, 选择常用的执行时间和能量消耗作为数据处理的性能指标。接下来, 将分别讨论计算任务在终端本地执行和卸载到 MEC 服务器上执行所需的时间和能耗。

对于本地计算, 完成第  $m$  个终端的计算任务所需时间可以表示为

$$T_m^{\text{loc}} = \vartheta_{m,0} \frac{D_m C_m}{f_m^{\text{loc}}}, \quad (5.5)$$

其中  $f_m^{\text{loc}}$  是第  $m$  个终端的计算能力 (单位为每秒执行 CPU 周期数, 即 CPU cycles/s)。本

地计算相应的能耗（单位为焦耳，即 J）可以表示为

$$E_m^{\text{loc}} = \vartheta_{m,0} \kappa_m^U D_m C_m (f_m^{\text{loc}})^2, \quad (5.6)$$

其中， $\kappa_m^U$  是第  $m$  个终端上与芯片架构相关的能量系数。

对于卸载计算，执行时间主要包含两部分，即通过上行链路传输数据信号的时间  $T_m^{\text{tra}}$  和在 MEC 服务器上执行数据处理的时间  $T_m^{\text{ser}}$ 。注意，由于计算结果远远小于原始任务的数据尺寸，基站通过下行链路将计算结果回传给相应终端的时间可以忽略。因此，完成第  $m$  个终端卸载的计算任务所需时间可以表示为

$$T_m^{\text{off}} = T_m^{\text{tra}} + T_m^{\text{ser}} = \sum_{n=1}^N \vartheta_{m,n} \frac{D_m}{R_{m,n}} + \sum_{n=1}^N \vartheta_{m,n} \frac{D_m C_m}{f_{m,n}^{\text{ser}}}, \quad (5.7)$$

其中  $R_{m,n} = B \log_2 \left( 1 + p_m \sum_{l=1}^L \sum_{k=1}^K |h_{m,l,n,k}|^2 / \sigma_w^2 \right)$  是第  $m$  个终端将其计算任务卸载到第  $n$  个基站的数据传输速率， $f_{m,n}^{\text{ser}}$  是第  $n$  个基站上的 MEC 服务器分配给来自第  $m$  个终端的计算任务的计算能力。相应地，卸载计算的能耗也主要包括两部分，即通过上行链路传输数据的能耗  $E_m^{\text{tra}}$  和在 MEC 服务器上处理数据的能耗  $E_m^{\text{ser}}$ 。因此，远程完成第  $m$  个终端的计算任务所需能耗可以表示为

$$E_m^{\text{off}} = E_m^{\text{tra}} + E_m^{\text{ser}} = \sum_{n=1}^N \vartheta_{m,n} p_m \frac{D_m}{R_{m,n}} + \sum_{n=1}^N \vartheta_{m,n} \kappa_n^M (f_{m,n}^{\text{ser}})^2 D_m C_m, \quad (5.8)$$

其中， $\kappa_n^M$  是第  $n$  个基站上的 MEC 服务器上与芯片架构相关的能量系数。

观察式 (5.5)-(5.8)，可以发现发射功率  $p_m$ 、计算能力  $f_m^{\text{loc}}$  和  $f_{m,n}^{\text{ser}}$  的资源分配对完成计算任务所需的执行时间和能耗有着不同的影响。具体而言，发射功率越大，计算能力越强，执行计算任务所需时间就越短，但相应能耗却越高。再者，卸载选择  $\vartheta_{m,n}$  也是影响计算任务性能的关键因素。当终端选择本地计算时，虽然消耗的能量较少，但由于自身计算能力的限制，可能会导致较高的延迟，从而无法满足时间敏感型应用的实时性需求。

### 5.3 联合卸载选择和资源分配设计

从上述分析可以了解到，感知任务和计算任务的性能都受到计算卸载选择  $\vartheta_{m,n}$  和终端的发射功率  $p_m$  的影响。并且，感知任务和计算任务是相互作用的。举个简单的例子，



选择更多的终端进行计算卸载会提高定位的精度，但同时也会带来较高的能耗。与此同时，终端上计算能力  $f_m^{\text{loc}}$  和 MEC 上计算能力  $f_{m,n}^{\text{ser}}$  的分配对完成计算任务所需时间和能耗有显著影响。因此，可以设计一种联合卸载选择和资源分配算法，以在资源有限的情况下共同提高感知精度和计算效率。

### 5.3.1 问题构造

考虑到感知任务和计算任务之间的竞争性和耦合性，该设计以最小化所有计算任务的加权总能耗为目标，同时需要保证计算任务的执行时间和感知任务的定位精度要求。具体地，该设计可以表述为如下优化问题：

$$\min_{\mathbf{p}, \boldsymbol{\vartheta}, \mathbf{f}, \mathbf{F}} \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{off}}) \quad (5.9a)$$

$$\text{s.t.} \quad T_m^{\text{loc}} + T_m^{\text{off}} \leq \varsigma_m, \forall m \in \mathcal{M}, \quad (5.9b)$$

$$\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) \leq \eta_l, \forall l \in \mathcal{L}, \quad (5.9c)$$

$$\vartheta_{m,n} \in \{0, 1\}, \forall m \in \mathcal{M}, \forall n \in \mathcal{N}', \quad (5.9d)$$

$$\sum_{n=0}^N \vartheta_{m,n} = 1, \forall m \in \mathcal{M}, \quad (5.9e)$$

$$\sum_{m=1}^M \vartheta_{m,n} f_{m,n}^{\text{ser}} \leq F_n^M, \forall n \in \mathcal{N}, \quad (5.9f)$$

$$0 \leq f_{m,n}^{\text{ser}} \leq F_n^M, \forall m \in \mathcal{M}, \forall n \in \mathcal{N}, \quad (5.9g)$$

$$0 \leq f_m^{\text{loc}} \leq F_m^U, \forall m \in \mathcal{M}, \quad (5.9h)$$

$$0 \leq p_m \leq P_m^{\text{max}}, \forall m \in \mathcal{M}, \quad (5.9i)$$

其中， $\mathcal{M} = \{1, 2, \dots, M\}$ ， $\mathcal{L} = \{1, 2, \dots, L\}$ ， $\mathcal{N} = \{1, 2, \dots, N\}$  和  $\mathcal{N}' = \{0, 1, \dots, N\}$  为定义的集合。 $\mathbf{p} = \{p_m, \forall m \in \mathcal{M}\}$ ， $\boldsymbol{\vartheta} = \{\vartheta_{m,n}, \forall m \in \mathcal{M}, \forall n \in \mathcal{N}'\}$ ， $\mathbf{f} = \{f_m^{\text{loc}}, \forall m \in \mathcal{M}\}$  和  $\mathbf{F} = \{f_{m,n}^{\text{ser}}, \forall m \in \mathcal{M}, \forall n \in \mathcal{N}\}$  为优化变量。式 (5.9a) 中的优化目标函数为完成所有计算任务的加权总能耗，其中  $\lambda_m > 0$  为第  $m$  个终端的权重。约束条件 (5.9b) 和 (5.9c) 分别是计算任务的执行时间要求和感知任务的定位精度要求，其中  $\varsigma_m$  为第  $m$  个终端最大可容忍的时延， $\eta_l$  是第  $l$  个目标最大可容忍的 CRB 矩阵的迹。约束条件 (5.9d) 和 (5.9e) 均为计算卸载选择限制，表示每个计算任务可以选择在本地执行，或者选择其中一个基站进行计算卸载。约束条件 (5.9f)、(5.9g) 和 (5.9h) 是计算能力限制，其中  $F_m^U$  和  $F_n^M$  分别是第  $m$  个终端和第  $n$  个基站的 MEC 服务器的总计算能力。最后，约束条件 (5.9i) 是终端发射功率

限制，其中  $P_m^{\max}$  为第  $m$  个终端的最大发射功率预算。

可以看出，问题 (5.9) 是一个计算困难的混合整数非线性规划（Mixed Integer Nonlinear Programming, MINLP），很难在多项式时间内获得其最优解。在这种情况下，利用问题 (5.9) 中目标函数和限制条件的结构，可以通过 AO 方法获得一个尽可能接近最优解的可行次优解。下面详细描述算法设计的过程。

### 5.3.2 算法设计

首先将原本复杂的 MINLP 问题解耦成两个子问题，即卸载选择和资源优化。接着，在迭代中依次求解这两个子问题。最后，当原问题的目标值收敛时，获得可行的次优解。对于卸载选择，通过固定发射功率和计算能力，该问题可以表述为

$$\begin{aligned} \min_{\vartheta} \quad & \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{off}}) \\ \text{s.t.} \quad & (5.9b) - (5.9f). \end{aligned} \quad (5.10)$$

显然，问题 (5.10) 是一个经典的整数规划问题。因此，可以采取最常用的分支定界（Branch and Bound, BnB）法来解决这个问题<sup>[127]</sup>。具体而言，问题 (5.10) 可以看做 BnB 树的根节点，而 BnB 树是通过分支和定界的方式被创建并搜索的。对于分支，通过在叶节点上添加互斥的约束，可以将一个父问题分为两个子问题，即一个叶节点添加约束  $\vartheta_{m,n} = 1$ ，另一个叶节点添加约束  $\vartheta_{m,n} = 0$ 。对于定界，通过检查子问题目标的上界和下界来确定是否保留或修剪当前分支。特别地，每个子问题目标的下界可以通过求解其松弛问题来获得，其中原问题的二进制离散约束被松弛为连续约束。为了获得问题 (5.10) 目标值的下界，需要解决如下松弛问题：

$$\begin{aligned} \min_{\vartheta} \quad & \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{off}}) \\ \text{s.t.} \quad & (5.9b), (5.9c), (5.9e), (5.9f), \end{aligned} \quad (5.11a)$$

$$0 \leq \vartheta_{m,n} \leq 1, \forall m \in \mathcal{M}, \forall n \in \mathcal{N}', \quad (5.11b)$$

其中，约束条件 (5.9c) 具有复杂的非凸结构，参见式 (5.4)。为了处理这个约束条件，利用式 (5.3) 中 CRB 矩阵的特殊结构，即 FIM  $\mathbf{J}_l$  的逆，引入一个辅助矩阵变量  $\Theta_l$ ，其满足以下条件：

$$\Theta_l \succeq \mathbf{J}_l^{-1} \quad (5.12)$$

基于舒尔补引理（参见引理 3.1），附加的约束条件 (5.12) 可以重新表示为

$$\begin{bmatrix} \Theta_l & \mathbf{I} \\ \mathbf{I} & \mathbf{J}_l \end{bmatrix} \succeq \mathbf{0}. \quad (5.13)$$

在这种情况下，约束条件 (5.9c) 可以被替换为

$$\text{tr}(\Theta_l) \leq \eta. \quad (5.14)$$

因此，问题 (6.10) 可以等价地转换为如下凸优化问题：

$$\begin{aligned} \min_{\vartheta, \Theta} \quad & \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{off}}) \\ \text{s.t.} \quad & (5.9b), (5.9e), (5.9f), (5.11b), (5.13), (5.14). \end{aligned} \quad (5.15)$$

根据 **BnB** 算法的原理，问题 (5.15) 的最优解和目标值可以分别看作是问题 (5.10) 的松弛解和目标值的下界。基于此，可以在分支和定界的过程中找到问题 (5.10) 的最优解。综上，用于解决卸载选择子问题的 **BnB** 算法总结在算法 3 中。

接着，对于给定的卸载选择策略，资源分配子问题可以表示为

$$\begin{aligned} \min_{\mathbf{p}, \mathbf{f}, \mathbf{F}, \Theta} \quad & \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{off}}) \\ \text{s.t.} \quad & (5.9b), (5.9f) - (5.9i), (5.13), (5.14), \end{aligned} \quad (5.16)$$

其中凸约束条件 (5.13) 和 (5.14) 替换了原问题 (5.9) 中的约束条件 (5.9c)。然而，由于复杂的非凸目标函数，问题 (5.16) 仍然是非凸的。特别是，为卸载计算进行上行链路数据传输所消耗的能量，可以看作是发射功率  $p_m, \forall m$  的非凸函数，即

$$E_m^{\text{tra}}(p_m) = \frac{(1 - \vartheta_{m,0})p_m D_m}{B \log_2(1 + p_m \chi_m)}, \quad (5.17)$$

其中  $\chi_m = \sum_{l=1}^L \sum_{k=1}^K \sum_{n=1}^N \vartheta_{m,n} |h_{m,l,n,k}|^2 / \sigma_w^2$ 。为了解决  $E_m^{\text{tra}}(p_m)$  的非凸性，需要引进一个辅助变量  $q_m = \frac{1 - \vartheta_{m,0}}{\log_2(1 + p_m \chi_m)} \geq 0$ 。因此， $E_m^{\text{tra}}(p_m)$  等价于

$$E_m^{\text{tra}}(q_m) = \frac{D_m}{\chi_m B} q_m \left( 2^{\frac{1 - \vartheta_{m,0}}{q_m}} - 1 \right). \quad (5.18)$$

**算法 3**: 卸载选择设计算法

**输入**:  $N, M, K, L, B, \sigma_w^2, \lambda_m, D_m, C_m, \tau_m, \kappa_m^U, F_n^M, \kappa_n^M, \eta_l, \mathbf{p}, \mathbf{f}, \mathbf{F}, \forall m, n, l$ 
**输出**:  $\vartheta$ 

- 1: 初始化 BnB 树节点  $\mathcal{W} = \{W_0\}$ , 节点约束  $\mathcal{C}_0 = \emptyset$ , 目标值的上界  $Z_{\text{ub}} = \infty$ , 节点索引  $i = -1$ , 向量化解矩阵  $\vartheta$  为等价的解向量  $\mathbf{v} = [v_1, v_2, \dots, v_{M*(N+1)}]^T$ , 其中  $v_{m+(m*n)} = \vartheta_{m,n}$ ;
- 2: **while**  $\mathcal{W} \neq \emptyset$  **do**
- 3:    $i \leftarrow i + 1$ ;
- 4:    $W_i \leftarrow$  从 BnB 树  $\mathcal{W}$  中弹出第一个节点;
- 5:   解决伴随着节点  $W_i$  上约束  $\mathcal{C}_i$  的问题 (5.15), 获得松弛解  $\mathbf{v}^{(i)}$  并更新目标值  $Z^{(i)}$ ;
- 6:   **if** 没有可行解 **then**
- 7:     **return**
- 8:   **end if**
- 9:   **if**  $\mathbf{v}^{(i)}$  不是整数 **then**
- 10:     **if**  $Z^{(i)} < Z_{\text{ub}}$  **then**
- 11:       在  $\mathbf{v}^{(i)}$  中找到第一个非整数元素  $v_j^{(i)}, \forall j \in \{1, 2, \dots, M*(N+1)\}$ ;
- 12:       将节点  $W_i$  分支成两个子节点  $W_i^{(0)}$  和  $W_i^{(1)}$ , 分别添加约束条件  $\mathcal{C}_i^{(0)} : \{v_j^{(i)} = 0\}$  和  $\mathcal{C}_i^{(1)} : \{v_j^{(i)} = 1\}$ ;
- 13:        $\mathcal{W} \leftarrow \mathcal{W} \cup \{W_i^{(0)}, W_i^{(1)}\}$
- 14:     **end if**
- 15:   **else**
- 16:     **if**  $Z^{(i)} < Z_{\text{ub}}$  **then**
- 17:        $Z_{\text{ub}} \leftarrow Z^{(i)}$ ;
- 18:       得到最优解  $\mathbf{v}^* \leftarrow \mathbf{v}^{(i)}$ ;
- 19:     **break**
- 20:   **end if**
- 21: **end if**
- 22: **end while**
- 23: 恢复解矩阵  $\vartheta^* \leftarrow \mathbf{v}^*$ .

相应地, 约束条件 (5.9b) 和 (5.13) 也可以分别被替换为

$$T_m^{\text{loc}} + \frac{q_m D_m}{B(1 - \vartheta_{m,0})} + T_m^{\text{ser}} \leq \tau_m, \quad (5.19)$$

和

$$\begin{bmatrix} \Theta_l & \mathbf{I} \\ \mathbf{I} & \bar{\mathbf{J}}_l(\mathbf{q}) \end{bmatrix} \succeq \mathbf{0}, \quad (5.20)$$

其中,  $\bar{\mathbf{J}}_l(\mathbf{q}) = \sum_{m=1}^M \frac{1 - \vartheta_{m,0}}{\chi_m} \left( 2^{\frac{1 - \vartheta_{m,0}}{q_m}} - 1 \right) \mathbf{G}_{l,m}$ ,  $\mathbf{q} = \{q_m, \forall m \in \mathcal{M}\}$ 。然而, 转换后的约束条件 (5.20) 就辅助变量  $\mathbf{q}$  而言是非凸的。为了解决这个问题, 需要采用连续凸逼近 (Successive Convex Approximation, SCA) 技术来处理约束条件 (5.20)。具体地,  $f(q_m) =$

**算法 4**: 资源分配设计算法

输入:  $N, M, K, L, B, \sigma_w^2, P_m^{\max}, \lambda_m, D_m, C_m, \tau_m, F_m^U, \kappa_m^U, F_n^M, \kappa_n^M, \eta_l, \vartheta, \forall m, n, l$

输出:  $\mathbf{p}, \mathbf{f}, \mathbf{F}$

- 1: 初始化 迭代索引  $i = 0$ , 泰勒展开点  $\tilde{q}_m^{(i)} = \frac{1 - \vartheta_{m,0}}{\log(1 + P_m^{\max} \chi_m)}$ ;
- 2: **repeat**
- 3: 通过固定  $\tilde{\mathbf{q}}^{(i)}$ , 解决问题 (5.23), 获得  $\mathbf{q}^{(i+1)}$ ,  $\mathbf{f}^{(i+1)}$  和  $\mathbf{F}^{(i+1)}$ ;
- 4: 更新  $\tilde{\mathbf{q}}^{(i+1)} = \mathbf{q}^{(i+1)}$ ;
- 5:  $i = i + 1$ ;
- 6: **until** 收敛
- 7: 根据式 (5.24) 计算  $\mathbf{p}$ .

$2^{\frac{1 - \vartheta_{m,0}}{q_m}}$  在点  $\tilde{q}_m$  处的一阶泰勒展开式可以表示为

$$\tilde{f}(q_m, \tilde{q}_m) = 2^{\frac{1 - \vartheta_{m,0}}{\tilde{q}_m}} + \frac{(1 - \vartheta_{m,0}) \ln 2}{\tilde{q}_m^2} \cdot 2^{\frac{1 - \vartheta_{m,0}}{\tilde{q}_m}} (\tilde{q}_m - q_m), \quad (5.21)$$

其中  $\tilde{\mathbf{q}} = \{\tilde{q}_m, \forall m \in \mathcal{M}\}$ ,  $\tilde{q}_m$  是  $q_m$  在上一轮迭代中的值。然后, 通过用  $\bar{\mathbf{J}}_l(\mathbf{q}, \tilde{\mathbf{q}}) = \sum_{m=1}^M \frac{1 - \vartheta_{m,0}}{\chi_m} \left( \tilde{f}(q_m, \tilde{q}_m) - 1 \right) \mathbf{G}_{l,m}$  代替  $\bar{\mathbf{J}}_l(\mathbf{q})$ , 约束条件 (5.20) 可以重新表示为

$$\begin{bmatrix} \Theta_l & \mathbf{I} \\ \mathbf{I} & \bar{\mathbf{J}}_l(\mathbf{q}, \tilde{\mathbf{q}}) \end{bmatrix} \succeq \mathbf{0}. \quad (5.22)$$

因此, 资源分配子问题最终可以表述为如下凸优化问题<sup>5</sup>:

$$\begin{aligned} \min_{\mathbf{q}, \mathbf{f}, \mathbf{F}, \Theta} \quad & \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{tra}}(q_m) + E_m^{\text{ser}}) \\ \text{s.t.} \quad & (5.9f) - (5.9i), (5.14), (5.19), (5.22). \end{aligned} \quad (5.23)$$

因此, 可以使用一些现成的凸优化工具直接获得其最优解。得到解后, 基站的发射功率  $p_m$  可以根据如下式子计算:

$$p_m = \frac{1}{\chi_m} \left( 2^{\frac{1 - \vartheta_{m,0}}{q_m}} - 1 \right), \forall m \in \{m | \vartheta_{m,0} = 0\}. \quad (5.24)$$

综上, 资源分配算法的设计步骤总结在算法 4 中。此外, 基于迭代优化的卸载选择和资源分配的联合设计总结为算法 5。

<sup>5</sup>问题 (5.23) 的凸性证明请参阅附录 A.3.3

**算法 5**：联合卸载选择和资源分配设计算法

 输入:  $N, M, K, L, B, \sigma_w^2, P_m^{\max}, \lambda_m, D_m, C_m, \tau_m, F_m^U, \kappa_m^U, F_n^M, \kappa_n^M, \eta_l, \forall m, n, l$ 

 输出:  $\vartheta, \mathbf{p}, \mathbf{f}, \mathbf{F}$ 

- 1: 初始化 迭代索引  $t = 0$ ,  $p_m^{(t)} = P_m^{\max}, f_m^{\text{loc}(t)} = F_m^U, f_{m,n}^{\text{ser}(t)} = F_n^M / M, \forall m, n$ ;
- 2: **repeat**
- 3: 固定  $\mathbf{p}^{(t)}, \mathbf{f}^{(t)}$  和  $\mathbf{F}^{(t)}$ , 运行算法 3, 获得  $\vartheta^{(t+1)}$ ;
- 4: 固定  $\vartheta^{(t+1)}$ , 运行算法 4, 获得  $\mathbf{p}^{(t+1)}, \mathbf{f}^{(t+1)}$  和  $\mathbf{F}^{(t+1)}$ ;
- 5:  $t = t + 1$ ;
- 6: **until** 收敛

## 5.3.3 算法分析

本节分别从收敛性和复杂度两方面分析所提出的联合卸载选择和资源分配设计算法。

**收敛性分析:** 定义问题 (5.9) 的目标值在第  $t$  轮迭代时为  $\mathcal{F}(\vartheta^{(t)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)})$ 。根据算法 5 的第 3 步, 对于给定的  $\mathbf{p}^{(t)}, \mathbf{f}^{(t)}$  和  $\mathbf{F}^{(t)}$ , 利用算法 3 可获得最优解  $\vartheta^{(t+1)}$ , 则

$$\mathcal{F}(\vartheta^{(t)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)}) \geq \mathcal{F}(\vartheta^{(t+1)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)}). \quad (5.25)$$

接着, 根据算法 5 的第 4 步可得

$$\begin{aligned} \mathcal{F}(\vartheta^{(t+1)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)}) &\stackrel{(a)}{=} \mathcal{F}_{ub}(\vartheta^{(t+1)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)}) \\ &\stackrel{(b)}{\geq} \mathcal{F}_{ub}(\vartheta^{(t+1)}, \mathbf{p}^{(t+1)}, \mathbf{f}^{(t+1)}, \mathbf{F}^{(t+1)}) \\ &\stackrel{(c)}{\geq} \mathcal{F}(\vartheta^{(t+1)}, \mathbf{p}^{(t+1)}, \mathbf{f}^{(t+1)}, \mathbf{F}^{(t+1)}), \end{aligned} \quad (5.26)$$

其中  $\mathcal{F}_{ub}$  表示问题 (5.23) 的目标值。等式 (a) 成立是因为在给定  $\mathbf{p}^{(t)}, \mathbf{f}^{(t)}$  和  $\mathbf{F}^{(t)}$  的情况下, 问题 (5.23) 就变量  $\vartheta$  而言与原问题 (5.9) 相同。不等式 (b) 成立是因为问题 (5.23) 是凸的, 可以获得其最优解。不等式 (c) 成立在于使用了一阶泰勒近似, 即问题 (5.23) 的目标值是问题 (5.16) 的目标值的上界。结合式 (5.25) 和式 (5.26), 可以得出

$$\mathcal{F}(\vartheta^{(t)}, \mathbf{p}^{(t)}, \mathbf{f}^{(t)}, \mathbf{F}^{(t)}) \geq \mathcal{F}(\vartheta^{(t+1)}, \mathbf{p}^{(t+1)}, \mathbf{f}^{(t+1)}, \mathbf{F}^{(t+1)}), \quad (5.27)$$

这意味着加权总能耗在每次迭代后是非增的。考虑到感知任务和计算任务的性能要求, 问题 (5.9) 的目标值, 即加权总能耗, 是有下界的。因此, 算法 5 的收敛性可以得到保证。此外, 图 5.3 展示了所提算法不同终端数量下的收敛行为, 这也验证了算法 5 的收敛性。

**复杂度分析:** 由于算法 5 是由算法 3 和算法 4 组成的, 且每次迭代中的执行步骤

是相同的，因此可以重点分析每次迭代中算法 3 和算法 4 的计算复杂度。首先，算法 3 主要基于 **BnB** 算法，其计算复杂度可以表示为  $I_1 C_1^{\text{cost}}$ ，其中  $I_1$  指为寻找最优解所遍历 **BnB** 树节点的数量， $C_1^{\text{cost}}$  是在每个节点上解决相应的松弛优化问题 (5.15) 的计算复杂度。其次，由于使用了 **SCA** 技术，算法 4 也是迭代的，因此其计算复杂度可以表示为  $I_2 C_2^{\text{cost}}$ ，其中  $I_2$  为收敛所需的迭代次数， $C_2^{\text{cost}}$  指在每轮迭代中解决问题 (5.23) 的计算复杂度。此外，由于凸优化问题 (5.15) 和 (5.23) 均只有 **LMI** 约束条件，可以采取最差情况下使用内点法<sup>[107]</sup>求解的计算成本来刻画所提算法的计算复杂度<sup>[111]</sup>。具体而言，问题 (5.15) 有  $(MN + 2M + N)$  个维度为 1 的 **LMI** 约束， $L$  个维度为 3 的 **LMI** 约束和  $L$  个维度为 6 的 **LMI** 约束。因此，对于给定的精度  $\varepsilon_1 > 0$ ，使用内点法解决问题 (5.15) 在最差情况下的计算复杂度为  $C_1^{\text{cost}} = \sqrt{\Xi_1 + 9L} \ln(1/\varepsilon_1) [(\Xi_1 + 243L)n_1 + (\Xi_1 + 45L)n_1^2 + n_1^3]$ ，其中  $\Xi_1 = MN + 2M + N$ ，决策变量  $n_1 = \mathcal{O}(MN)$ 。相似地，问题 (5.23) 有  $3M + N$  个维度为 1 的 **LMI** 约束， $L$  个维度为 3 的 **LMI** 约束和  $L$  个维度为 6 的 **LMI** 约束。因此，对于给定的精度  $\varepsilon_2 > 0$ ，使用内点法解决问题 (5.23) 在最差情况下的计算复杂度为  $C_2^{\text{cost}} = \sqrt{\Xi_2 + 9L} \ln(1/\varepsilon_2) [(\Xi_2 + 243L)n_2 + (\Xi_2 + 45L)n_2^2 + n_2^3]$ ，其中  $\Xi_2 = 3M + N$ ，决策变量  $n_2 = \mathcal{O}(2M + 9L)$ 。

## 5.4 仿真结果

本节通过大量的数值仿真来验证所提出算法的有效性。不失一般性，假设所有终端、基站和目标随机分布在  $300 \text{ m} \times 300 \text{ m} \times 300 \text{ m}$  的三维空间内。路径损耗采用 **3GPP** 标准中的模型，即  $\text{PL}_{\text{dB}} = 128.1 + 37.6 \log_{10}(d)^{[112]}$ ，其中  $d$  (**km**) 表示发射机和接收机之间的距离。为了便于分析，假设所有终端对计算任务的执行时间要求相同，最大计算能力相同，发射功率预算相同，分别表示为  $\tau_m = \tau_0$ 、 $F_m^U = F_0^U$  和  $P_k^{\text{max}} = P_0$ 。同时假设对所有感知目标的定位要求相同，所有 **MEC** 服务器的最大计算能力相同，分别表示为  $\eta_l = \eta_0$  和  $F_n^M = F_0^M$ 。除非另有说明，默认的仿真参数根据表 5.1 设置。

首先，图 5.3 展示了算法 5（外循环）和算法 4（内循环）在不同终端数目下的收敛行为。可以看出，算法 5 几乎只需要 1 到 2 次迭代就可以收敛，算法 4 在 10 次迭代内即可收敛，这验证了所提的算法均具有较快的收敛速度，因此其在边缘智能网络中实现的成本是可以承受的。

然后，在图 5.4 中展示了所提的联合算法与 4 种基线算法在性能上的比较。四种基线算法分别为固定基站发射功率  $p_m = P_m^{\text{max}}/2$  的“固定发射功率算法”，固定计算能力  $f_m^{\text{loc}} = F_m^U$  和  $f_{m,n}^{\text{ser}} = F_n^M/M$  的“固定计算能力算法”，所有终端均随机选择卸载的“随机

表 5.1 算法 5 的仿真参数设置

| 参数        | 取值  |
|-----------|---|
| 基站数量      | $N = 3$   |
| 基站天线数     | $K = 64$  |
| 终端数量      | $M = 10$  |
| 目标数量      | $L = 2$   |
| 带宽        | $B = 10$ MHz  |
| 加权系数      | $\lambda_m = 1$                                     |
| 噪声方差      | $\sigma_w^2 = -110$ dBm                             |
| 终端的最大发射功率 | $P_0 = 30$ dBm                                      |
| 计算任务大小    | $D_m \in [40, 60]$ KB                               |
| 计算强度      | $C_m \in [100, 150]$ CPU cycles/bit                 |
| 计算能力      | $F_0^U = 1$ GHz, $F_0^M = 20$ GHz                   |
| 能量系数      | $\kappa_0^U = \kappa_0^M = 5 \times 10^{-29}$ [128] |
| QoS 要求    | $\tau_0 = 50$ ms, $\eta_0 = 0.05$ m <sup>2</sup>    |

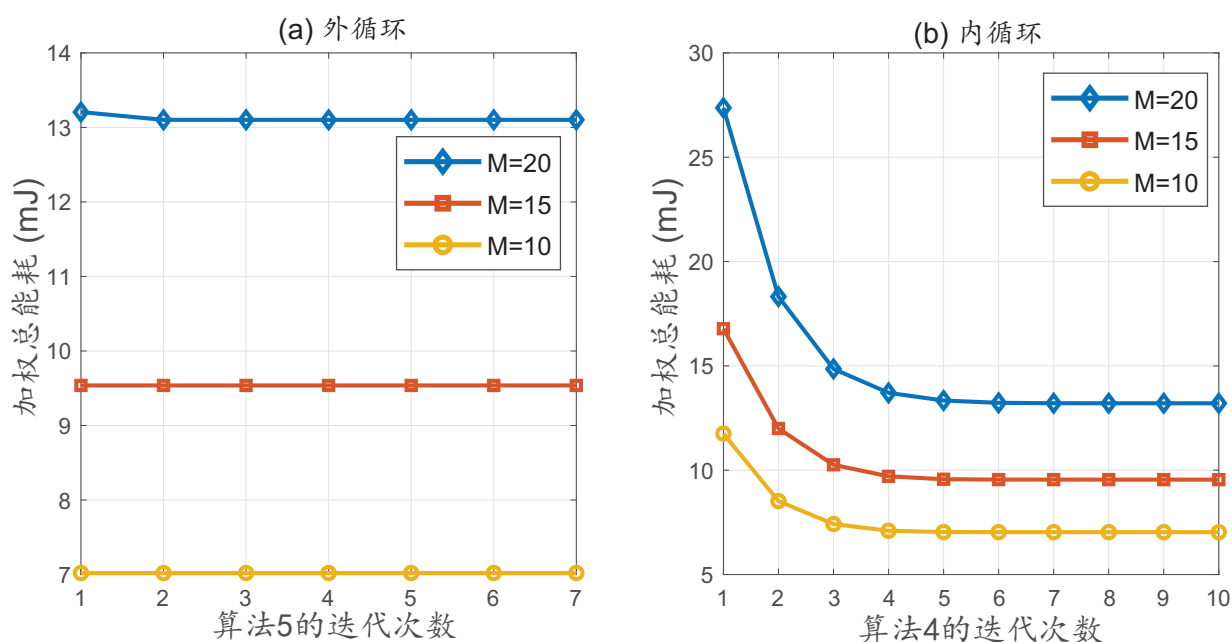


图 5.3 算法 5 和算法 4 的收敛行为

卸载选择算法”和没有考虑感知功能的 MEC 相关工作中的资源分配算法<sup>[61, 129]</sup>，并采取所提的卸载选择算法来满足定位需求的“现存资源分配算法”。从图中可见，算法 5 在所有算法中表现最好，现存资源分配算法表现中等，固定资源分配(包括发射功率和计算能力)算法表现最差，这验证了所提的联合算法的有效性。虽然对比算法的复杂度均小于所



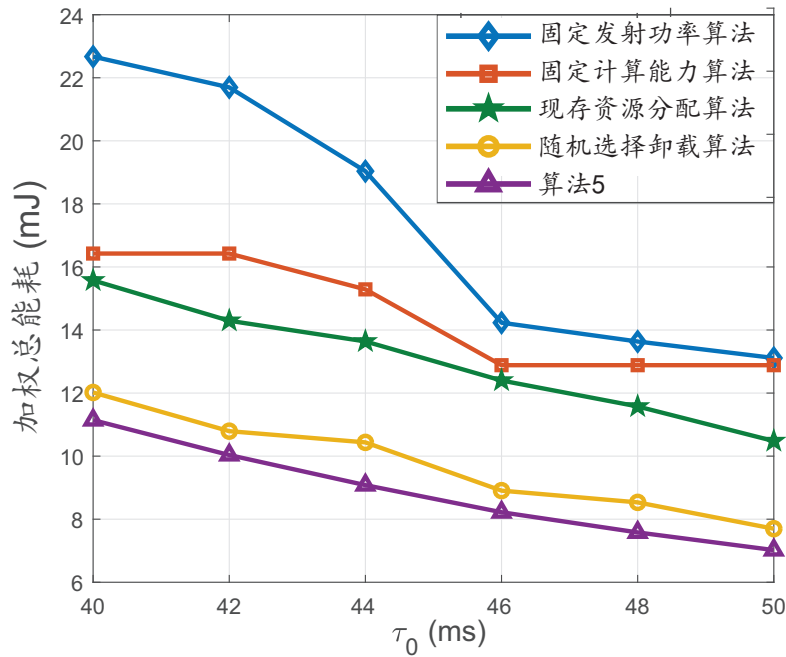


图 5.4 算法 5 与不同算法的性能比较

提算法，但是它们与所提算法的性能差距也较大。此外，随机卸载算法虽然在加权总能耗上接近所提的联合算法，但由于其没有很好地利用终端本地的计算资源，性能差距也是不可忽略的。相比于随机卸载算法，所提的联合算法提升不明显的原因在于卸载选择是典型的离散优化问题，相比于资源分配这样的连续优化问题，对整体性能的影响较小。

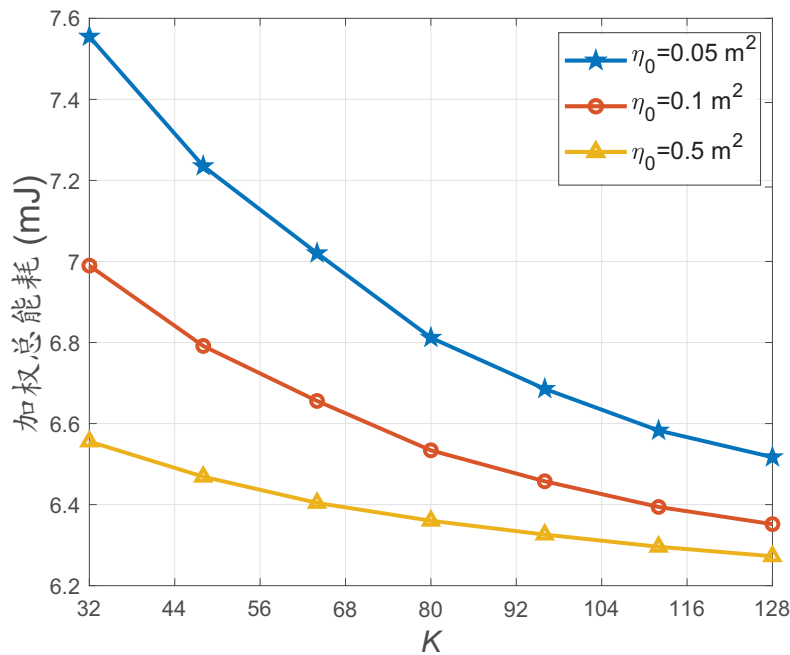


图 5.5 基站天线数和感知任务要求对加权总能耗的影响

图 5.5 研究了不同基站天线数量  $K$  和不同的定位精度要求  $\eta_0$  对加权总能耗的影响。可以看出，感知任务要求的阈值越小，加权总能耗越高。这是因为感知任务要求的阈值越小，对定位精度的要求就越高，为了保证高定位精度，需要更多的终端参与协作定位或发射更高的功率传输信号，因此总能耗也就越大。此外，加权总能量消耗随着基站天线数量的增加而减小。这是因为增加天线数量可以提高卸载计算任务的数据传输速率，同时提高感知目标的定位精度。然而，随着基站天线数量的增加，曲线的斜率变小，即利用增加基站天线所获得的性能增益变小。特别是，基站配备更多天线，会导致更高的射频链成本。因此，选择适当数量的基站天线可以在性能和成本之间取得平衡。

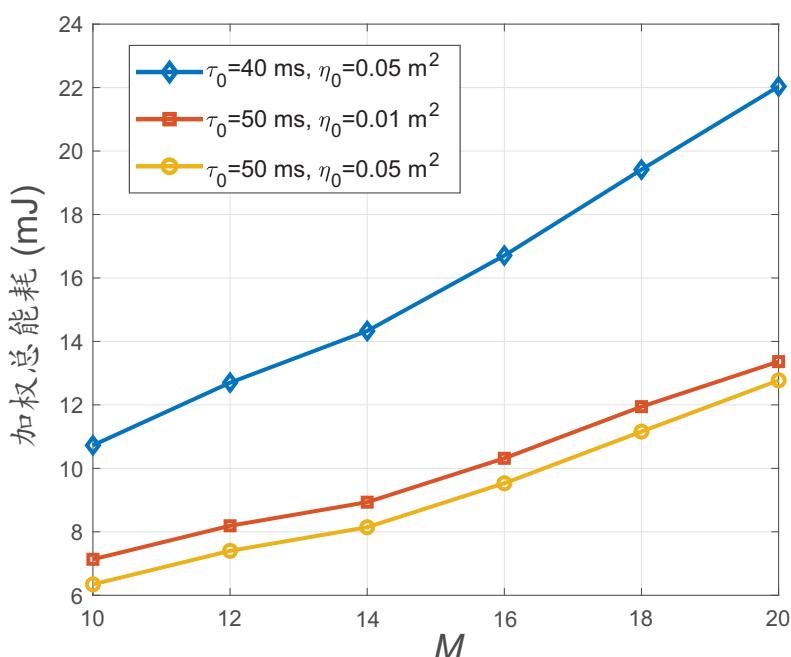


图 5.6 终端数量和 QoS 要求对加权总能耗的影响

接着，图 5.6 中给出了终端数量  $M$ 、计算任务的最大可容忍延迟  $\tau_0$  和感知任务的最大可容忍 CRB 矩阵的迹  $\eta_0$  对加权总能耗的影响，其中基站数量  $N = 4$ 。显然，加权总能耗随着终端数量的增加而增加。此外， $\tau_0 = 50$  ms 且  $\eta_0 = 0.05$  m<sup>2</sup> 时的加权总能耗最低， $\tau_0 = 40$  ms 且  $\eta_0 = 0.05$  m<sup>2</sup> 时的总能耗最高。换句话说，较小的  $\tau_0$  表示对计算任务的执行时间有更严格的限制，而  $\eta_0$  越大表示对感知目标的定位精度要求较低。因此，可以根据实际系统偏好适当放松或收紧感知任务和计算任务的性能要求，以达到期望的系统性能。

图 5.7 探究了计算任务最大的可容忍时延  $\tau_0$  和其数据大小  $D_m$  对加权总能耗和本地计算比例的影响。可以看出，随着最大可容忍时延的增加，加权总能耗减少，但本地计算的比例提高。这是因为  $\tau_0$  的值越高，意味着对计算任务的时间要求越宽松，导致选择本地计算的终端数量更多，总能耗就更少。此外，在相同的 QoS 要求下，计算任务的数据量越

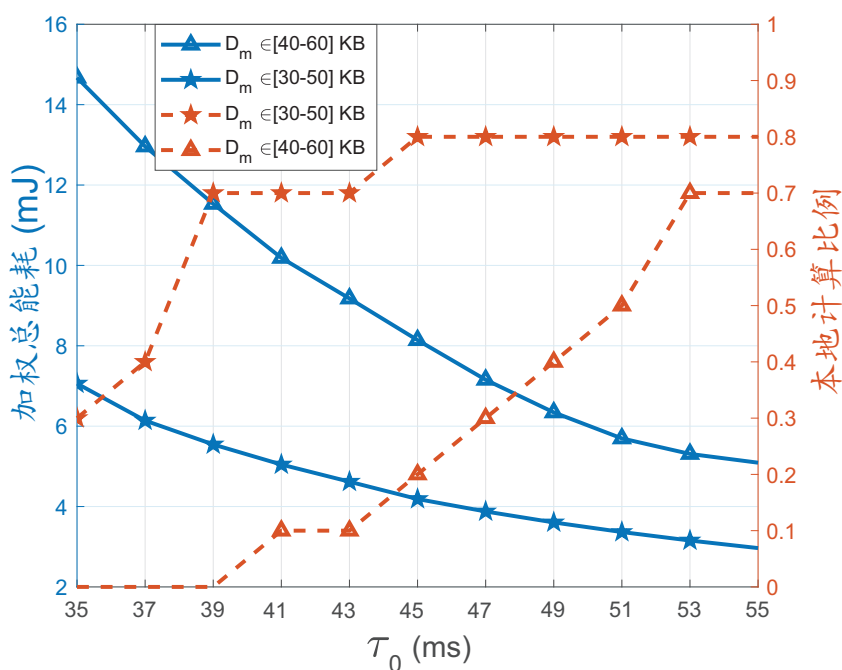


图 5.7 计算任务大小和时延要求对加权总能耗和本地计算比例的影响

大，执行本地计算的终端数量就越少，完成该任务所需的能量成本就越高。这说明一些小规模的计算任务可以在本地执行以节省总能耗，而大规模的计算任务则需要卸载到 MEC 服务器上以保证延迟需求。

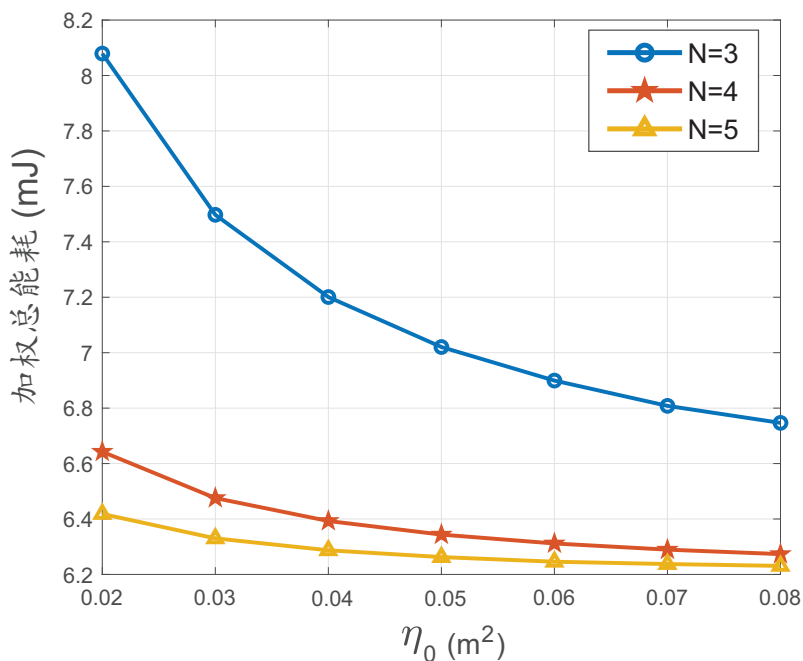


图 5.8 基站数量和感知任务要求对加权总能耗的影响

图 5.8 检验了最大可容忍 CRB 矩阵的迹  $\eta_0$  和基站数量  $N$  对系统性能的影响。可以看

出，加权总能耗随着定位要求的放宽而减小，即最大可容忍 **CRB** 矩阵的迹增大。此外，通过增加基站的数量可以有效降低加权总能耗。这是因为基站的增加可以提高定位精度，并为卸载计算任务提供了更多的计算资源，从而以更少的能耗来满足定位和计算的要求。另外，可以发现  $N = 4$  和  $N = 5$  之间的性能差距明显小于  $N = 3$  和  $N = 4$  之间的性能差距，这意味着通过增加基站数量来提高性能是有限的。因此，需要部署合适数量的基站，以平衡通信和感知一体化系统的成本和性能。

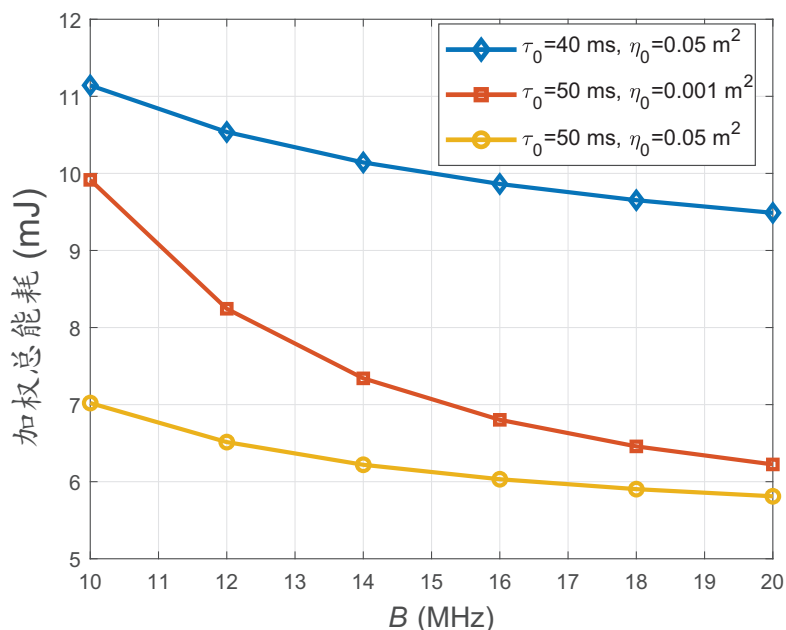


图 5.9 带宽在不同 QoS 要求下对加权总能耗的影响

最后，图 5.9 研究了不同 QoS 要求下加权总能耗与带宽  $B$  之间的关系。可以看出，加权总能耗随着带宽的增加而降低。这说明带宽大小对计算任务数据的传输速率和感知目标的定位精度有着很大的影响。此外，随着带宽的增加，要求为  $\tau_0 = 50$  ms 且  $\eta_0 = 0.001$  m<sup>2</sup> 的情况与要求为  $\tau_0 = 50$  ms 且  $\eta_0 = 0.05$  m<sup>2</sup> 的情况之间的性能差距减小，而要求为  $\tau_0 = 40$  ms 且  $\eta_0 = 0.05$  m<sup>2</sup> 的情况与要求为  $\tau_0 = 50$  ms 且  $\eta_0 = 0.05$  m<sup>2</sup> 的情况之间的性能差距基本保持不变。这意味着在整个带宽  $B$  的范围内放松计算任务的时延要求可以降低加权总能耗，特别是在较小带宽区域放松感知任务的定位要求相比于较大带宽区域可以更有效地降低加权总能耗。

## 5.5 本章小结

本章研究了通信辅助的感知和计算一体化设计。为了提高感知和计算一体化的整体性能，在满足计算任务执行时间和感知任务的定位精度要求的前提下，以最小化加权总能耗

为目标，提出了联合卸载选择和资源分配方案设计，该设计是一个计算困难的 **MINLP** 问题。为了解决这个问题，设计了一种基于迭代优化的算法来获得可行解。理论分析表明，该算法具有较快的收敛性能。此外，大量的仿真结果证实了该算法在边缘智能网络中的有效性。



## 6 功能复用的通信、感知和计算一体化

### 6.1 引言

在 6G 愿景的推动下，越来越多的智能终端和一体化网元部署在网络边缘，由此涌现出一批对性能有着极高要求的新兴智能业务，如无人驾驶、全息通信和扩展现实等。这些新兴智能业务，要求边缘智能网络及时提供信息通信、目标感知和智能计算的复合功能。在过去传统的系统中，通信、感知和计算功能是独立存在的，例如雷达系统只负责数据获取，通信系统只负责数据传输，云计算系统只负责数据处理。这种分离化设计存在频谱资源与硬件资源的浪费，功能相互独立也会带来信息处理时延较高的问题，从而无法满足未来新兴智能业务的需求。在这种情况下，需要研究通信、感知和计算功能的融合设计，以提高边缘智能网络的资源利用率。基于此，本章利用大规模 MIMO 技术，使得边缘智能网络在相同的无线资源上同时执行多目标感知、多维计算和多流通信。此外，配有大规模天线阵列的基站对接收到的各类混叠信号进行联合处理，从而实现三者功能的融合与复用。为了在有限的系统资源下实现三种不同功能的预期性能，需要设计相应的波束成形算法来协调干扰。基于此，本章首先分析和揭示了关键系统参数对通信、感知和计算一体化性能的影响，然后根据边缘智能网络的不同应用场景具有差异化需求的特点，制定了两种典型的联合发射和接收波束成形优化算法来实现通信、感知和计算三者性能之间的平衡。

内容概述：本章提出了一种功能复用的通信、感知和计算一体化架构，利用配置大规模天线阵列的基站所具有的空间自由度，在同一无线资源块上同时实现多目标感知、多维计算和多流通信。针对功能复用导致的多业务的同信道干扰和资源竞争问题，提出了两种联合发射和接收波束成形设计算法，以提高系统的总体性能和降低系统的总发射功率。具体结构安排如下：6.2 节介绍了功能复用的通信、感知和计算一体化系统模型，6.3 节从最大化加权总性能和最小化发射总功率的角度分别给出了两种联合发射波束和接收波束设计算法，6.4 进行了数值仿真以验证所提算法的有效性，6.5 节总结了本章内容。

## 6.2 系统模型

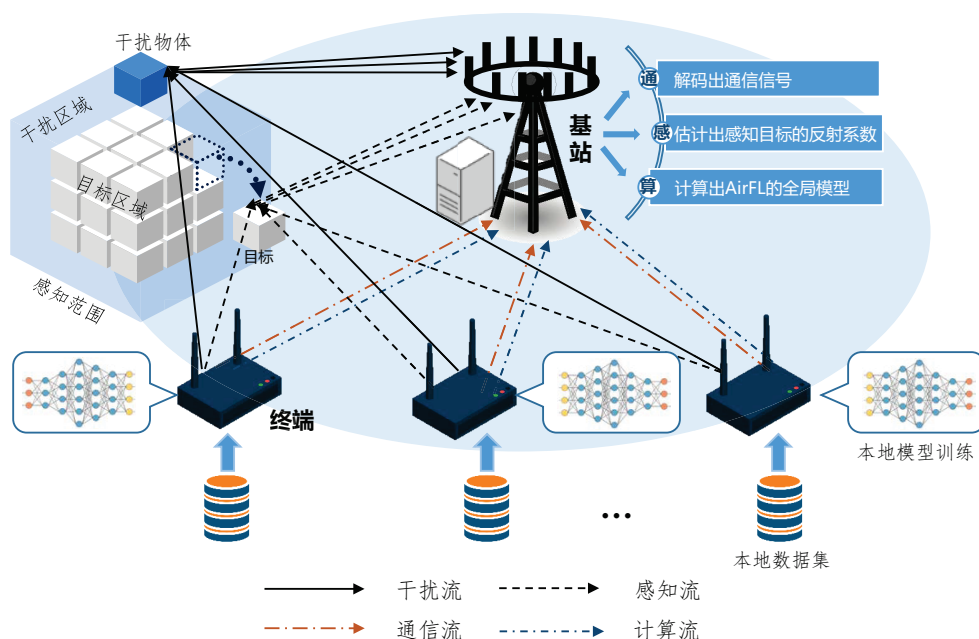


图 6.1 功能复用的通信、感知和计算一体化系统模型

如图 6.1 所示，考虑一个功能复用的通信、感知和计算一体化系统模型，它由配备  $N$  根天线的多功能基站和  $K$  个多功能智能终端组成，每个终端配备  $M$  根天线，其中  $KM \leq N$ 。将包括目标感知区域和干扰区域在内的兴趣范围（Range of Interest, ROI）划分为多个大小相等的立方体，每个立方体代表一个像素点<sup>1</sup> [130]。其中，感知区域内有  $I$  个需要被感知的目标，同时在干扰范围内有  $O$  个杂波作为干扰存在<sup>2</sup>。该系统有三项基本任务，即信息通信、目标感知和模型计算，需要由多个智能终端和多天线基站协同完成。具体而言，基于所获取的环境信息<sup>3</sup>，每个智能终端首先进行空间定向波束成形，以控制其通信信号、感知信号和计算信号的叠加发射方向，使信号根据信道条件自适应地到达基站。对于信息通信，基站设计相应的通信接收机，对各个智能终端传输的通信信号进行解码，以得到期望的通信数据。对于目标感知，基站接收到包含目标区域环境信息的反射信号，然后设计相应的感知接收机来估计目标的反射系数<sup>4</sup>。对于模型计算，通过利用基于空中计算的联邦学习（Over-the-Air Federated Learning, AirFL）架构，基站设计相应的计算

<sup>1</sup>像素的大小表示感知的精度。实际上，像素点不一定是立方体形式存在。根据感知精度的要求，可以基于信号波长、基站天线数、ROI 的范围大小和无线环境等因素来确定像素点的大小和形状。

<sup>2</sup>一般来说，杂波可以被认为是无线信道上的散射体。因此，杂波数量和信道状态信息可以通过信道估计获得<sup>[131]</sup>。

<sup>3</sup>预先通过环境检测可以粗略掌握环境信息。例如，通过信道估计可以得到信道条件<sup>[132]</sup>，通过波束扫描可以得到 ROI 的位置<sup>[133]</sup>。

<sup>4</sup>由于 ROI 中所有像素点的大小、形状和排列方式都相同，因此反射系数是唯一的识别特征。例如，基站可以通过准确获取的反射系数来进行目标分类<sup>[124]</sup>或目标成像<sup>[134]</sup>。



接收机，以聚合所有智能终端训练的本地模型，得到全局模型。

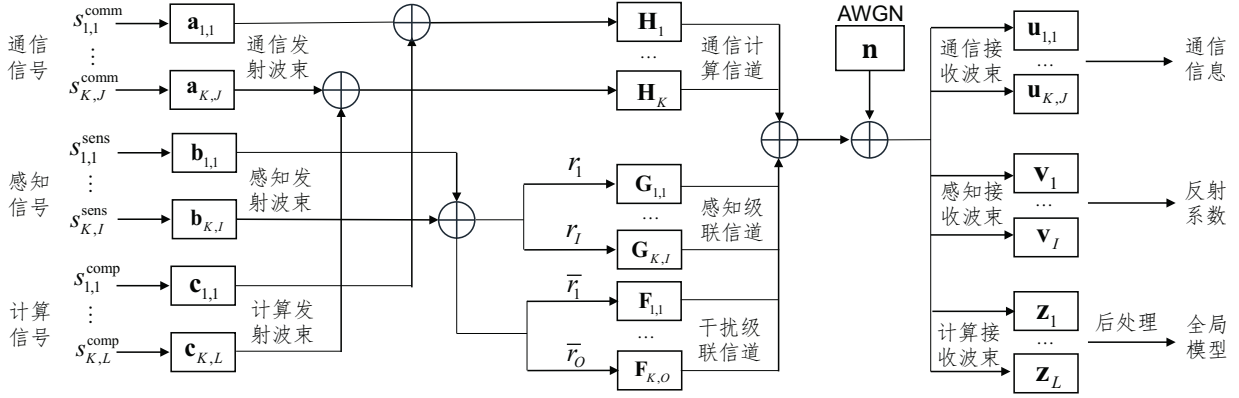


图 6.2 功能复用的通信、感知和计算一体化系统流程图

如图 6.2 所示，假设第  $k$  个终端生成数据  $\mathbf{s}_k^{\text{comm}} = [s_{k,1}^{\text{comm}}, \dots, s_{k,j}^{\text{comm}}, \dots, s_{k,J}^{\text{comm}}]^T \in \mathbb{R}^{J \times 1}$  作为通信信号来传递信息，同时构建感知信号  $\mathbf{s}_k^{\text{sens}} = [s_{k,1}^{\text{sens}}, \dots, s_{k,i}^{\text{sens}}, \dots, s_{k,I}^{\text{sens}}]^T \in \mathbb{R}^{I \times 1}$  去感知  $I$  个目标，并训练其数据集得到本地模型参数  $\mathbf{s}_k^{\text{comp}} = [s_{k,1}^{\text{comp}}, \dots, s_{k,l}^{\text{comp}}, \dots, s_{k,L}^{\text{comp}}]^T \in \mathbb{R}^{L \times 1}$  作为计算信号参与 AirFL。注意三种类型的信号的功能是不同的，其中  $s_{k,j}^{\text{comm}}$  是第  $k$  个终端上携带第  $j$  个测量数据的通信信号，它需要在基站处精确解码。 $s_{k,i}^{\text{sens}}$  是来自第  $k$  个终端用于第  $i$  个目标的感知信号，它不携带任何数据，只是用来估计目标的反射系数。 $s_{k,l}^{\text{comp}}$  是在第  $k$  个终端上经预处理的第  $l$  个本地模型参数，由终端训练其本地数据集生成。为了便于分析，假设通信信号、感知信号和计算信号是相互独立的，且均服从单位方差的高斯分布，即  $\mathbb{E} \{ s_{k,i}^{\text{sens}} (s_{k,i}^{\text{sens}})^H \} = \mathbb{E} \{ s_{k,l}^{\text{comp}} (s_{k,l}^{\text{comp}})^H \} = \mathbb{E} \{ s_{k,j}^{\text{comm}} (s_{k,j}^{\text{comm}})^H \} = 1, \forall k \in \Omega_K, i \in \Omega_I, l \in \Omega_L, j \in \Omega_J$ 。接着，第  $k$  个终端进行叠加编码，构造发射信号为

$$\mathbf{x}_k = \underbrace{\sum_{j=1}^J \mathbf{a}_{k,j} s_{k,j}^{\text{comm}}}_{\text{通信}} + \underbrace{\sum_{i=1}^I \mathbf{b}_{k,i} s_{k,i}^{\text{sens}}}_{\text{感知}} + \underbrace{\sum_{l=1}^L \mathbf{c}_{k,l} s_{k,l}^{\text{comp}}}_{\text{计算}}, \quad (6.1)$$

其中  $\mathbf{a}_{k,j} \in \mathbb{C}^{M \times 1}$ ,  $\mathbf{b}_{k,i} \in \mathbb{C}^{M \times 1}$ ,  $\mathbf{c}_{k,l} \in \mathbb{C}^{M \times 1}$  分别为通信信号、感知信号和计算信号的发射波束成形向量。然后，基站上接收到的混叠信号可以表示为

$$\mathbf{y} = \underbrace{\sum_{k=1}^K \sum_{j=1}^J \mathbf{H}_k \mathbf{a}_{k,j} s_{k,j}^{\text{comm}}}_{\text{通信}} + \underbrace{\sum_{k=1}^K \sum_{i=1}^I r_i \mathbf{G}_{k,i} \mathbf{b}_{k,i} s_{k,i}^{\text{sens}}}_{\text{感知}} + \underbrace{\sum_{k=1}^K \sum_{o=1}^O r_o \mathbf{F}_{k,o} \sum_{m=1}^I \mathbf{b}_{k,m} s_{k,m}^{\text{sens}}}_{\text{杂波干扰}}$$

$$+ \underbrace{\sum_{k=1}^K \sum_{l=1}^L \mathbf{H}_k \mathbf{c}_{k,l} S_{k,l}^{\text{comp}}}_{\text{计算}} + \underbrace{\mathbf{n}}_{\text{AWGN}}, \quad (6.2)$$

其中  $\mathbf{n}$  是均值为 0、方差为  $\sigma_n^2$  的 AWGN,  $r_i \in \mathbb{C}, \forall i \in \Omega_I$  和  $\bar{r}_o \in \mathbb{C}, \forall o \in \Omega_O$  分别是第  $i$  个目标和第  $o$  个杂波的反射系数。  $\mathbf{H}_k \in \mathbb{C}^{N \times M}$  是从第  $k$  个终端到基站的信道增益矩阵。  $\mathbf{G}_{k,i} = \mathbf{g}'_{k,i} \mathbf{g}_{k,i}^H \in \mathbb{C}^{N \times M}$  代表级联的感知信道,  $\mathbf{g}_{k,i}$  和  $\mathbf{g}'_{k,i}$  分别为从第  $k$  个终端到第  $i$  个目标和从第  $i$  个目标到基站的信道增益向量。类似地,  $\mathbf{F}_{k,o} = \mathbf{f}'_{k,o} \mathbf{f}_{k,o}^H \in \mathbb{C}^{N \times M}$  表示级联的干扰信道,  $\mathbf{f}_{k,o}$  和  $\mathbf{f}'_{k,o}$  分别为从第  $k$  个终端到第  $o$  个杂波和从第  $o$  个杂波到基站的信道增益向量。此外, 假设  $\mathbf{H}_k$ 、 $\mathbf{G}_{k,i}$  和  $\mathbf{F}_{k,o}$  均在一个时隙内保持不变, 但在连续时隙独立衰落。接下来, 逐一讨论信息通信、目标感知和模型计算三个基本任务的性能指标。

### 6.2.1 信息通信

为了提升通信信号的传输质量, 基站进行相应的通信接收波束成形来减轻干扰的影响。具体地, 在基站上经波束成形后接收到的通信信号  $s_{k,j}^{\text{comm}}, \forall k \in \Omega_K, j \in \Omega_J$  可以表示为

$$\begin{aligned} y_{k,j}^{\text{comm}} &= \mathbf{u}_{k,j}^H \mathbf{y} = \mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{c}_{k,j} S_{k,j}^{\text{comm}} + \sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J \mathbf{u}_{k,j}^H \mathbf{c}_{i,n} S_{i,n}^{\text{comm}} + \sum_{i=1}^K \sum_{l=1}^L \mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{b}_{i,l} S_{i,l}^{\text{comp}} \\ &+ \mathbf{u}_{k,j}^H \sum_{k=1}^K \sum_{i=1}^I r_i \mathbf{G}_{k,i} \mathbf{a}_{k,i} S_{k,i}^{\text{sens}} + \mathbf{u}_{k,j}^H \sum_{k=1}^K \sum_{o=1}^O \bar{r}_o \mathbf{F}_{k,o} \sum_{m=1}^I \mathbf{a}_{k,m} S_{k,m}^{\text{sens}} + \mathbf{u}_{k,j}^H \mathbf{n}, \end{aligned} \quad (6.3)$$

其中  $\mathbf{u}_{k,j} \in \mathbb{C}^{N \times 1}$  是关于来自第  $k$  个终端的第  $j$  个通信信号的通信接收波束成形向量。接收 SINR 决定了通信信号的质量, 因此通常被选择作为信息通信的性能指标。关于通信信号  $s_{k,j}^{\text{comm}}$  的接收 SINR 可以表示为

$$\Gamma_{k,j} = \frac{|\mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{a}_{k,j}|^2}{\sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J |\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{a}_{i,n}|^2 + X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2}, \quad \forall k, j, \quad (6.4)$$

其中,  $X_{k,j} = \sum_{i=1}^K \sum_{m=1}^I R_m^2 |\mathbf{u}_{k,j}^H \mathbf{G}_{i,m} \mathbf{b}_{i,m}|^2 + \sum_{i=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 |\mathbf{u}_{k,j}^H \mathbf{F}_{i,o} \mathbf{b}_{i,m}|^2 + \sum_{i=1}^K \sum_{l=1}^L |\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{c}_{i,l}|^2$ ,  $R_i$  和  $\bar{R}_o$  分别是基于先验概率得到的目标和杂波反射系数  $r_i$  和  $\bar{r}_o$  的均方根 (Root Mean Square, RMS) [124]。

### 6.2.2 目标感知

为了提高目标感知的精度，所有终端对目标进行协作感知。接着，在基站处部署一个线性无偏估计器来估计目标的反射系数。为了获得精确的反射系数  $r_i, i \in \Omega_I$ ，基站进行相应的感知接收波束成形，以增强期望的信号同时抑制同频干扰。因此，估计的第  $i$  个目标的反射系数  $\hat{r}_i$  可以表示为

$$\begin{aligned}\hat{r}_i &= \mathbf{v}_i^H \mathbf{y} \\ &= \mathbf{v}_i^H r_i \sum_{k=1}^K \mathbf{G}_{k,i} \mathbf{b}_{k,i} s_{k,i}^{\text{sens}} + \mathbf{v}_i^H \sum_{k=1}^K \sum_{m=1, m \neq i}^I r_m \mathbf{G}_{k,m} \mathbf{a}_{k,m} s_{k,m}^{\text{sens}} + \mathbf{v}_i^H \sum_{k=1}^K \sum_{o=1}^O \bar{r}_o \mathbf{F}_{k,o} \sum_{m=1}^I \mathbf{b}_{k,m} s_{k,m}^{\text{sens}} \\ &\quad + \mathbf{v}_i^H \sum_{k=1}^K \mathbf{H}_k \left( \sum_{l=1}^L \mathbf{c}_{k,l} s_{k,l}^{\text{comp}} + \sum_{j=1}^J \mathbf{a}_{k,j} s_{k,j}^{\text{comm}} \right) + \mathbf{v}_i^H \mathbf{n}, \forall i,\end{aligned}\quad (6.5)$$

其中  $\mathbf{v}_i \in \mathbb{C}^{N \times 1}$  是关于第  $i$  个目标的感知接收波束成形向量。为了评估感知任务的性能，采用 **MSE** 作为目标感知的性能指标。从式 (6.5) 可以看出  $\hat{r}_i$  不仅是  $s_{k,i}^{\text{sens}}$  的函数，还是来自不同终端对于不同目标的其他感知信号、计算信号和通信信号的函数。在这种情况下，需要最小化估计的反射系数  $\hat{r}_i$  和真实的反射系数  $r_i$  之间的 **MSE** 来抑制其他信号的干扰，实现高精度的目标感知。具体地，第  $i$  个目标的感知 **MSE** 可以表示为

$$\begin{aligned}\text{MSE}_i^{\text{sens}} &= \mathbb{E} \left\{ (\hat{r}_i - r_i) (\hat{r}_i - r_i)^H \right\} \\ &= R_i^2 \left| \sum_{k=1}^K \mathbf{v}_i^H \mathbf{G}_{k,i} \mathbf{b}_{k,i} - 1 \right|^2 \\ &\quad + \sum_{k=1}^K \sum_{m=1, m \neq i}^I R_m^2 |\mathbf{v}_i^H \mathbf{G}_{k,m} \mathbf{b}_{k,m}|^2 + \sum_{k=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 |\mathbf{v}_i^H \mathbf{F}_{k,o} \mathbf{b}_{k,m}|^2 \\ &\quad + \sum_{k=1}^K \left( \sum_{l=1}^L |\mathbf{v}_i^H \mathbf{H}_k \mathbf{c}_{k,l}|^2 + \sum_{j=1}^J |\mathbf{v}_i^H \mathbf{H}_k \mathbf{a}_{k,j}|^2 \right) + \sigma_n^2 \|\mathbf{v}_i^H\|^2, \forall i.\end{aligned}\quad (6.6)$$

为了更直观地表达反射系数估计精度的影响，在此展示一个实际应用案例，即目标成像<sup>[134]</sup>。如图 6.3 所示，感知精度越高，即估计反射系数与实际反射系数之间的 **MSE** 越小，目标成像精度越高。

### 6.2.3 模型计算

对于模型计算，利用 **AirFL** 架构，多个智能终端和基站协同训练一个全局模型

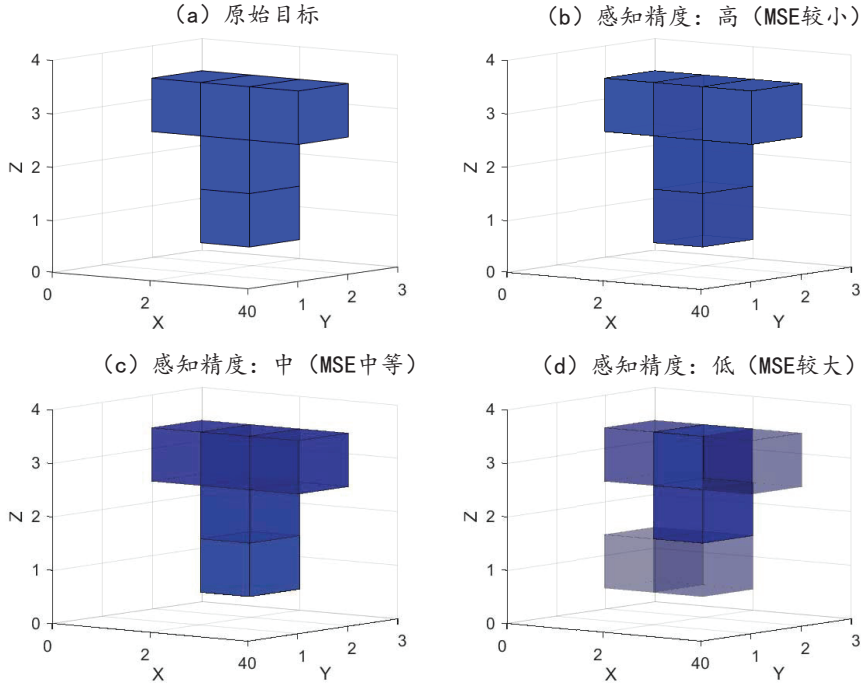


图 6.3 感知精度对目标成像效果的影响示例

$\mathbf{s}^{\text{comp}} \in \mathbb{R}^{L \times 1}$ 。定义每个终端的本地数据集为  $\mathcal{D}_k, \forall k \in \Omega_K$ ，本地数据集总和为  $\mathcal{D}$ ，其中  $|\mathcal{D}| = \sum_{k=1}^K |\mathcal{D}_k|$ 。基于 AirFL 的原理，为了促使全局模型的融合，每个本地模型参数需要在传输之前乘上一个权重系数  $\xi_k = \frac{|\mathcal{D}_k|}{|\mathcal{D}|}$  进行预处理。因此，在无干扰条件下理想的全局模型可以表示为

$$\mathbf{s}_l^{\text{comp}} = \frac{1}{K} \sum_{k=1}^K \mathbf{s}_{k,l}^{\text{comp}}, \forall l \in \Omega_L \quad (6.7)$$

其中  $\mathbf{s}_{k,l}^{\text{comp}}$  为第  $k$  个终端上产生的经预处理的第  $l$  个本地模型参数相关的计算信号。为了减小在干扰和信道衰落情况下的计算误差，需要在基站上进行计算接收波束成形。因此，在基站上接收的计算信号可以表示为

$$\begin{aligned}
 \hat{\mathbf{s}}_l^{\text{comp}} &= \frac{\mathbf{z}_l^H}{K} \mathbf{y} = \frac{\mathbf{z}_l^H}{K} \sum_{k=1}^K \mathbf{H}_k \mathbf{c}_{k,l} \mathbf{s}_{k,l}^{\text{comp}} + \frac{\mathbf{z}_l^H}{K} \sum_{k=1}^K \sum_{i=1, i \neq l}^L \mathbf{H}_k \mathbf{c}_{k,i} \mathbf{s}_{k,i}^{\text{comp}} + \frac{\mathbf{z}_l^H}{K} \sum_{k=1}^K \sum_{i=1}^I r_i \mathbf{G}_{k,i} \mathbf{b}_{k,i} \mathbf{s}_{k,i}^{\text{sens}} \\
 &+ \frac{\mathbf{z}_l^H}{K} \sum_{k=1}^K \sum_{o=1}^O \bar{r}_o \mathbf{F}_{k,o} \sum_{m=1}^I \mathbf{b}_{k,m} \mathbf{s}_{k,m}^{\text{sens}} + \frac{\mathbf{z}_l^H}{K} \sum_{k=1}^K \sum_{j=1}^J \mathbf{H}_k \mathbf{a}_{k,j} \mathbf{s}_{k,j}^{\text{comm}} + \frac{\mathbf{z}_l^H}{K} \mathbf{n}, \forall l, \quad (6.8)
 \end{aligned}$$




其中  $\mathbf{z}_l \in \mathbb{C}^{N \times 1}$  是关于第  $l$  个本地模型参数的计算接收波束成形向量。一般来说，模型聚

合的失真程度可以由  $s_l^{\text{comp}}$  和  $\hat{s}_l^{\text{comp}}$  之间的 MSE 来衡量，即

$$\begin{aligned}
\text{MSE}_l^{\text{comp}} &= \mathbb{E} \left\{ (\hat{s}_l^{\text{comp}} - s_l^{\text{comp}}) (\hat{s}_l^{\text{comp}} - s_l^{\text{comp}})^H \right\} \\
&= \frac{1}{K^2} \sum_{k=1}^K |\mathbf{z}_l^H \mathbf{H}_k \mathbf{c}_{k,l} - 1|^2 + \frac{1}{K^2} \sum_{m=1, m \neq l}^L \sum_{k=1}^K |\mathbf{z}_l^H \mathbf{H}_k \mathbf{c}_{k,m}|^2 \\
&+ \frac{1}{K^2} \sum_{k=1}^K \sum_{i=1}^I R_i^2 |\mathbf{z}_l^H \mathbf{G}_{k,i} \mathbf{b}_{k,i}|^2 + \frac{1}{K^2} \sum_{k=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 |\mathbf{z}_l^H \mathbf{F}_{k,o} \mathbf{b}_{k,m}|^2 \\
&+ \frac{1}{K^2} \sum_{k=1}^K \sum_{j=1}^J |\mathbf{z}_l^H \mathbf{H}_k \mathbf{a}_{k,j}|^2 + \frac{\sigma_n^2}{K^2} \|\mathbf{z}_l\|^2, \forall l.
\end{aligned} \tag{6.9}$$

在此，也同样提供一个经典的应用案例，即基于 AirFL 的图像识别来更直观地展示模型聚合误差对计算性能的影响<sup>[15]</sup>。从表 6.1 中可以看到计算精度越高，即经无线信道传输聚合后的全局模型与期望的全局模型之间的 MSE 越小，图像识别越准确。

表 6.1 计算精度对图像识别效果的影响示例

|               |   |   |   |  |   |   |   |
|---------------|---|---|---|--|---|---|---|
| 数字图像          |  |  |  |  |  |  |  |
| 计算精度：高(MSE小)  | 0   | 5   | 1   | 9  | 4   | 6   | 4   |
| 计算精度：中(MSE中等) | 6(×)  | 5   | 1   | 9  | 4   | 6   | 4   |
| 计算精度：低(MSE大)  | 0   | 8(×)  | 1   | 4(×)   | 8(×)  | 4(×)  | 4   |
| 数字图像          |  |  |  |  |  |  |  |
| 计算精度：高(MSE小)  | 7   | 1   | 8   | 5  | 2   | 9   | 3   |
| 计算精度：中(MSE中等) | 7   | 1   | 8   | 5  | 2   | 4(×)  | 3   |
| 计算精度：低(MSE大)  | 4(×)  | 1   | 8   | 8(×)   | 2   | 4(×)  | 8(×)  |

观察式 (6.4)、(6.6) 和 (6.9)，不难发现通信、感知和计算的性能由终端上的发射波束成形向量  $\{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}\}$  和基站上的接收波束成形向量  $\{\mathbf{u}_{k,j}, \mathbf{v}_i, \mathbf{z}_l\}$  共同决定。虽然通信、感知和计算三个功能的高性能都是系统所期望的，但是在有限的系统资源下，它们之间具有竞争关系。因此，需要进行联合发射和接收波束成形设计，以有限的边缘智能网络资源提高通信、感知和计算一体化的整体性能。

### 6.3 联合发射和接收波束成形设计

本节旨在联合优化终端和基站上的发射和接收波束。由于边缘智能网络在不同应用场景下对通信、感知和计算的性能有差异化的需求，本节设计了两种典型的 MOOP 来实现

它们三者性能之间的权衡。第一个 **MOOP** 是在发射功率有限的情况下，最大化系统加权总性能。第二个 **MOOP** 是最小化系统的发射总功率，同时保证通信、感知和计算的 **QoS** 要求。

### 6.3.1 加权总性能最大化设计

在进行加权总性能最大化设计之前，先研究它的三个基本功能模块，即关于通信、感知和计算单方面性能优化的 **SOOP**。第一个 **SOOP** 的目标是在终端发射功率的限制下，最大化通信速率加权和，即

*S-1'*: 通信速率加权和最大化

$$\begin{aligned} \max_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{u}_{k,j}} \quad & \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} r_{k,j} \\ \text{s.t. C1:} \quad & \sum_{i=1}^J \|\mathbf{a}_{k,j}\|^2 + \sum_{l=1}^L \|\mathbf{b}_{k,i}\|^2 + \sum_{j=1}^L \|\mathbf{c}_{k,l}\|^2 \leq P_{\max,k}, \forall k \in \Omega_K, \end{aligned} \quad (6.10)$$

其中  $r_{k,j} = \log_2(1 + \Gamma_{k,j})$  是来自第  $k$  个终端的第  $j$  个通信信号的传输速率， $\theta_{k,j}^{\text{comm}}$  是关于  $r_{k,j}$  的权重系数， $P_{\max,k}$  是第  $k$  个终端的最大发射功率预算。

**注6.1.** 每个 **SOOP** 只关注通信、感知或计算中的一个功能的性能，而不考虑其他两个功能的性能。例如，从问题 *S-1'* 可以看出感知发射波束  $\mathbf{b}_{k,i}$  和计算发射波束  $\mathbf{c}_{k,l}$  会自然取零来使得通信速率加权和最大化。因此，在通信 **SOOP** 中无需考虑感知和计算性能的约束条件。同样的，在感知 **SOOP** 和计算 **SOOP** 中均不需要考虑其他功能的性能要求。

不难发现，问题 *S-1'* 是一个经典的通信传输速率加权和最大化问题<sup>[135]</sup>。由于目标函数涉及分数结构的对数函数，且发射波束成形向量  $\{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}\}$  和通信接收波束成形向量  $\mathbf{u}_{k,j}$  以耦合的形式存在于约束条件和目标函数中，该问题是非凸的。为了解决这些问题，提出以下定理将问题 *S-1'* 转化成等价的 **MSE** 的对数最小化问题。

**定理6.1.** 对于通信信号  $s_{k,j}^{\text{comm}}$ ，其接收 **SINR**  $\Gamma_{k,j}$  与最小均方误差 (*Minimum MSE, MMSE*)  $e_{k,j}^{\text{comm}}$  的关系可以表示为

$$1 + \Gamma_{k,j} = (e_{k,j}^{\text{comm}})^{-1}, \forall k, j. \quad (6.11)$$

证明: 请参阅附录 A.4.1。  $\square$

根据定理 6.1, 可以将问题  $S-I'$  的目标函数转换为

$$\min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{u}_{k,j}} \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} \log_2 (e_{k,j}^{\text{comm}}). \quad (6.12)$$

使用附录 A.4.1 中推导的 MMSE 接收机, 即  $\mathbf{u}_{k,j} = \Xi^{-1} \mathbf{H}_k \mathbf{a}_{k,j}$ , 转换后的目标函数 (6.12) 等价于最小化通信 MSE 加权和, 即

$$\min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}} \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} \log_2 (\text{MSE}_{k,j}^{\text{comm}}). \quad (6.13)$$

然而, 由于对数函数和的结构, 再次转换后的目标函数 (6.13) 仍然保持非凸性。为此, 引入一个关于  $\text{MSE}_{k,j}^{\text{comm}}$  的权重变量  $\omega_{k,j}$ <sup>[136]</sup>, 将目标函数 (6.13) 进一步转化为

$$\min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \omega_{k,j}} \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} [\omega_{k,j} \text{MSE}_{k,j}^{\text{comm}} - \log_2 (\omega_{k,j})]. \quad (6.14)$$

注意到关于权重变量  $\omega_{k,j}$  的最小化 MSE 问题是一个无约束的优化问题, 因此令目标函数 (6.14) 关于  $\omega_{k,j}$  的一阶导数为零, 可以得到最优的权重变量为

$$\omega_{k,j}^* = (\ln 2 \cdot \text{MSE}_{k,j}^{\text{comm}})^{-1}, \quad (6.15)$$

这使得目标函数 (6.13) 与 (6.14) 等价。经过上述转换, 问题  $S-I'$  可以被替代为

$S-I$ : 转换后的通信 MSE 加权和最小化

$$\begin{aligned} \min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{u}_{k,j}, \omega_{k,j}} & \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} [\omega_{k,j} \text{MSE}_{k,j}^{\text{comm}} - \log_2 (\omega_{k,j})] \\ \text{s.t.} & \text{ C1.} \end{aligned} \quad (6.16)$$

接着, 第二个 SOOP 的目标为在终端发射功率的限制下, 最小化感知 MSE 的加权和, 即

$S-2$ : 感知 MSE 加权和最小化

$$\min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{v}_i} \sum_{i=1}^I \theta_i^{\text{sens}} \text{MSE}_i^{\text{sens}}$$

$$\text{s.t. C1,} \quad (6.17)$$

其中  $\theta_i^{\text{sens}}$  是关于  $\text{MSE}_i^{\text{sens}}$  的权重系数， $\text{MSE}_i^{\text{sens}}$  为第  $i$  个目标的估计反射系数和真实反射系数之间的 MSE，其定义在式 (6.6) 中。

相似地，第三个 SOOP 是在有限的发射功率约束下，最小化计算 MSE 的加权和，即

*S-3: 计算MSE加权和最小化*

$$\begin{aligned} \min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{z}_l} \quad & \sum_{l=1}^L \theta_l^{\text{comp}} \text{MSE}_l^{\text{comp}} \\ \text{s.t.} \quad & \text{C1,} \end{aligned} \quad (6.18)$$

其中  $\theta_l^{\text{comp}}$  是关于  $\text{MSE}_l^{\text{comp}}$  的权重系数， $\text{MSE}_l^{\text{comp}}$  为聚合的计算信号与期望的计算信号之间的 MSE，其定义在式 (6.9) 中。

为了联合优化通信、感知和计算的性能，使用加权和的方法构造相应的 MOOP<sup>[137]</sup>，其可以表述为

*M-1: 加权总性能最大化*

$$\begin{aligned} \min_{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{v}_i, \mathbf{u}_{k,j}, \mathbf{z}_l, \omega_{k,j}} \quad & \alpha_1 \Psi_1 + \alpha_2 \Psi_2 + \alpha_3 \Psi_3 \\ \text{s.t.} \quad & \text{C1,} \end{aligned} \quad (6.19)$$

其中  $\Psi_p, p = 1, 2, 3$  分别为问题 *S-1*、*S-2* 和 *S-3* 经归一化后的目标函数<sup>5</sup>， $\alpha_p \geq 0$  是第  $p$  个目标函数的权重，代表了系统优化的偏好，并且满足  $\alpha_1 + \alpha_2 + \alpha_3 = 1$ 。通过改变权重  $\alpha_p$  的值，求解问题 *M-1* 可以得到不同的解。注意到当  $\alpha_p = 1$  且  $\alpha_q = 0, \forall p \neq q$  时，问题 *M-1* 等价于问题 *S-p*，这意味着问题 *M-1* 是问题 *S-1*、*S-2* 和 *S-3* 的一般形式。因此，在给定功能优先级  $\alpha_p$  的情况下，求解问题 *M-1* 可以取得期望的系统性能。

由于多个优化变量在问题 *M-1* 的目标函数中耦合，即发射波束  $\{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}\}$ 、接收波束  $\{\mathbf{u}_{k,j}, \mathbf{v}_i, \mathbf{z}_l\}$  和权重变量  $\{\omega_{k,j}\}$ ，问题 *M-1* 是非凸的，难以在多项式时间内得到其最优解。幸运的是，尽管目标函数不是所有变量的联合凸函数，但分别仅考虑发射波束、接收波束或权重变量时，它是一个凸函数。基于这个观察，可以将问题 *M-1* 解耦为三个子问题，即通过固定发射波束和权重变量来优化接收波束，通过固定发射波束和接收波束来优

<sup>5</sup>由于问题 *S-1*、*S-2* 和 *S-3* 的目标值有不同的取值范围，需要对每个问题的目标函数执行归一化操作，以协调通信、感知和计算性能，并促进问题 *M-1* 目标值的收敛。为了方便表示，定义  $\mathcal{F}_p, p = 1, 2, 3$ ，分别为问题 *S-p* 的目标函数。同时，定义  $\Psi_p = (\mathcal{F}_p - \mathcal{F}_p^*) / |\mathcal{F}_p^*|$  为归一化的目标函数，其中  $\mathcal{F}_p^*$  为问题 *M-1* 中通信、感知和计算的性能极限，可以分别使用解决问题 *M-1* 的算法去解决问题 *S-p* 来获得。



化权重变量，以及通过固定接收波束和权重变量来优化发射波束。通过在迭代中依次求解这三个子问题，当目标值收敛时，可以获得一个可行的次优解<sup>[138]</sup>。接下来，分别处理这三个子问题。

首先，对于优化接收波束  $\{\mathbf{u}_{k,j}, \mathbf{v}_i, \mathbf{z}_l\}$  的子问题，通过对问题 *M-1* 应用 KKT 条件，推导出最优接收波束的必要条件为

$$\begin{aligned}\frac{\partial}{\partial \mathbf{u}_{k,j}} \left( \sum_{p=1}^3 \alpha_p \Psi_p \right) &= \mathbf{0} \Rightarrow \frac{\partial}{\partial \mathbf{u}_{k,j}} (\text{MSE}_{k,j}^{\text{comm}}) = \mathbf{0}. \\ \frac{\partial}{\partial \mathbf{v}_i} \left( \sum_{p=1}^3 \alpha_p \Psi_p \right) &= \mathbf{0} \Rightarrow \frac{\partial}{\partial \mathbf{v}_i} (\text{MSE}_i^{\text{sens}}) = \mathbf{0}. \\ \frac{\partial}{\partial \mathbf{z}_l} \left( \sum_{p=1}^3 \alpha_p \Psi_p \right) &= \mathbf{0} \Rightarrow \frac{\partial}{\partial \mathbf{z}_l} (\text{MSE}_l^{\text{comp}}) = \mathbf{0}.\end{aligned}\quad (6.20)$$

基于式 (6.20)，可以获得最优的接收波束，也称为 MMSE 接收波束，即

$$\mathbf{u}_{k,j} = \Xi^{-1} \mathbf{H}_k \mathbf{a}_{k,j}, \quad (6.21)$$

$$\mathbf{v}_i = \left[ R_i^2 \left( \sum_{k=1}^K \mathbf{G}_{k,i} \mathbf{b}_{k,i} \right) \left( \sum_{k=1}^K \mathbf{b}_{k,i}^H \mathbf{G}_{k,i}^H \right) + \sum_{k=1}^K \sum_{n=1, n \neq i}^I R_n^2 \mathbf{G}_{k,n} \mathbf{b}_{k,n} \mathbf{b}_{k,n}^H \mathbf{G}_{k,n}^H + \Phi \right]^{-1} R_i^2 \sum_{k=1}^K \mathbf{G}_{k,i} \mathbf{b}_{k,i}, \quad (6.22)$$

和

$$\mathbf{z}_l = \Xi^{-1} \sum_{k=1}^K \mathbf{H}_k \mathbf{c}_{k,l}, \quad (6.23)$$

$$\begin{aligned}\text{其中, } \Phi &= \sum_{k=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 \mathbf{F}_{k,o} \mathbf{a}_{k,m} \mathbf{a}_{k,m}^H \mathbf{F}_{k,o}^H + \sum_{k=1}^K \sum_{l=1}^L \mathbf{H}_k \mathbf{b}_{k,l} \mathbf{b}_{k,l}^H \mathbf{H}_k^H + \sum_{k=1}^K \sum_{m=1}^J \mathbf{H}_k \mathbf{c}_{k,m} \mathbf{c}_{k,m}^H \mathbf{H}_k^H \\ &+ \sigma_n^2 \mathbf{I}_N, \quad \Xi = \Phi + \sum_{k=1}^K \sum_{n=1}^I R_n^2 \mathbf{G}_{k,n} \mathbf{b}_{k,n} \mathbf{b}_{k,n}^H \mathbf{G}_{k,n}^H.\end{aligned}$$

接着，对于优化权重变量  $\{\omega_{k,j}\}$  的子问题，最优解已在式 (6.15) 中给出，即  $\omega_{k,j}^* = (\ln 2 \cdot \text{MSE}_{k,j}^{\text{comm}})^{-1}$ 。

最后，对于给定的 MMSE 接收波束和权重变量，优化发射波束  $\{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}\}$  的子问题是一个典型的二次约束二次规划 (Quadratically Constrained Quadratic Programming, QCQP) 问题，可以利用内点法来求解。具体而言，通过在目标函数中添加对数障碍函数，将该问题转化为一个无约束的凸优化问题。转换后的无约束凸优化问题可以表述为

$$\min_{\mathbf{a}_{k,i}, \mathbf{b}_{k,l}, \mathbf{c}_{k,j}} \varepsilon(\alpha_1 \Psi_1 + \alpha_2 \Psi_2 + \alpha_3 \Psi_3) - \sum_{k=1}^K \log(-f_k^{\text{con}}), \quad (6.24)$$

其中  $\varepsilon > 0$  是障碍参数,  $f_k^{\text{con}} = \sum_{i=1}^I \|\mathbf{a}_{k,i}\|^2 + \sum_{l=1}^L \|\mathbf{b}_{k,l}\|^2 + \sum_{j=1}^J \|\mathbf{c}_{k,j}\|^2 - P_{\max,k}$ . 然后, 采用牛顿迭代法来获得该问题的最优解, 其需要用到目标函数关于变量的梯度  $\nabla_{\mathbf{x}}$  和海森矩阵  $\nabla_{\mathbf{x}}^2$ , 它们分别为

$$\begin{aligned}\nabla_{\mathbf{a}_{k,j}} &= 2\varepsilon \left( \mathbf{T}_{k,l}^a \mathbf{a}_{k,j} - \tilde{\alpha}_3 \theta_{k,j}^{\text{comm}} \omega_{k,j} \mathbf{H}_k^H \mathbf{u}_{k,j} \right) + \frac{2\mathbf{a}_{k,j}}{f_k^{\text{con}}}, \\ \nabla_{\mathbf{b}_{k,i}} &= 2\varepsilon \left( \mathbf{T}_{k,i}^b \mathbf{b}_{k,i} - R_i^2 \tilde{\alpha}_1 \theta_i^{\text{sens}} \mathbf{G}_{k,i}^H \mathbf{v}_i \right) + \frac{2\mathbf{b}_{k,i}}{f_k^{\text{con}}}, \\ \nabla_{\mathbf{c}_{k,l}} &= \frac{2\varepsilon}{K^2} \left( \mathbf{T}_{k,l}^c \mathbf{c}_{k,l} - \tilde{\alpha}_2 \theta_l^{\text{comp}} \mathbf{H}_k^H \mathbf{z}_l \right) + \frac{2\mathbf{c}_{k,l}}{f_k^{\text{con}}},\end{aligned}\quad (6.25)$$

和

$$\begin{aligned}\nabla_{\mathbf{a}_{k,j}}^2 &= 2\varepsilon \mathbf{T}_{k,l}^a + \frac{4}{(f_k^{\text{con}})^2} \mathbf{a}_{k,j} \mathbf{a}_{k,j}^T + \frac{2}{f_k^{\text{con}}} \mathbf{I}, \\ \nabla_{\mathbf{b}_{k,i}}^2 &= 2\varepsilon \mathbf{T}_{k,i}^b + \frac{4}{(f_k^{\text{con}})^2} \mathbf{b}_{k,i} \mathbf{b}_{k,i}^T + \frac{2}{f_k^{\text{con}}} \mathbf{I}, \\ \nabla_{\mathbf{c}_{k,l}}^2 &= \frac{2\varepsilon}{K^2} \mathbf{T}_{k,l}^c + \frac{4}{(f_k^{\text{con}})^2} \mathbf{c}_{k,l} \mathbf{c}_{k,l}^T + \frac{2}{f_k^{\text{con}}} \mathbf{I},\end{aligned}\quad (6.26)$$

其中,

$$\mathbf{T}_{k,j}^a = \tilde{\alpha}_1 \sum_{i=1}^I \theta_i^{\text{sens}} \mathbf{H}_k^H \mathbf{v}_i \mathbf{v}_i^H \mathbf{H}_k + \frac{\tilde{\alpha}_2}{K^2} \sum_{l=1}^L \theta_l^{\text{comp}} \mathbf{H}_k^H \mathbf{z}_l \mathbf{z}_l^H \mathbf{H}_k + \tilde{\alpha}_3 \theta_{k,j}^{\text{comm}} \omega_{k,j} \mathbf{H}_k^H \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H \mathbf{H}_k, \quad (6.27)$$

$$\begin{aligned}\mathbf{T}_{k,i}^b &= \tilde{\alpha}_1 \theta_i^{\text{sens}} \left( R_i^2 \mathbf{G}_{k,i}^H \mathbf{v}_i \mathbf{v}_i^H \mathbf{G}_{k,i} + \sum_{o=1}^O \bar{R}_o^2 \mathbf{F}_{k,o}^H \mathbf{v}_i \mathbf{v}_i^H \mathbf{F}_{k,o} \right) \\ &+ \frac{\tilde{\alpha}_2}{K^2} \sum_{l=1}^L \theta_l^{\text{comp}} \left( R_i^2 \mathbf{G}_{k,i}^H \mathbf{z}_l \mathbf{z}_l^H \mathbf{G}_{k,i} + \sum_{o=1}^O \bar{R}_o^2 \mathbf{F}_{k,o}^H \mathbf{z}_l \mathbf{z}_l^H \mathbf{F}_{k,o} \right) \\ &+ \tilde{\alpha}_3 \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} \omega_{k,j} \left( R_i^2 \mathbf{H}_k^H \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H \mathbf{H}_k + \sum_{o=1}^O \bar{R}_o^2 \mathbf{F}_{k,o}^H \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H \mathbf{F}_{k,o} \right),\end{aligned}\quad (6.28)$$

和

$$\begin{aligned}\mathbf{T}_{k,l}^c &= \tilde{\alpha}_1 K^2 \sum_{i=1}^I \theta_i^{\text{sens}} \mathbf{H}_k^H \mathbf{v}_i \mathbf{v}_i^H \mathbf{H}_k + \tilde{\alpha}_2 \theta_l^{\text{comp}} \mathbf{H}_k^H \mathbf{z}_l \mathbf{z}_l^H \mathbf{H}_k \\ &+ \tilde{\alpha}_3 K^2 \sum_{k=1}^K \sum_{j=1}^J \theta_{k,j}^{\text{comm}} \omega_{k,j} \mathbf{H}_k^H \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H \mathbf{H}_k.\end{aligned}\quad (6.29)$$

**算法 6**: 基于加权总性能最大化的联合发射和接收波束设计

输入:  $N, M, K, O, I, L, J, \sigma_n^2, \theta_{k,j}^{\text{comm}}, \theta_{k,i}^{\text{sens}}, \theta_{k,l}^{\text{comp}}, P_{\max,k}, \alpha_p, \forall k, i, l, j, p$

输出:  $\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{v}_i, \mathbf{u}_{k,j}, \mathbf{z}_l$

- 1: 初始化 发射波束  $\mathbf{a}_{k,j}^{(0)} = \mathbf{b}_{k,i}^{(0)} = \mathbf{c}_{k,l}^{(0)} = [\sqrt{\frac{P_{\max,k}}{3M}}, 0, \dots, 0]^T, \forall k, i, l, j$ , 迭代索引  $t = 1$ ;
- 2: **repeat**
- 3: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.21) 计算  $\mathbf{u}_{k,j}^{(t)}$ ;
- 4: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.22) 计算  $\mathbf{v}_i^{(t)}$ ;
- 5: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.23) 计算  $\mathbf{z}_l^{(t)}$ ;
- 6: 基于  $\mathbf{u}_{k,j}^{(t)}, \mathbf{v}_i^{(t)}, \mathbf{z}_l^{(t)}, \mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.15) 计算  $\omega_{k,j}^{(t)}$ ;
- 7: 初始化 迭代索引  $m = 1$ , 障碍参数  $\varepsilon = 1$ , 放大因子  $\nu = 10$ ;
- 8: **repeat**
- 9: 选择合适的步长  $\eta_{k,j}^{a(m)}, \eta_{k,i}^{b(m)}$  和  $\eta_{k,l}^{c(m)}$ ;
- 10: 基于  $\eta_{k,j}^{a(m)}, \mathbf{a}_{k,j}^{(m-1)}, \mathbf{u}_{k,j}^{(t)}, \mathbf{v}_i^{(t)}, \mathbf{z}_l^{(t)}$  和  $\omega_{k,j}^{(t)}$ , 根据式 (6.30) 计算  $\mathbf{a}_{k,j}^{(m)}$ ;
- 11: 基于  $\eta_{k,i}^{b(m)}, \mathbf{b}_{k,i}^{(m-1)}, \mathbf{u}_{k,j}^{(t)}, \mathbf{v}_i^{(t)}, \mathbf{z}_l^{(t)}$  和  $\omega_{k,j}^{(t)}$ , 根据式 (6.31) 计算  $\mathbf{b}_{k,i}^{(m)}$ ;
- 12: 基于  $\eta_{k,l}^{c(m)}, \mathbf{c}_{k,l}^{(m-1)}, \mathbf{u}_{k,j}^{(t)}, \mathbf{v}_i^{(t)}, \mathbf{z}_l^{(t)}$  和  $\omega_{k,j}^{(t)}$ , 根据式 (6.32) 计算  $\mathbf{c}_{k,l}^{(m)}$ ;
- 13: **if** 满足中心性条件 **then**
- 14: 更新  $\varepsilon = \varepsilon * \nu$ ;
- 15: **end if**
- 16: 更新  $m = m + 1$ ;
- 17: **until** 对偶间隙收敛
- 18: 更新  $\mathbf{a}_{k,j}^{(t)} = \mathbf{a}_{k,j}^{(m)}, \mathbf{b}_{k,i}^{(t)} = \mathbf{b}_{k,i}^{(m)}$  和  $\mathbf{c}_{k,l}^{(t)} = \mathbf{c}_{k,l}^{(m)}$ ;
- 19: 更新  $t = t + 1$ ;
- 20: **until** 目标值收敛

上式中的  $\tilde{\alpha}_1 = \frac{\alpha_1}{|\mathcal{F}_1^*|}$ ,  $\tilde{\alpha}_2 = \frac{\alpha_2}{|\mathcal{F}_2^*|}$ ,  $\tilde{\alpha}_3 = \frac{\alpha_3}{|\mathcal{F}_3^*|}$ 。有了梯度和海森矩阵, 发射波束可以根据以下更新规则计算得到:

$$\mathbf{a}_{k,j}^{(t+1)} = \mathbf{a}_{k,j}^{(t)} - \eta_{k,j}^a \left( \nabla_{\mathbf{a}_{k,j}}^2 \right)^{-1} \nabla_{\mathbf{a}_{k,j}}, \quad (6.30)$$

$$\mathbf{b}_{k,i}^{(t+1)} = \mathbf{b}_{k,i}^{(t)} - \eta_{k,i}^b \left( \nabla_{\mathbf{b}_{k,i}}^2 \right)^{-1} \nabla_{\mathbf{b}_{k,i}}, \quad (6.31)$$

$$\mathbf{c}_{k,l}^{(t+1)} = \mathbf{c}_{k,l}^{(t)} - \eta_{k,l}^c \left( \nabla_{\mathbf{c}_{k,l}}^2 \right)^{-1} \nabla_{\mathbf{c}_{k,l}}, \quad (6.32)$$

其中,  $\eta_{k,j}^a$ ,  $\eta_{k,i}^b$  和  $\eta_{k,l}^c$  分别为发射波束  $\mathbf{a}_{k,j}$ ,  $\mathbf{b}_{k,i}$  和  $\mathbf{c}_{k,l}$  的更新步长。综上, 基于加权总性能最大化的联合发射和接收波束设计过程总结在算法 6 中。

### 6.3.2 发射总功率最小化设计

以发射总功率最小化为目标, 同时分别满足通信、感知和计算的 QoS 条件的 MOOP 设计可以表述为:

## M-2: 发射总功率最小化

$$\begin{aligned}
 & \underset{\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{v}_i, \mathbf{u}_{k,j}, \mathbf{z}_l}{\text{minimize}} && \sum_{k=1}^K \left( \sum_{j=1}^J \|\mathbf{a}_{k,j}\|^2 + \sum_{i=1}^I \|\mathbf{b}_{k,i}\|^2 + \sum_{l=1}^L \|\mathbf{c}_{k,l}\|^2 \right) \\
 & \text{s.t.} && \text{C2: } \Gamma_{k,j} \geq \gamma_{k,j}, \forall k \in \Omega_K, j \in \Omega_J, \\
 & && \text{C3: } \text{MSE}_i^{\text{sens}} \leq \delta_i, \forall i \in \Omega_I, \\
 & && \text{C4: } \text{MSE}_l^{\text{comp}} \leq \chi_l, \forall l \in \Omega_L,
 \end{aligned} \tag{6.33}$$

其中  $\gamma_{k,j} > 0$  是关于通信信号  $s_{k,j}^{\text{comm}}$  的最小 SINR 要求,  $\delta_i > 0$  是关于第  $i$  个目标的最大可容忍感知误差,  $\chi_l > 0$  是关于第  $l$  个模型参数的最大可容忍计算误差。问题 M-2 的目标函数是最小化发射总功率, 约束条件 C2, C3 和 C4 分别是通信、感知和计算的 QoS 要求。然而, 由于优化变量, 即接收波束和发射波束, 在约束 C2、C3 和 C4 中相互耦合, 问题 M-2 是非凸的。因此, 同样需要将问题 M-2 解耦成两个子问题, 然后在迭代中依次求解, 直至收敛。基于平衡系统性能和计算复杂度的考虑, 对于优化接收波束的子问题, 仍然使用 MMSE 接收波束, 其定义在式 (6.21)、(6.22) 和 (6.23) 中。对于另一个子问题, 即固定接收波束优化发射波束, 需要处理非凸约束条件 C2, 它等价于

$$\frac{1}{\gamma_{k,j}} |\mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{a}_{k,j}|^2 \geq \sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J |\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{a}_{i,n}|^2 + X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2. \tag{6.34}$$

为了解决非凸性, 需要引入一个辅助变量  $\mathbf{A}_{k,j} = \mathbf{a}_{k,j} \mathbf{a}_{k,j}^H$ , 然后将其代入式 (6.34) 中, 得到

$$\frac{1}{\gamma_{k,j}} \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{A}_{k,j} \mathbf{H}_k^H \mathbf{u}_{k,j}) \geq \sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{A}_{i,n} \mathbf{H}_i^H \mathbf{u}_{k,j}) + X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2. \tag{6.35}$$

基于此, 发射波束优化子问题可以转换成一个标准的半正定规划 (Semidefinite Programming, SDP) 问题:

## M-2': 优化发射波束等价子问题

$$\begin{aligned}
 & \underset{\mathbf{A}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}}{\text{min}} && \sum_{k=1}^K \left( \sum_{j=1}^J \text{tr}(\mathbf{A}_{k,j}) + \sum_{i=1}^I \|\mathbf{b}_{k,i}\|^2 + \sum_{l=1}^L \|\mathbf{c}_{k,l}\|^2 \right) \\
 & \text{s.t.} && \text{C5: } (6.35), \\
 & && \text{C6: } \overline{\text{MSE}}_i^{\text{sens}} \leq \delta_i, \forall i \in \Omega_I, \\
 & && \text{C7: } \overline{\text{MSE}}_l^{\text{comp}} \leq \chi_l, \forall l \in \Omega_L,
 \end{aligned}$$

$$\begin{aligned} \text{C8: } & \mathbf{A}_{k,j} \succeq \mathbf{0}, \forall k \in \Omega_K, j \in \Omega_J, \\ \text{C9: } & \text{Rank}(\mathbf{A}_{k,j}) = 1, \forall k \in \Omega_K, j \in \Omega_J, \end{aligned} \quad (6.36)$$

$$\begin{aligned} \text{其中 } \overline{\text{MSE}}_i^{\text{sens}} &= \text{MSE}_i^{\text{sens}} - \sum_{k=1}^K \sum_{j=1}^J |\mathbf{v}_i^H \mathbf{H}_k \mathbf{a}_{k,j}|^2 + \sum_{k=1}^K \sum_{j=1}^J \text{tr}(\mathbf{v}_i^H \mathbf{H}_k \mathbf{A}_{k,j} \mathbf{H}_k^H \mathbf{v}_i), \quad \overline{\text{MSE}}_l^{\text{comp}} = \\ & \text{MSE}_l^{\text{comp}} - \frac{1}{K^2} \sum_{k=1}^K \sum_{j=1}^J |\mathbf{z}_l^H \mathbf{H}_k \mathbf{a}_{k,j}|^2 + \frac{1}{K^2} \sum_{k=1}^K \sum_{j=1}^J \text{tr}(\mathbf{z}_l^H \mathbf{H}_k \mathbf{A}_{k,j} \mathbf{H}_k^H \mathbf{z}_l). \end{aligned}$$

注意到矩阵  $\mathbf{A}_{k,j}$  的秩约束条件 C9 是非凸的，这使得问题 M-2' 仍然是非凸的。因此，可以采用半正定松弛（Semidefinite Relaxation, SDR）技术，即舍弃约束 C9。这样，问题 M-2' 可以重新表述为

M-2'': 优化发射波束转换子问题

$$\begin{aligned} \min_{\mathbf{A}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}} & \sum_{k=1}^K \left( \sum_{j=1}^J \text{tr}(\mathbf{A}_{k,j}) + \sum_{i=1}^I \|\mathbf{b}_{k,i}\|^2 + \sum_{l=1}^L \|\mathbf{c}_{k,l}\|^2 \right) \\ \text{s.t. } & \text{C5} - \text{C8}. \end{aligned} \quad (6.37)$$

可以发现，问题 M-2'' 是发射波束  $\{\mathbf{A}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}\}$  的联合凸问题，因此可以获得其最优解。值得一提的是，对于问题 M-2'' 的最优解  $\mathbf{A}_{k,j}^*$  的秩，有以下定理：

**定理6.2.** 问题 M-2'' 的最优解  $\mathbf{A}_{k,j}^*$  的秩始终满足  $\text{Rank}(\mathbf{A}_{k,j}^*) = 1, \forall k, j$ 。

证明：请参阅附录 A.4.2。 □

根据定理 6.2，通过对  $\mathbf{A}_{k,j}^*$  进行 EVD，可以得到问题 M-2 的唯一最优解  $\mathbf{a}_{k,j}^*$ ，即

$$\mathbf{a}_{k,j}^* = \sqrt{\lambda_{k,j}^{\max}(\mathbf{A}_{k,j}^*)} \boldsymbol{\xi}_{k,j}^{\max}, \quad (6.38)$$

其中  $\lambda_{k,j}^{\max}(\mathbf{A}_{k,j}^*)$  为  $\mathbf{A}_{k,j}^*$  的最大特征值， $\boldsymbol{\xi}_{k,j}^{\max}$  代表相应的单位特征向量。综上，基于发射总功率最小化的联合发射和接收波束设计过程总结在算法 7 中。

### 6.3.3 算法分析

本小节将简要分析所提的算法 6 和算法 7 的收敛性和复杂度。

**收敛性分析：**对于算法 6，由于问题 M-1 分别对每个变量而言是凸的，可以找到每个问题的最优解。换句话说，每次迭代的解在下次迭代中都是可行的。作为结果，问题 M-1 的目标值在迭代中是单调非增的。此外，由于终端的发射功率限制，问题 M-1 的目

**算法 7**: 基于发射总功率最小化的联合发射和接收波束设计

 输入:  $K, N, M, I, O, L, J, \sigma_n^2, P_0, \delta_i, \chi_l, \gamma_{k,j}, \forall k, i, l, j$ 

 输出:  $\mathbf{a}_{k,j}, \mathbf{b}_{k,i}, \mathbf{c}_{k,l}, \mathbf{u}_{k,j}, \mathbf{v}_i, \mathbf{z}_l, \forall k, i, l, j$ 

- 1: 初始化 发射波束  $\mathbf{a}_{k,i}^{(0)} = \mathbf{b}_{k,l}^{(0)} = \mathbf{c}_{k,j}^{(0)} = [\sqrt{\frac{P_0}{3M}}, 0, \dots, 0]^T, \forall k, i, l, j$ , 迭代索引  $t = 1$ ;
- 2: **repeat**
- 3: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.21) 计算  $\mathbf{u}_{k,j}^{(t)}$
- 4: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.22) 计算  $\mathbf{v}_i^{(t)}$ ;
- 5: 基于  $\mathbf{a}_{k,j}^{(t-1)}, \mathbf{b}_{k,i}^{(t-1)}$  和  $\mathbf{c}_{k,l}^{(t-1)}$ , 根据式 (6.23) 计算  $\mathbf{z}_l^{(t)}$ ;
- 6: 基于  $\mathbf{u}_{k,j}^{(t)}, \mathbf{v}_i^{(t)}$  和  $\mathbf{z}_l^{(t)}$ , 通过解决问题  $M-2''$  获得  $\mathbf{A}_{k,j}^{(t)}, \mathbf{b}_{k,i}^{(t)}$  和  $\mathbf{c}_{k,l}^{(t)}$ ;
- 7: 根据式 (6.38), 通过对  $\mathbf{A}_{k,j}^{(t)}$  进行 EVD 获得  $\mathbf{a}_{k,j}^{(t)}$ ;
- 8:  $t = t + 1$ ;
- 9: **until** 目标值收敛

标值是有下界的。因此, 基于单调有界收敛定理<sup>[139]</sup>, 算法 6 会在适当次数的迭代后收敛。对于算法 7, 采取的 MMSE 接收波束是能够有效保证 QoS 要求的, 同时有助于降低发射总功率。对于给定的 MMSE 接收波束, 问题  $M-2''$  对于发射波束是凸的, 因此可以得到其最优解, 保证在第  $t + 1$  轮迭代中的目标值小于第  $t$  轮迭代中得到的目标值。这意味着发射总功率在迭代中是单调非增的。再者, 由于约束条件 C2、C3 和 C4 中的 QoS 要求, 发射总功率是有下界的。因此, 算法 7 的收敛性能够得到保证。

**复杂度分析:** 由于算法 6 和算法 7 都具有迭代性质, 且每次迭代过程中的执行步骤都是相同的。因此, 只需要分析每次迭代中的执行复杂度。注意到算法 6 的计算复杂度主要来自于步骤 7 至步骤 17, 即寻找问题  $M-1$  的最优发射波束  $\{\mathbf{a}_{k,j}^{(t)}, \mathbf{b}_{k,i}^{(t)}, \mathbf{c}_{k,l}^{(t)}\}$ , 而算法 7 的计算复杂度主要来自于步骤 6, 即寻找问题  $M-2''$  的最优发射波束  $\{\mathbf{A}_{k,j}^{(t)}, \mathbf{b}_{k,i}^{(t)}, \mathbf{c}_{k,l}^{(t)}\}$ 。此外, 由于问题  $M-1$  和问题  $M-2''$  只含有 SOC 或者 LMI 约束, 可以采用最差情况下使用内点法求解这两个问题的计算成本来刻画算法 6 和算法 7 的计算复杂度<sup>[111]</sup>。具体地, 问题  $M-1$  有  $K$  个维度为 1 的 SOC 约束, 决策变量为  $n_1 = \mathcal{O}(K^2 M^2)$ 。问题  $M-2''$  有  $KJ$  个维度为 1 的 LMI 约束,  $KJ$  个维度为  $M$  的 LMI 约束和  $I + L$  个维度为 1 的 SOC 约束, 决策变量为  $n_2 = \mathcal{O}(KM^3)$ 。因此, 对于给定精度  $\epsilon_i > 0, i = 1, 2$ , 每次迭代中算法 6 和算法 7 的计算复杂度可以分别表示为  $\ln(1/\epsilon_i) \cdot \varsigma_i, i = 1, 2$ , 其中  $\varsigma_1 = \sqrt{2K} (n_1 K + n_1^3)$ ,  $\varsigma_2 = \sqrt{KJ(M+1) + 2(I+L)} \cdot n_2 \cdot [KJ(M^3 + n_2 M^2 + n_2 + 1) + I + L + n_2^2]$ 。

## 6.4 仿真结果

本节提供了一些数值仿真结果来验证所提两个算法的可行性和有效性。在不失一般

表 6.2 算法 6 和算法 7 的仿真参数设置

| 参数        | 取值   |
|-----------|--|
| 基站天线数     | $N = 64$   |
| 终端数       | $K = 20$   |
| 终端天线数     | $M = 3$  |
| 数据流数      | $O = 1, I = 1, L = 1, J = 1$   |
| 小区半径      | 500 m  |
| 反射系数的 RMS | $R_i = R_o = 1$  |
| 权重系数      | $\theta_{k,j}^{\text{comm}} = \theta_i^{\text{sens}} = \theta_l^{\text{comp}} = 1$ |
| 最大发射信噪比   | SNR = 5 dB   |
| QoS 要求    | $\gamma_0 = 0.1 \text{ dB}, \delta_0 = 0.01, \chi_0 = 0.01$                        |
| 性能偏好系数    | $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$   |

性的情况下，假设所有终端均随机分布在小区范围内。路径损耗采用 3GPP 标准中的模型，即  $\text{PL}_{\text{dB}} = 128.1 + 37.6 \log_{10}(d)^{[112]}$ ，其中  $d$  (km) 表示发射机和接收机之间的距离。为了便于分析，假设所有终端有相同的 QoS 要求和相同的最大发射功率预算，即  $\gamma_{k,j} = \gamma_0$ ， $\delta_i = \delta_0$ ， $\chi_l = \chi_0$  和  $P_{\max,k} = P_0$ 。此外，用  $\text{SNR} = 10 \log_{10}(P_0/\sigma_n^2)$  表示终端的发射信噪比（以 dB 为单位）。除非另有说明，默认的仿真参数根据表 6.2 设置。

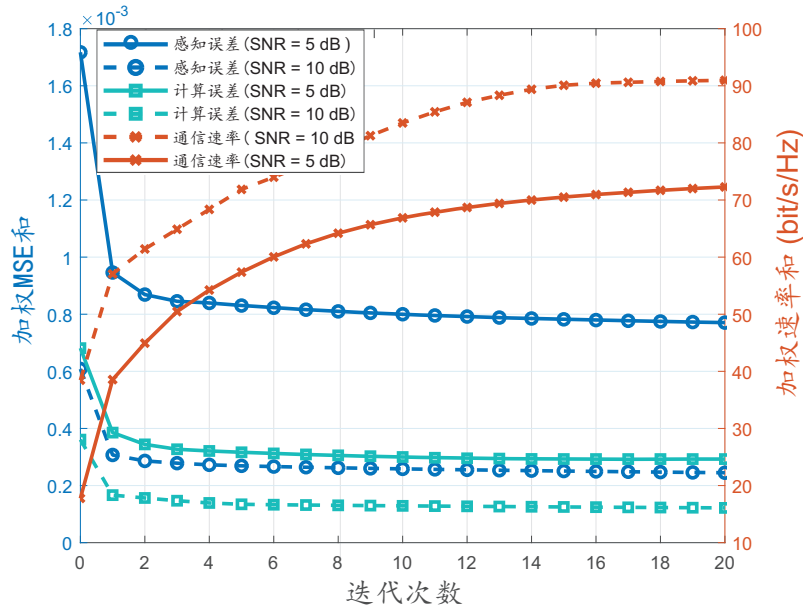


图 6.4 算法 6 的收敛行为

图 6.4 给出了算法 6 在不同发射信噪比下的收敛行为。可以看出，在迭代过程中，感知误差和计算误差单调减小，而通信速率单调增加，它们在几次迭代后都可以收敛到稳定

的平衡点。因此，算法 6 的实现成本对于实际的边缘智能网络来说是可以承受的。

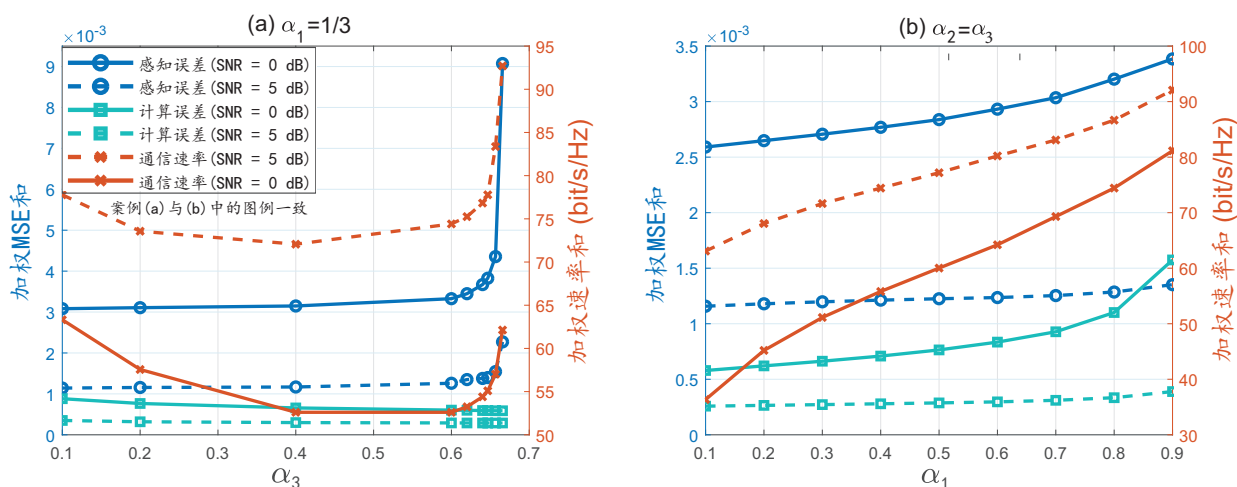


图 6.5 性能偏好系数和终端发射功率对系统性能的影响

图 6.5 研究了在不同发射信噪比下通信  $\alpha_1$ 、感知  $\alpha_2$  和计算  $\alpha_3$  优先级对系统性能的影响。在这里，给出了两种不同的案例。对于案例 (a)，固定通信加权系数  $\alpha_1 = 1/3$ 。在这种情况下， $\alpha_2 + \alpha_3 = 2/3$ ，然后改变计算  $\alpha_3$  的优先级值，以获得不同的解。在案例 (b) 中，固定  $\alpha_2 = \alpha_3$ ，同时改变通信  $\alpha_1$  的优先级以观察不同的结果。可以看出，在案例 (a) 中，随着  $\alpha_3$  增加时，三个功能的性能指标有不同的变化趋势。具体而言，计算误差缓慢减小，感知误差逐渐增大，而通信速率先减小后增大。在案例 (b) 中，随着  $\alpha_1$  的增加，感知和计算的性能逐渐降低，而通信的性能逐渐提高。因此，选择一组适当的优先级系数来平衡通信、感知和计算的性能是有意义的。此外，发现在信噪比为 5 dB 时和 0 dB 时，性能存在较大的差距，这意味着适当提高发射功率可以为系统的整体性能带来更大的增益。

图 6.6 给出了不同基站天线数  $N$  和不同终端数  $K$  对算法 6 性能的影响。在发射信噪比给定的情况下，基站处天线数越多，整体性能越好。这是因为利用增加的天线提供的空间多路复用增益可以有效地增强系统性能。此外，随着终端数量的增加，感知误差和计算误差减小，而通信速率增加。这意味着增加终端数量有利于达到更高的信息传输速率，更准确地估计目标物体的反射系数以及更精确地对 AirFL 进行模型聚合，这说明算法 6 对于有大量设备接入的边缘智能网络具有很大的吸引力。

图 6.7 展示了算法 6 与三种基线波束成形算法的性能对比。它们分别是一种固定形式的 MMSE 波束成形 (Fixed-MMSE Beamforming, F-MMSEBF) 算法，其中固定的 MMSE 接收波束只与信道相关、一种基于 AO 的匹配滤波波束成形 (Matched Filter Beamforming, MFBF) 算法，其接收机采用匹配滤波法和一种基于 AO 的迫零波束成形 (Zero Forcing Beamforming, ZFBF) 算法，其发射机采用迫零法。可以看出，F-MMSEBF 算法的性能



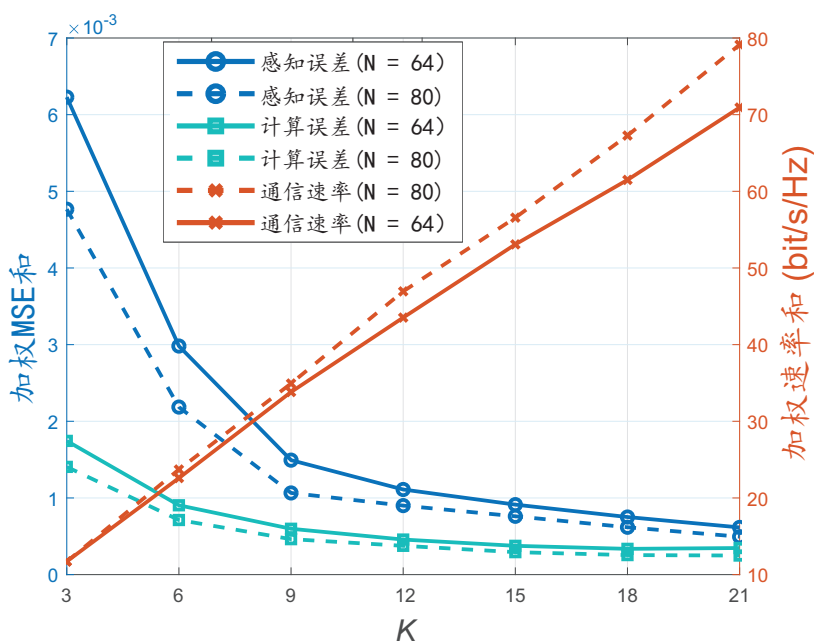


图 6.6 终端数量和基站天线数对系统性能的影响

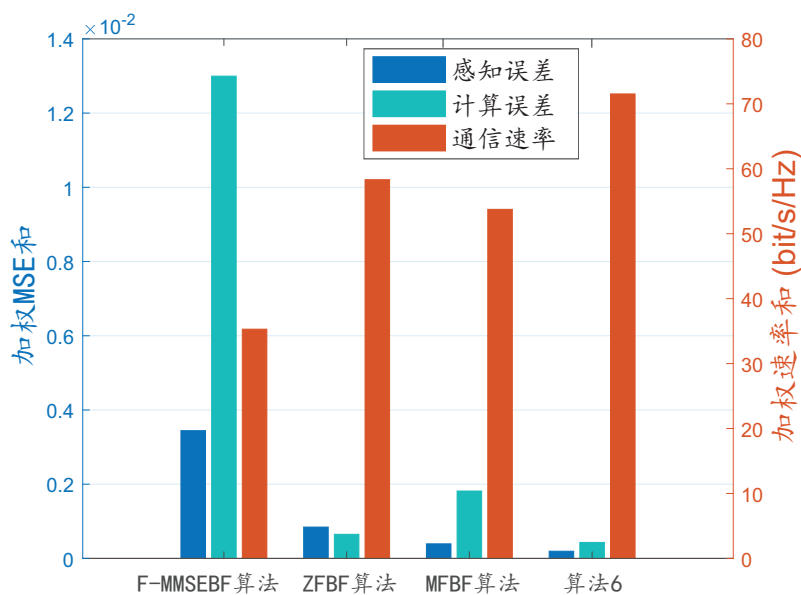


图 6.7 算法 6 与基线算法的性能比较

在所有算法中是最差的，由于其接收波束固定，没有考虑信道的特性，因此严重限制了系统性能。而 MFBF 算法在感知性能上优于 ZFBF 算法，但在计算和通信方面表现较差。注意到，虽然对比算法的复杂度均小于所提算法，但是它们与所提算法的性能差距也较大。可以看出，算法 6 在通信、感知和计算三方面的性能都是最好的，因为它根据信道特性，分别以最优的方式获得发射波束和接收波束。

图 6.8 检验了不同的通信 QoS 条件下算法 7 的收敛性能。可以观察到发射总功率在迭

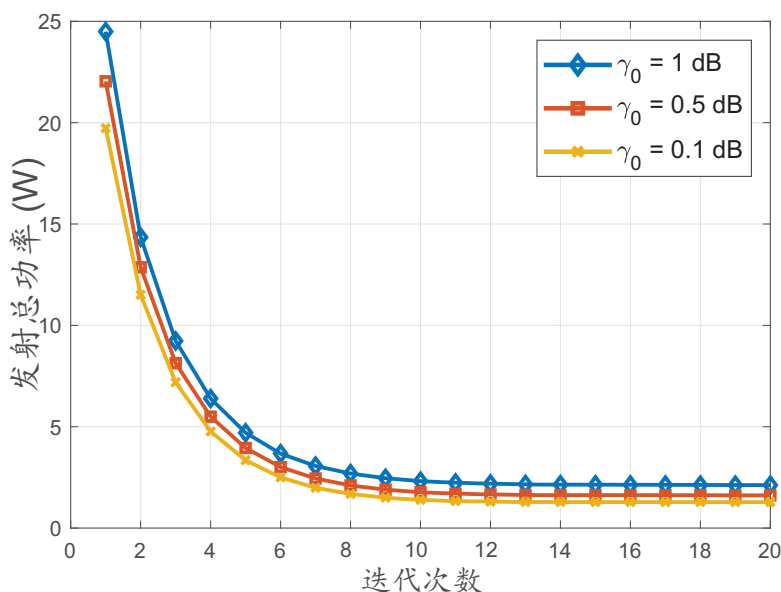


图 6.8 算法 7 的收敛行为

代过程中逐渐降低，且在不同的通信 QoS 条件下均不超过 10 次迭代就可以达到收敛点，这验证了算法 7 的可行性。

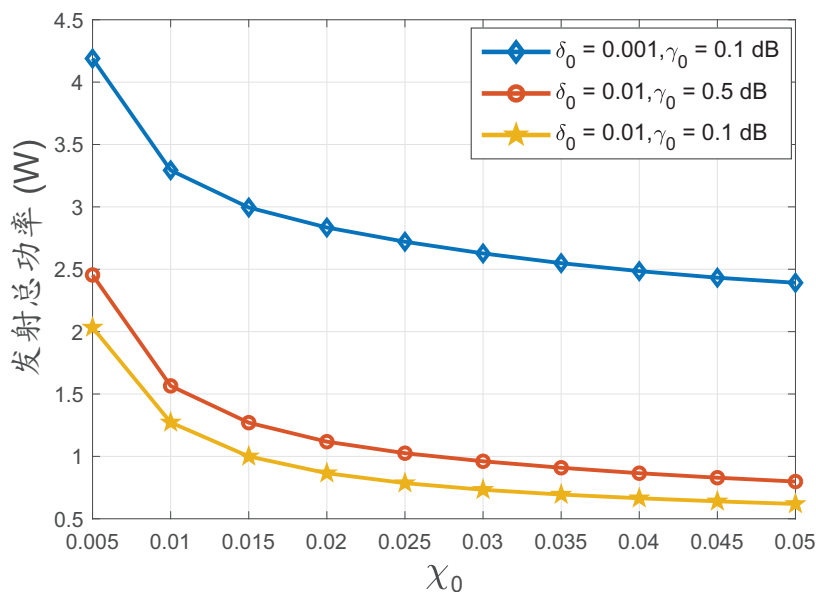


图 6.9 QoS 条件对发射总功率的影响

图 6.9 研究了最大可容忍感知误差  $\delta_0$ 、最大可容忍计算误差  $\chi_0$  和最小通信 SINR 要求  $\gamma_0$  对发射总功率的影响。可以看出，发射总功率随着最大可容忍计算误差的增大而减小，因为  $\chi_0$  越大，代表对计算精度的约束越宽松，因此发射功耗越低。此外，在  $\chi_0$  的整个区域内， $\delta_0 = 0.001$  和  $\gamma_0 = 0.1$  dB 时发射总功率最小，而  $\delta_0 = 0.01$  和  $\gamma_0 = 0.1$  dB 时发射总功率最大。这意味着较大的  $\delta_0$  对感知精度的限制较小，较大的  $\gamma_0$  表示对通信的要求较高。

因此，适当放宽性能要求以降低发射总功率是有意义的。

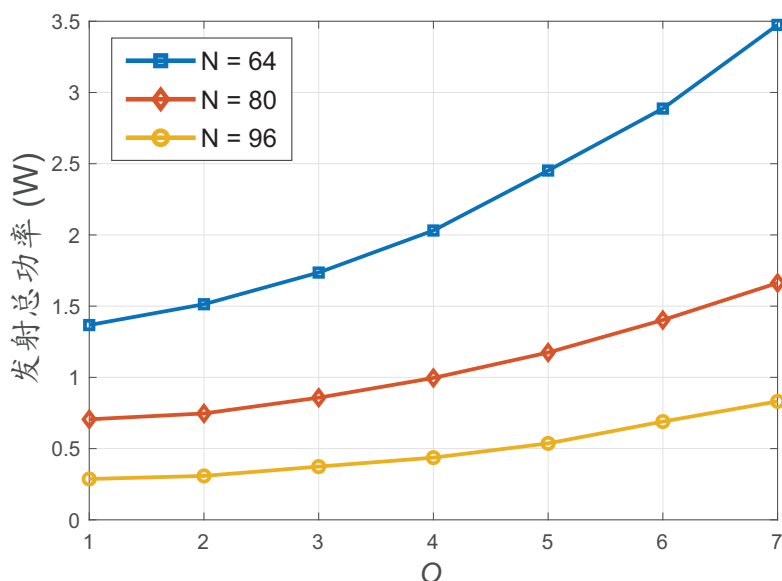


图 6.10 杂波数量和基站天线数对发射总功率的影响

接着，图 6.10 研究了算法 7 中杂波数  $O$  和基站天线数  $N$  对系统性能的影响。可以发现，随着杂波数量的增加，总发射功率显著增加。这是因为杂波作为干扰成分对通信、感知和计算的性能都是有害的，因此需要消耗更多的功率来满足 QoS 要求。此外，得益于更多的阵列增益，基站具有更多天线时总发射功率更低。再者， $N = 80$  和  $N = 96$  之间的性能增益差距小于  $N = 64$  和  $N = 80$  之间的性能增益差距，这说明通过增加基站天线数量所提供的性能增强并不是无限的。请注意，在实际系统中，配备更多天线的基站可以有效地降低发射总功耗，但这也导致了更高的射频链路成本。因此，部署适当数量的基站天线可以在性能和成本之间取得平衡。

最后，图 6.11 研究了不同算法在最小化发射总功率方面的性能比较。与之前相同，三种基线算法分别为 F-MMSEBF 算法、MFBF 算法和 ZFBF 算法。对于所有算法，发射总功率均随着通信 SINR 要求  $\gamma_0$  的增大而增加。这是因为对于给定的感知和计算性能要求，更大的  $\gamma_0$  意味着对通信质量有着更严格的要求，从而导致更大的传输功耗。可以看出，F-MMSEBF 算法在所有算法中表现最差，虽然 ZFBF 算法在  $\gamma_0$  的低值区域的性能优于 MFBF 算法，但其在  $\gamma_0$  的高值区域表现较差。此外，本文提出的算法 7 在  $\gamma_0$  的整个区域范围内均表现最佳，特别是在通信要求较高的情况下，仍然能以较低的发射功率满足系统性能，这验证了算法 7 在边缘智能网络中的优势。

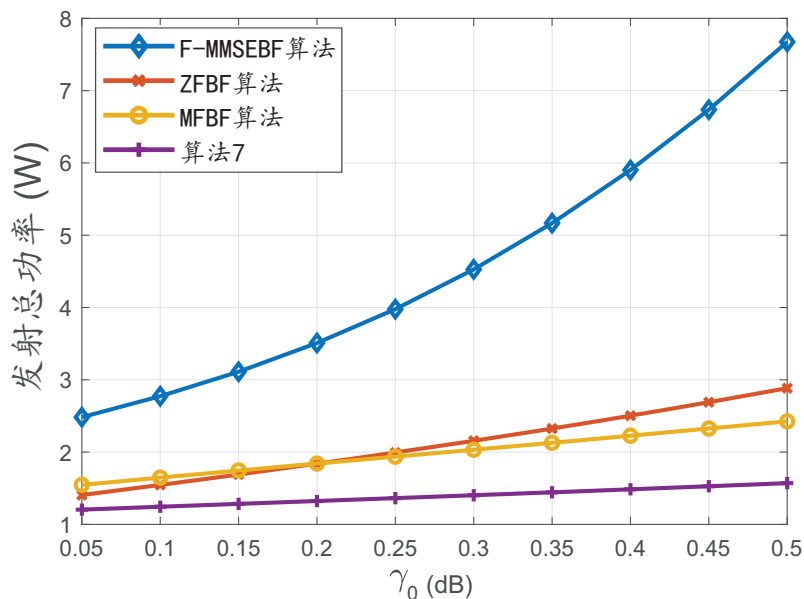


图 6.11 算法 7 与基线算法的性能比较

## 6.5 本章小结

本章将通信、感知和计算这三个孤立的功能进行了一般性的整合，为边缘智能网络提供了一种功能复用的通信、感知和计算一体化架构。为了提高系统整体性能，分别从最大化加权总性能和最小化发射总功率这两个典型的角度提出了联合发射和接收波束成形优化算法。根据边缘智能网络的实际应用需求，可以通过改变系统优先级来获得所期望的通信、感知和计算的性能。大量仿真结果验证了两种所提算法的优越性和有效性。

## 7 总结与展望

### 7.1 研究总结

为了满足各种新兴业务与智能应用的多维极致性能需求，边缘智能网络需要在有限的资源下实现对海量信息的高效获取、传输和处理。然而，在过去的系统中，通信、感知和计算功能都是单独设计的，这不仅浪费了稀缺的频谱资源，还制约了网络性能的提升，更无法满足未来数据驱动的智能服务的高性能需求。因此，有必要研究面向边缘智能网络的通信、感知和计算一体化理论与技术，以提升系统资源利用率和总体性能。面向边缘智能网络的通信、感知和计算一体化设计，关键是如何有效利用有限的无线资源和硬件资源实现通信、感知和计算功能之间的相互增强和互利互惠，从而达到基于业务特点和需求的自适应调度和复用。

基于此，本学位论文从理论和技术上深入研究了面向边缘智能网络的通信、感知和计算一体化，并从设计、分析和优化的角度提出了四个一体化典型架构，即面向感知的通信和计算一体化、计算辅助的通信和感知一体化、通信支撑的感知和计算一体化以及功能复用的通信、感知和计算一体化。具体研究内容和主要贡献如下：

1. 为了给面向边缘智能网络的通信、感知和计算一体化设计提供理论指导和技术支持，分析了通信、感知和计算一体化设计的基础理论和关键技术，包括三者功能的具体定义和性能评价指标、一体化的发展趋势和通用系统架构以及实现一体化的关键技术。
2. 为了对边缘智能网络中的海量感知数据进行高效处理，提出了一种面向感知的通信和计算一体化架构，其利用非正交传输和边缘 FL 实现通信和计算一体化，有效减少了数据的传输量并降低了数据处理时延，从而提高了信息感知有效性。通过分析信道衰落、干扰和噪声以及信道估计误差等不利因素对 FL 模型参数传递的影响，提出了一种联合设备选择和收发机设计的鲁棒优化算法来提高通信和计算一体化的整体性能。仿真结果证明了所提的鲁棒算法能够有效提高通信可靠性和学习准确性。

3. 为了解决边缘智能网络中资源共享导致的多路信号混叠问题，提出了一种计算辅助的通信和感知一体化架构，其利用多功能基站在同一频谱和相同硬件上实现了通信和感知的协同，有效提高了频谱和硬件的利用效率。通过分析通信内干扰与通信和感知功能间干扰对系统性能的影响，借助 AI 强大的计算能力，提出了一种基于 DL 的联合感知发射波形和通信接收波束设计算法来提高通信和感知一体化的整体性能。仿真结果证明了所提的 DL 算法具有低复杂性和高鲁棒性的特点。
4. 为了充分发挥边缘智能网络中大量分布的多功能边缘节点和一体化网元的能力，提出了一种通信支撑的感知和计算一体化架构，其利用多节点协作通信实现感知信息和计算数据的高效传递和汇聚。通过分析感知任务和计算任务之间的耦合关系，在边缘节点资源有限的情况下，提出了一种联合卸载选择和资源分配算法来增强感知和计算一体化的整体性能。仿真结果证明了所提算法可以在边缘智能网络中同时实现高精度感知和低时延计算。
5. 为了在边缘智能网络中实现通信、感知和计算三者功能的有效融合与复用，提出了一种功能复用的通信、感知和计算一体化架构，其利用配置大规模天线阵列的基站所具有的空间自由度，在同一资源块上同时实现多目标感知、多维计算和多流通信。通过分析功能复用情况下多业务的同信道干扰和资源竞争对系统性能的影响，提出了两种联合发射和接收波束设计算法，以在有限的无线资源下实现通信、感知和计算三者的性能折中。仿真结果证明了所提的两种算法分别能够有效提升系统总性能和降低系统总能耗。

## 7.2 未来展望

本文基于边缘智能网络的特点和需求，对通信、感知和计算一体化展开了深入的研究，并取得了一些创新性的成果，但还有较大的改进和提升空间。此外，通信、感知和计算一体化还在不断演进中，仍存在诸多挑战和待解决的问题。在本文的基础上，未来的研究方向包括但不限于如下几点：

1. 高移动性：本文提出的通信、感知和计算一体化设计和优化方案均考虑低移动性的场景，因此无法直接应用到具有高移动性的场景中。在高动态环境下，通信、感知和计算一体化的设计和优化会面临更大的挑战和难度。具体而言，在高移动性的场景中信道会产生快速时变，且受多径效应和多普勒效应影响大，导致 CSI 获取困难、

目标难以实时跟踪和数据传输速率受限等问题。因此，未来可以考虑高动态环境，针对其特点，对通信、感知和计算一体化进行设计和优化。

2. 信号同步：本文是在信号同步的前提下实现分布式节点的高效协作通信、感知和计算。例如，本文中采取的关键技术，如空中计算、协同感知和协作计算等，对分布式节点同步的要求非常苛刻。若系统无法满足同步的需求，便会大大降低数据传输的可靠性、感知的精确性以及计算的准确性。特别是，高精度感知要求时钟同步在纳秒级别，而传统通信往往无法达到这一要求。因此，未来可以研究更精确的同步算法、协议和技术，保证分布式系统的可靠性，以满足未来极高性能要求的通信、感知和计算服务。
3. 硬件设计：在本文所考虑的边缘智能网络中，多功能边缘节点和一体化网元同时具有通信、感知和计算的功能。然而，在实际硬件设计时面临着很大的挑战，不能简单地将感知和计算模块引入到传统的通信设备中。例如，对于感知功能而言，往往对器件有着高精度和高分辨率的要求，这会导致高昂的实现成本和极高的处理复杂度。对于计算功能而言，实现 AI 算法往往需要使用专门的 GPU，这难以满足通信设备对低功耗、平价计算的需求。因此，如何设计匹配通信、感知和计算能力和需求的硬件也是未来需要研究的重点方向之一。
4. 信息安全：本文研究的边缘智能网络中通信、感知和计算一体化未涉及到安全性问题。实际上，边缘智能网络中存在着大量的边缘节点，且分布在不同的位置，需要通过无线网络实现数据的交互和功能的协同。部分边缘节点结构简单、功能单一，它们的数据极易受到窃听、攻击或篡改，从而对网络进行渗透和破坏，导致信息安全的短板效应突出。此外，网络边缘产生了海量数据，这需要采取更强大的保护机制才能够保护数据的安全性。因此，下一步可以研究边缘智能网络中通信、感知和计算一体化的信息安全问题。
5. 星地融合：本文研究的面向边缘智能网络的通信、感知和计算一体化聚焦于地面网络。在未来，6G 将构建空天地一体的全球网络，实现全域无缝接入。卫星网络有着覆盖范围广、灵活性高、扩展性强的特点，能够有效弥补地面网络的不足，成为实现全球覆盖的重要手段。目前卫星网络已经在通信、导航、遥感等方面发挥了重要作用，如何有效管理和协调星地资源，提供更大连接数、更高可靠性、更低时延的通信、感知和计算服务是星地融合的关键。因此，面向星地融合网络的通信、感知和计算一体化也是一个很好的研究方向。

6. 新兴技术：为了助力未来 6G 网络实现“万物智联”的美好愿景，通信、感知和计算一体化可以与一些 6G 新兴关键技术有机结合、协同演进，从而发挥出更大的潜力。例如，利用智能超表面技术实现信号增强和覆盖增强，借助超大规模 MIMO 技术获得超高的空间自由度和极高的空间分辨率，使用太赫兹频段实现超高速率传输和超高精度感知，结合算网融合技术实现网络资源的统一编排和协同调度。因此，未来可以进一步探索通信、感知和计算一体化与新兴技术结合面临的挑战和关键问题。



## 参考文献

- [1] IMT-2030 (6G) 推进组. 通信感知一体化技术研究报告[R]. 2021.
- [2] Andrews J G, Buzzi S, Choi W, et al. What will 5G be?[J]. *IEEE Journal on Selected Areas in Communications*, 2014, 32(6): 1065 – 1082.
- [3] Series M. IMT Vision–Framework and overall objectives of the future development of IMT for 2020 and beyond: ITU-R M.2083-0[R]. 2015.
- [4] Palattella M R, Dohler M, Grieco A, et al. Internet of things in the 5G era: Enablers, architecture, and business models[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(3): 510 – 527.
- [5] Saad W, Bennis M, Chen M. A vision of 6G wireless systems: Applications, trends, technologies, and open research problems[J]. *IEEE Network*, 2019, 34(3): 134 – 142.
- [6] Dang S, Amin O, Shihada B, et al. What should 6G be?[J]. *Nature Electronics*, 2020, 3(1): 20 – 29.
- [7] IMT-2030 (6G) 推进组. 6G总体愿景与潜在关键技术白皮书[R]. 2021.
- [8] Statista. Number of IoT connected devices worldwide 2019-2021, with forecasts to 2030[EB]. 2022.
- [9] Series M. IMT traffic estimates for the years 2020 to 2030: ITU-R M.2370-0[R]: ITU, 2015.
- [10] You X, Zhang C, Tan X, et al. AI for 5G: research directions and paradigms[J]. *Science China Information Sciences*, 2019, 62: 1 – 13.
- [11] Li R, Zhao Z, Zhou X, et al. Intelligent 5G: When cellular networks meet artificial intelligence[J]. *IEEE Wireless Communications*, 2017, 24(5): 175 – 183.
- [12] Wang W, Kumar N, Chen J, et al. Realizing the potential of the internet of things for smart tourism with 5G and AI[J]. *IEEE Network*, 2020, 34(6): 295 – 301.
- [13] Lee H, Lee S H, Quek T Q, et al. Deep learning framework for wireless systems: Applications to optical wireless communications[J]. *IEEE Communications Magazine*, 2019, 57(3): 35 – 41.
- [14] Shi W, Cao J, Zhang Q, et al. Edge computing: Vision and challenges[J]. *IEEE internet of things journal*, 2016, 3(5): 637 – 646.
- [15] Xu D, Li T, Li Y, et al. Edge intelligence: Architectures, challenges, and applications[J]. *arXiv preprint arXiv:2003.12172*, 2020.
- [16] Zhou Z, Chen X, Li E, et al. Edge intelligence: Paving the last mile of artificial intelligence with edge

- computing[J]. *Proceedings of the IEEE*, 2019, 107(8): 1738–1762.
- [17] Deng S, Zhao H, Fang W, et al. Edge intelligence: The confluence of edge computing and artificial intelligence[J]. *IEEE Internet of Things Journal*, 2020, 7(8): 7457–7469.
- [18] Liu M, Fang S, Dong H, et al. Review of digital twin about concepts, technologies, and industrial applications[J]. *Journal of Manufacturing Systems*, 2021, 58: 346–361.
- [19] Clemm A, Vega M T, Ravuri H K, et al. Toward truly immersive holographic-type communication: Challenges and solutions[J]. *IEEE Communications Magazine*, 2020, 58(1): 93–99.
- [20] Xie H, Yan Z, Yao Z, et al. Data collection for security measurement in wireless sensor networks: A survey[J]. *IEEE Internet of Things Journal*, 2018, 6(2): 2205–2224.
- [21] Chen X, Zhang Z, Zhong C, et al. Exploiting multiple-antenna techniques for non-orthogonal multiple access[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(10): 2207–2220.
- [22] Dillon T, Wu C, Chang E. Cloud computing: issues and challenges[C] // 2010 24th IEEE International Conference on Advanced Information Networking and Applications. 2010: 27–33.
- [23] Dilley J, Maggs B, Parikh J, et al. Globally distributed content delivery[J]. *IEEE Internet Computing*, 2002, 6(5): 50–58.
- [24] Satyanarayanan M, Bahl P, Caceres R, et al. The case for vm-based cloudlets in mobile computing[J]. *IEEE pervasive Computing*, 2009, 8(4): 14–23.
- [25] Bonomi F, Milito R, Zhu J, et al. Fog computing and its role in the internet of things[C] // Proceedings of the first edition of the MCC workshop on Mobile cloud computing. 2012: 13–16.
- [26] Chang H, Hari A, Mukherjee S, et al. Bringing the cloud to the edge[C] // 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). 2014: 346–351.
- [27] Patel M, Naughton B, Chan C, et al. Mobile-edge computing introductory technical white paper[J]. White paper, mobile-edge computing (MEC) industry initiative, 2014, 29: 854–864.
- [28] Hu Y C, Patel M, Sabella D, et al. Mobile edge computing—A key technology towards 5G[J]. *ETSI white paper*, 2015, 11(11): 1–16.
- [29] Mao Y, You C, Zhang J, et al. A survey on mobile edge computing: The communication perspective[J]. *IEEE communications surveys & tutorials*, 2017, 19(4): 2322–2358.
- [30] Mach P, Becvar Z. Mobile edge computing: A survey on architecture and computation offloading[J]. *IEEE communications surveys & tutorials*, 2017, 19(3): 1628–1656.
- [31] Turing A M. Computing machinery and intelligence[J]. *Mind*, 1950, LIX(236): 433–460.
- [32] Murphy R R. Introduction to AI robotics[M]. USA: MIT Press, 2019.
- [33] Jackson P. Introduction to expert systems[M]. USA: Addison-Wesley Pub. Co. Reading Mass., 1986.
- [34] Forsyth D A, Ponce J. Computer vision: A modern approach[M]. USA: Prentice Hall Professional Tech-

- nical Reference, 2002.
- [35] C M, H S. Foundations of statistical natural language processing[M]. USA : MIT Press, 1999.
- [36] Mitchell T M. Machine learning[M]. USA : McGraw-Hill New York, 2007.
- [37] Caruana R, Niculescu-Mizil A. An empirical comparison of supervised learning algorithms[C] // 2006 23rd international conference on Machine learning. 2006 : 161 – 168.
- [38] Figueiredo M A T, Jain A K. Unsupervised learning of finite mixture models[J]. IEEE Transactions on pattern analysis and machine intelligence, 2002, 24(3) : 381 – 396.
- [39] Sutton R S, Barto A G. Reinforcement learning: An introduction[M]. USA : MIT Press, 2018.
- [40] LeCun Y, Bengio Y, Hinton G. Deep learning[J]. Nature, 2015, 521(7553) : 436 – 444.
- [41] Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks[J]. Pattern recognition, 2018, 77 : 354 – 377.
- [42] Pascanu R, Mikolov T, Bengio Y. On the difficulty of training recurrent neural networks[C] // 2013 International conference on machine learning. 2013 : 1310 – 1318.
- [43] Gardner M W, Dorling S. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences[J]. Atmospheric environment, 1998, 32(14-15) : 2627 – 2636.
- [44] Gartner. Hype Cycle for the Internet of Things[EB]. 2018.
- [45] Yi S, Hao Z, Qin Z, et al. Fog computing: Platform and applications[C] // 2015 3rd IEEE workshop on hot topics in web systems and technologies (HotWeb). 2015 : 73 – 78.
- [46] Ha K, Chen Z, Hu W, et al. Towards wearable cognitive assistance[C] // 2014 12th Annual International Conference on Mobile Systems, Applications, and Services. 2014 : 68 – 81.
- [47] Lane N D, Bhattacharya S, Mathur A, et al. Squeezing deep learning into mobile and embedded devices[J]. IEEE Pervasive Computing, 2017, 16(3) : 82 – 88.
- [48] Hard A, Rao K, Mathews R, et al. Federated learning for mobile keyboard prediction[J]. arXiv preprint arXiv:1811.03604, 2018.
- [49] 世界5G大会. Integration of sensing, communication and computing toward 6G[R]. 2020.
- [50] 闫实, 彭木根, 王文博. 通信-感知-计算融合: 6G 愿景与关键技术[J]. 北京邮电大学学报, 2021, 44(4) : 1 – 10.
- [51] 中国通信学会. 通感算一体化网络前沿报告[EB]. 2022.
- [52] 王文博, 侯跃峰, 田峰. 面向扩展现实的通信感知计算一体化视频传输及 QoE 评价[J]. 移动通信, 2022, 46(05) : 84 – 88.
- [53] 马忠贵, 李卓, 梁彦鹏. 自动驾驶车联网中通感算融合研究综述与展望[J]. 工程科学学报, 2023, 45(01) : 137 – 149.
- [54] 廖佳佳, 向路平, 胡杰, 等. 基于无人机的 6G 通感算研究概述[J]. 移动通信, 2023, 47(03) : 92 – 100.

- [55] 景毅, 姜春晓, 詹亚锋. 面向卫星通信的 6G 通感算融合架构、技术与挑战[J]. 无线电通信技术, 2023, 49(01): 12–20.
- [56] 滕博宇, 袁晓军. 智能超表面辅助 6G 通感算深度融合关键技术[J]. 移动通信, 2023, 47(03): 25–30.
- [57] Nazer B, Gastpar M. Computation over multiple-access channels[J]. IEEE Transactions on information theory, 2007, 53(10): 3498–3516.
- [58] Gastpar M. Uncoded transmission is exactly optimal for a simple Gaussian “sensor” network[J]. IEEE Transactions on Information Theory, 2008, 54(11): 5247–5251.
- [59] Chen L, Zhao N, Chen Y, et al. Over-the-air computation for IoT networks: Computing multiple functions with antenna arrays[J]. IEEE Internet of Things journal, 2018, 5(6): 5296–5306.
- [60] Wen D, Zhu G, Huang K. Reduced-dimension design of MIMO over-the-air computing for data aggregation in clustered IoT networks[J]. IEEE Transactions on Wireless Communications, 2019, 18(11): 5255–5268.
- [61] Saleem U, Liu Y, Jangsher S, et al. Latency minimization for D2D-enabled partial computation offloading in mobile edge computing[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 4472–4486.
- [62] Tran T X, Pompili D. Joint task offloading and resource allocation for multi-server mobile-edge computing networks[J]. IEEE Transactions on Vehicular Technology, 2018, 68(1): 856–868.
- [63] Dai Y, Xu D, Maharjan S, et al. Joint computation offloading and user association in multi-task mobile edge computing[J]. IEEE Transactions on Vehicular Technology, 2018, 67(12): 12313–12325.
- [64] Ye H, Li G Y, Juang B-H. Power of deep learning for channel estimation and signal detection in OFDM systems[J]. IEEE Wireless Communications Letters, 2017, 7(1): 114–117.
- [65] Li J, Gao H, Lv T, et al. Deep reinforcement learning based computation offloading and resource allocation for MEC[C] // 2018 IEEE wireless communications and networking conference (WCNC). 2018: 1–6.
- [66] Xia W, Zheng G, Zhu Y, et al. A deep learning framework for optimization of MISO downlink beamforming[J]. IEEE Transactions on Communications, 2019, 68(3): 1866–1880.
- [67] Paulraj A J, Kailath T. Increasing capacity in wireless broadcast systems using distributed transmission/directional reception (DTDR): U.S. Patent, No. 5,345,599[P]. [1994].
- [68] Rappaport T S, Sun S, Mayzus R, et al. Millimeter wave mobile communications for 5G cellular: It will work![J]. IEEE access, 2013, 1: 335–349.
- [69] Mealey R M. A method for calculating error probabilities in a radar communication system[J]. IEEE Transactions on Space Electronics and Telemetry, 1963, 9(2): 37–42.
- [70] Roberton M, Brown E. Integrated radar and communications based on chirped spread-spectrum techniques[C] // 2003 IEEE MTT-S International Microwave Symposium Digest: Vol 1. 2003: 611–614.
- [71] Han L, Wu K. Radar and radio data fusion platform for future intelligent transportation system[C] // 2010

- 7th European Radar Conference. 2010: 65–68.
- [72] Li B, Kumar H, Petropulu A P. A joint design approach for spectrum sharing between radar and communication systems[C] // 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2016: 3306–3310.
- [73] Deng H, Himed B. Interference mitigation processing for spectrum-sharing between radar and wireless communications systems[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2013, 49(3): 1911–1919.
- [74] Mahal J A, Khawar A, Abdelhadi A, et al. Spectral coexistence of MIMO radar and MIMO cellular system[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, 53(2): 655–668.
- [75] Lu L, Li G Y, Swindlehurst A L, et al. An overview of massive MIMO: Benefits and challenges[J]. *IEEE journal of selected topics in signal processing*, 2014, 8(5): 742–758.
- [76] Jiang F, Chen J, Swindlehurst A L, et al. Massive MIMO for wireless sensing with a coherent multiple access channel[J]. *IEEE Transactions on Signal Processing*, 2015, 63(12): 3005–3017.
- [77] Swindlehurst A L, Ayanoglu E, Heydari P, et al. Millimeter-wave massive MIMO: The next wireless revolution?[J]. *IEEE Communications Magazine*, 2014, 52(9): 56–62.
- [78] Guan J, Madani S, Jog S, et al. Through fog high-resolution imaging using millimeter wave radar[C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020: 11464–11473.
- [79] Molisch A F, Ratnam V V, Han S, et al. Hybrid beamforming for massive MIMO: A survey[J]. *IEEE Communications Magazine*, 2017, 55(9): 134–141.
- [80] Han K, Hong S. High-resolution phased-subarray MIMO radar with grating lobe cancellation technique[J]. *IEEE Transactions on Microwave Theory and Techniques*, 2022, 70(5): 2775–2785.
- [81] Cui Y, Liu F, Jing X, et al. Integrating sensing and communications for ubiquitous IoT: Applications, trends, and challenges[J]. *IEEE Network*, 2021, 35(5): 158–167.
- [82] Liu F, Cui Y, Masouros C, et al. Integrated sensing and communications: Towards dual-functional wireless networks for 6G and beyond[J]. *IEEE Journal on Selected Areas in Communications*, 2022.
- [83] Liu A, Huang Z, Li M, et al. A survey on fundamental limits of integrated sensing and communication[J]. *IEEE Communications Surveys & Tutorials*, 2022, 24(2): 994–1034.
- [84] Zhou T, Tang D, Zhu H, et al. Reinforcement learning with composite rewards for production scheduling in a smart factory[J]. *IEEE Access*, 2020, 9: 752–766.
- [85] Hu L, Miao Y, Wu G, et al. iRobot-Factory: An intelligent robot factory based on cognitive manufacturing and edge computing[J]. *Future Generation Computer Systems*, 2019, 90: 569–577.
- [86] Wang J, Jiang C, Han Z, et al. Internet of vehicles: Sensing-aided transportation information collection and diffusion[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(5): 3813–3825.

- [87] Nie L, Ning Z, Wang X, et al. Data-driven intrusion detection for intelligent internet of vehicles: A deep convolutional neural network-based method[J]. *IEEE Transactions on Network Science and Engineering*, 2020, 7(4): 2219–2230.
- [88] Liu F, Chen J, Zhang Q, et al. Online MEC offloading for V2V networks[J/OL]. *IEEE Transactions on Mobile Computing*(Early Access), 2022: 1–13.  
<http://dx.doi.org/10.1109/TMC.2022.3186893>.
- [89] Ghazal T M, Hasan M K, Alshurideh M T, et al. IoT for smart cities: Machine learning approaches in smart healthcare—A review[J]. *Future Internet*, 2021, 13(8): 218.
- [90] Catarinucci L, De Donno D, Mainetti L, et al. An IoT-aware architecture for smart healthcare systems[J]. *IEEE internet of things journal*, 2015, 2(6): 515–526.
- [91] Abdellatif A A, Mohamed A, Chiasserini C F, et al. Edge computing for smart health: Context-aware approaches, opportunities, and challenges[J]. *IEEE Network*, 2019, 33(3): 196–203.
- [92] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. Wireless sensor networks: a survey[J]. *Computer networks*, 2002, 38(4): 393–422.
- [93] Liyanaarachchi S D, Barneto C B, Riihonen T, et al. Joint OFDM waveform design for communications and sensing convergence[C] // *2020 IEEE International Conference on Communications (ICC)*. 2020: 1–6.
- [94] Wei Z, Yuan W, Li S, et al. Orthogonal time-frequency space modulation: A promising next-generation waveform[J]. *IEEE wireless communications*, 2021, 28(4): 136–144.
- [95] Barrenechea P, Elferink F, Janssen J. FMCW radar with broadband communication capability[C] // *2007 European Radar Conference*. 2007: 130–133.
- [96] Chen X, Wang X, Xu S, et al. A novel radar waveform compatible with communication[C] // *2011 International Conference on Computational Problem-Solving (ICCP)*. 2011: 177–181.
- [97] Palomar DP, Cioffi J M, Lagunas M A. Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization[J]. *IEEE Transactions on Signal Processing*, 2003, 51(9): 2381–2401.
- [98] Zhang Z, Chai X, Long K, et al. Full duplex techniques for 5G networks: self-interference cancellation, protocol design, and relay selection[J]. *IEEE Communications Magazine*, 2015, 53(5): 128–137.
- [99] Liu F, Zhou L, Masouros C, et al. Toward dual-functional radar-communication systems: Optimal waveform design[J]. *IEEE Transactions on Signal Processing*, 2018, 66(16): 4264–4279.
- [100] Qi Q, Chen X, Zhong C, et al. Integrated sensing, computation and communication in B5G cellular Internet of Things[J]. *IEEE Transactions on Wireless Communications*, 2020, 20(1): 332–344.
- [101] Buck R C. Approximate complexity and functional representation: ADA031972[R]: Wisconsin Univ Madison Mathematics Research Center, 1976.

- [102] Zhu G, Liu D, Du Y, et al. Toward an intelligent edge: Wireless communication meets machine learning[J]. *IEEE communications magazine*, 2020, 58(1): 19–25.
- [103] Niknam S, Dhillon H S, Reed J H. Federated learning for wireless communications: Motivation, opportunities, and challenges[J]. *IEEE Communications Magazine*, 2020, 58(6): 46–51.
- [104] Goldenbaum M, Boche H, Stańczak S. Nomographic functions: Efficient computation in clustered Gaussian sensor networks[J]. *IEEE Transactions on Wireless Communications*, 2014, 14(4): 2093–2105.
- [105] Sun H, Zhou F, Hu R Q, et al. Robust beamforming design in a NOMA cognitive radio network relying on SWIPT[J]. *IEEE Journal on Selected Areas in Communications*, 2018, 37(1): 142–155.
- [106] Qi Q, Chen X, Ng D W K. Robust beamforming for NOMA-based cellular massive IoT with SWIPT[J]. *IEEE Transactions on Signal Processing*, 2019, 68: 211–224.
- [107] Boyd S, Boyd S P, Vandenberghe L. *Convex optimization*[M]. UK: Cambridge University Press, 2004.
- [108] Razaviyayn M, Hong M, Luo Z-Q. A unified convergence analysis of block successive minimization methods for nonsmooth optimization[J]. *SIAM Journal on Optimization*, 2013, 23(2): 1126–1153.
- [109] Both J W. On the rate of convergence of alternating minimization for non-smooth non-strongly convex optimization in Banach spaces[J]. *Optimization Letters*, 2022, 16(2): 729–743.
- [110] Grant M B S. *CVX: Matlab software for disciplined convex programming*[EB/OL]. 2014. <http://cvxr.com/cvx>.
- [111] Ben-Tal A, Nemirovski A. *Lectures on modern convex optimization: Analysis, algorithms, and engineering applications*[M]. USA: SIAM, 2001.
- [112] 3GPP. Coordinated multi-point operation for LTE physical layer aspects (Rel 11): No. 36.819, Version 11.2.0[R]. 2013.
- [113] Amiri M M, Gündüz D, Kulkarni S R, et al. Convergence of federated learning over a noisy downlink[J]. *IEEE Transactions on Wireless Communications*, 2021, 21(3): 1422–1437.
- [114] Yang K, Jiang T, Shi Y, et al. Federated learning via over-the-air computation[J]. *IEEE Transactions on Wireless Communications*, 2020, 19(3): 2022–2035.
- [115] McMahan B, Moore E, Ramage D, et al. Communication-efficient learning of deep networks from decentralized data[C] // *2017 Artificial intelligence and statistics*. 2017: 1273–1282.
- [116] Tang B, Li J. Spectrally constrained MIMO radar waveform design based on mutual information[J]. *IEEE Transactions on Signal Processing*, 2019, 67(3): 821–834.
- [117] Ouyang C, Liu Y, Yang H. On the performance of uplink ISAC systems[J]. *IEEE Communications Letters*, 2022, 26(8): 1769–1773.
- [118] Molchanov P, Tyree S, Karras T, et al. Pruning convolutional neural networks for resource efficient inference[J]. *arXiv preprint arXiv:1611.06440*, 2016.

- [119] Tang B, Naghsh M M, Tang J. Relative entropy-based waveform design for MIMO radar detection in the presence of clutter and interference[J]. *IEEE transactions on signal processing*, 2015, 63(14): 3783–3796.
- [120] Qi Q, Chen X, Lei L, et al. Outage-constrained robust design for sustainable B5G cellular internet of things[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(12): 5780–5790.
- [121] Godrich H, Haimovich A M, Blum R S. Target localization accuracy gain in MIMO radar-based systems[J]. *IEEE Transactions on Information Theory*, 2010, 56(6): 2783–2803.
- [122] Van Trees H L, Bell K L. Bayesian bounds for parameter estimation and nonlinear filtering/tracking[J]. *AMC*, 2007, 10(12): 10–1109.
- [123] Shen Y, Dai W, Win M Z. Power optimization for network localization[J]. *IEEE/ACM Transactions on Networking*, 2013, 22(4): 1337–1350.
- [124] Alirezaei G, Taghizadeh O, Mathar R. Optimum power allocation in sensor networks for active radar applications[J]. *IEEE Transactions on Wireless Communications*, 2015, 14(5): 2854–2867.
- [125] He Q, Blum R S, Haimovich A M. Noncoherent MIMO radar for location and velocity estimation: More antennas means better performance[J]. *IEEE Transactions on Signal Processing*, 2010, 58(7): 3661–3680.
- [126] Tajer A, Jajamovich G H, Wang X, et al. Optimal joint target detection and parameter estimation by MIMO radar[J]. *IEEE Journal of Selected Topics in Signal Processing*, 2010, 4(1): 127–145.
- [127] Clausen J. Branch and bound algorithms-principles and examples[J]. Department of Computer Science, University of Copenhagen, 1999: 1–30.
- [128] Fang F, Wang K, Ding Z, et al. Energy-efficient resource allocation for NOMA-MEC networks with imperfect CSI[J]. *IEEE Transactions on Communications*, 2021, 69(5): 3436–3449.
- [129] Yu J-J, Zhao M, Li W-T, et al. Joint offloading and resource allocation for time-sensitive multi-access edge computing network[C] // 2020 IEEE Wireless Communications and Networking Conference (WCNC). 2020: 1–6.
- [130] Luo C-G, Deng B, Wang H-Q, et al. High-resolution terahertz coded-aperture imaging for near-field three-dimensional target[J]. *Applied optics*, 2019, 58(12): 3293–3300.
- [131] Guerci J, Bergin J, Guerci R, et al. A new MIMO clutter model for cognitive radar[C] // 2016 IEEE Radar Conference (RadarConf). 2016: 1–6.
- [132] Huang Z, Wang K, Liu A, et al. Joint pilot optimization, target detection and channel estimation for integrated sensing and communication systems[J]. *IEEE Transactions on Wireless Communications*, 2022, 21(12): 10351–10365.
- [133] Wild T, Braun V, Viswanathan H. Joint design of communication and sensing for beyond 5G and 6G systems[J]. *IEEE Access*, 2021, 9: 30845–30857.



- 
- [134] Tian Y, Zhang Z, Huang C, et al. Intelligent reflecting surface aided computational imaging exploiting reed-muller sequences[C] //2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall). 2021 : 1–5.
- [135] Luo Z-Q, Zhang S. Dynamic spectrum management: Complexity and duality[J]. IEEE journal of selected topics in signal processing, 2008, 2(1) : 57–73.
- [136] Shi Q, Razaviyayn M, Luo Z-Q, et al. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel[J]. IEEE Transactions on Signal Processing, 2011, 59(9) : 4331–4340.
- [137] Marler R T, Arora J S. Survey of multi-objective optimization methods for engineering[J]. Structural and multidisciplinary optimization, 2004, 26 : 369–395.
- [138] Bezdek J C, Hathaway R J. Convergence of alternating optimization[J]. Neural, Parallel & Scientific Computations, 2003, 11(4) : 351–368.
- [139] Yeh J. Real analysis: Theory of measure and integration second edition[M]. Singapore : World Scientific Publishing Company, 2006.
- [140] Matsaglia G, PH Styan G. Equalities and inequalities for ranks of matrices[J]. Linear and multilinear Algebra, 1974, 2(3) : 269–292.
- [141] Yu W, Rhee W, Boyd S, et al. Iterative water-filling for Gaussian vector multiple-access channels[J]. IEEE Transactions on Information Theory, 2004, 50(1) : 145–152.
- [142] Garcia N, Haimovich A M, Coulon M, et al. Resource allocation in MIMO radar with multiple targets for non-coherent localization[J]. IEEE Transactions on Signal Processing, 2014, 62(10) : 2656–2666.
- [143] Godrich H, Petropulu A P, Poor H V. Power allocation strategies for target localization in distributed multiple-radar architectures[J]. IEEE Transactions on Signal Processing, 2011, 59(7) : 3226–3240.



## 附录

### A.1 第 3 章相关证明

#### A.1.1 引理 3.2 的证明

在证明之前，先给出一个有用的引理：

**引理A.1.** (*S-procedure*<sup>[107]</sup>) 考虑如下矢量函数  $\mathbf{f}_m(\mathbf{x})$ ：

$$\mathbf{f}_m(\mathbf{x}) = \mathbf{x}^H \mathbf{A}_m \mathbf{x} + 2 \operatorname{Re} \{ \mathbf{b}_m^H \mathbf{x} \} + \mathbf{c}_m, m \in \{1, 2\}, \mathbf{x} \in \mathbb{C}^{N \times 1}, \quad (\text{A.1})$$

其中  $\mathbf{A}_m \in \mathbb{C}^{N \times N}$ ,  $\mathbf{b}_m \in \mathbb{C}^{N \times 1}$ ,  $\mathbf{c}_m \in \mathbb{C}^{N \times 1}$ 。当且仅当存在  $\tau \geq 0$  时，推导  $\mathbf{f}_1(\mathbf{x}) \leq 0 \Rightarrow \mathbf{f}_2(\mathbf{x}) \leq 0$  成立，则有

$$\tau \begin{bmatrix} \mathbf{A}_1 & \mathbf{b}_1 \\ \mathbf{b}_1^H & \mathbf{c}_1 \end{bmatrix} - \begin{bmatrix} \mathbf{A}_2 & \mathbf{b}_2 \\ \mathbf{b}_2^H & \mathbf{c}_2 \end{bmatrix} \succeq \mathbf{0}. \quad (\text{A.2})$$

根据半正定矩阵的定义，对于  $\forall \mathbf{d} \neq \mathbf{0}$ ，有

$$\mathbf{d}^H \mathbf{A} \mathbf{d} - \mathbf{d}^H (\mathbf{B}^H \mathbf{x} \mathbf{c}^H + \mathbf{c} \mathbf{x}^H \mathbf{B}) \mathbf{d} \geq 0. \quad (\text{A.3})$$

这表明  $\mathbf{F}(\mathbf{x}) \succeq \mathbf{0}$ ,  $\forall \mathbf{x} : \|\mathbf{x}\| \leq \varpi$  成立。基于引理 3.1，式 (A.3) 等价于

$$\begin{aligned} \mathbf{d}^H \mathbf{A} \mathbf{d} &\geq \max_{\|\mathbf{x}\| \leq \varpi} \mathbf{d}^H (\mathbf{B}^H \mathbf{x} \mathbf{c}^H + \mathbf{c} \mathbf{x}^H \mathbf{B}) \mathbf{d} \\ &= 2\varpi |\mathbf{c}^H \mathbf{d}| \|\mathbf{B} \mathbf{d}\|. \end{aligned} \quad (\text{A.4})$$

利用柯西-施瓦兹不等式<sup>[140]</sup>, 式 (A.4) 可以进一步转换为

$$\mathbf{d}^H \mathbf{A} \mathbf{d} - 2\varpi \xi \mathbf{c}^H \mathbf{d} \geq 0, \forall \xi: |\xi| \leq \|\mathbf{B} \mathbf{d}\|. \quad (\text{A.5})$$

接着, 利用引理 A.1 和  $\varpi^2 - \mathbf{d}^H \mathbf{B}^H \mathbf{B} \mathbf{d} \leq 0$ , 当且仅当存在  $\lambda \geq 0$  时, 满足式 (A.5), 使得

$$\begin{bmatrix} \mathbf{A} - \lambda \mathbf{c} \mathbf{c}^H & -\varpi \mathbf{B}^H \\ -\varpi \mathbf{B} & \lambda \mathbf{I}_n \end{bmatrix} \succeq \mathbf{0}. \quad (\text{A.6})$$

至此, 引理 3.2 证明完毕。

## A.2 第 4 章相关证明

### A.2.1 引理 4.1 的证明

为了便于推导, 定义  $\mathbf{b}_i$  为  $\mathbf{B}$  的第  $i$  列向量,  $\mathbf{c}_j$  为  $\mathbf{C}$  的第  $j$  列向量, 其中  $c_{i,j}$  表示  $\mathbf{C}$  在第  $i$  行第  $j$  列的元素。根据克罗内克积的定义, 可以得到

$$\begin{aligned} (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}) &= \begin{bmatrix} c_{11} \mathbf{A} & c_{21} \mathbf{A} & \cdots & c_{p1} \mathbf{A} \\ c_{12} \mathbf{A} & c_{22} \mathbf{A} & \cdots & c_{p2} \mathbf{A} \\ \vdots & \vdots & \ddots & \vdots \\ c_{1q} \mathbf{A} & c_{2q} \mathbf{A} & \cdots & c_{pq} \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A} \sum_{j=1}^p c_{j1} \mathbf{b}_1 \\ \mathbf{A} \sum_{j=1}^p c_{j2} \mathbf{b}_2 \\ \vdots \\ \mathbf{A} \sum_{j=1}^p c_{jq} \mathbf{b}_p \end{bmatrix} = \begin{bmatrix} \mathbf{A} \mathbf{B} \mathbf{c}_1 \\ \mathbf{A} \mathbf{B} \mathbf{c}_2 \\ \cdots \\ \mathbf{A} \mathbf{B} \mathbf{c}_q \end{bmatrix} = \text{vec}(\mathbf{A} \mathbf{B} \mathbf{C}). \end{aligned} \quad (\text{A.7})$$

至此, 引理 4.1 证明完毕。

### A.2.2 引理 4.2 的证明

根据克罗内克积和矩阵行列式的性质, 可以推出

$$\begin{aligned} |\mathbf{I}_{mn} + \mathbf{A} \mathbf{B} \otimes \mathbf{C} \mathbf{D}| &= |\mathbf{I}_{mn} + (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})| \\ &= 1 \times \begin{vmatrix} \mathbf{I}_{mn} + (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) & \mathbf{0} \\ \mathbf{B} \otimes \mathbf{D} & \mathbf{I}_{mn} \end{vmatrix} \times 1 \end{aligned}$$

$$\begin{aligned}
&= \begin{vmatrix} \mathbf{I}_{mn} & -\mathbf{A} \otimes \mathbf{C} \\ \mathbf{0} & \mathbf{I}_{mn} \end{vmatrix} \begin{vmatrix} \mathbf{I}_{mn} + (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D}) & \mathbf{0} \\ \mathbf{B} \otimes \mathbf{D} & \mathbf{I}_{mn} \end{vmatrix} \begin{vmatrix} \mathbf{I}_{mn} & -\mathbf{A} \otimes \mathbf{C} \\ \mathbf{0} & \mathbf{I}_{mn} \end{vmatrix} \\
&= \begin{vmatrix} \mathbf{I}_{mn} & \mathbf{0} \\ \mathbf{B} \otimes \mathbf{D} & (\mathbf{B} \otimes \mathbf{D})(\mathbf{A} \otimes \mathbf{C}) + \mathbf{I}_{mn} \end{vmatrix} \\
&= |\mathbf{I}_{mn} + (\mathbf{B} \otimes \mathbf{D})(\mathbf{A} \otimes \mathbf{C})| \\
&= |\mathbf{I}_{mn} + \mathbf{BA} \otimes \mathbf{DC}|. \tag{A.8}
\end{aligned}$$

至此，引理 4.2 证明完毕。

### A.2.3 $M_s$ 和 $M_c$ 的推导

首先，推导最大可达感知速率  $M_s$ 。显然，当通信终端的数量为零时，即不存在通信干扰，可以获得最大感知速率。在这种情况下， $\sigma_{h,j} = \frac{1}{\sigma_n^2}, \forall j$ 。因此，感知速率最大化的优化问题可以表示为

$$\begin{aligned}
\max_{\sigma_s} \quad & \frac{N_r}{L} \sum_{i=1}^{N_t} \log_2 \left( 1 + \frac{\sigma_{t,i}}{\sigma_n^2} \sigma_{s,i} \right), \\
\text{s.t.} \quad & \sum_{i=1}^{N_t} \sigma_{s,i} \leq P_s.
\end{aligned} \tag{A.9}$$

对于该问题的求解，可以利用著名的注水算法来有效地获得最优解  $\sigma_{s,i}$  [141]。首先，构造相应的拉格朗日函数：

$$\mathcal{L}_s(\sigma_{s,i}) = \frac{N_r}{L} \sum_{i=1}^{N_t} \log_2 \left( 1 + \frac{\sigma_{t,i}}{\sigma_n^2} \sigma_{s,i} \right) + \mu \left( P_s - \sum_{i=1}^{N_t} \sigma_{s,i} \right). \tag{A.10}$$

其中， $\mu > 0$  为拉格朗日乘子。然后，基于 KKT 条件，令  $\mathcal{L}_s(\sigma_{s,i})$  的一阶导数为 0，可以得到

$$\sigma_{s,i} = \frac{L}{N_r \mu} - \frac{\sigma_n^2}{\sigma_{t,i}} = \hat{\mu} - \frac{\sigma_n^2}{\sigma_{t,i}}. \tag{A.11}$$

将式 (A.11) 代入取得最优点的条件  $P_s = \sum_{i=1}^{N_t} \sigma_{s,i}$  中，可以获得问题 A.9 的最优解为

$$\sigma_{s,i}^{\text{opt}} = \max(\hat{\mu} - \sigma_n^2 / \sigma_{t,i}, 0), \tag{A.12}$$

其中  $\hat{\mu} = (P_s + \sum_{i=1}^{N_t} \sigma_n^2 / \sigma_{t,i}) / N_t$ 。因此，可以得出最大化感知速率为

$$M_s = \frac{N_r}{L} \sum_{i=1}^{N_t} \log_2(1 + \sigma_{t,i} \sigma_{s,i}^{\text{opt}} / \sigma_n^2). \quad (\text{A.13})$$

接着，推导最大化通信速率  $M_c$ 。可以看出，当不存在感知发射信号时，即不存在感知干扰，可以获得最大通信速率  $M_c$ 。因此，将  $\sigma_{s,i} = 0, \forall i$  代入式(4.26)中，可以求得最大通信速率为

$$M_c = \frac{1}{K} \sum_{k=1}^K \log_2(1 + p_k \mathbf{h}_k^T \hat{\mathbf{R}}_k^{-1} \mathbf{h}_k^*), \quad (\text{A.14})$$

其中， $\hat{\mathbf{R}}_k = \sum_{i \neq k}^K p_i \mathbf{h}_i^* \mathbf{h}_i^T + \sigma_n^2 \mathbf{I}_{N_r}$ 。至此， $M_s$  和  $M_c$  的推导完毕。

### A.3 第 5 章相关证明

#### A.3.1 FIM的推导

对于未知参数向量  $\mathbf{u}_l = [x_l, y_l, z_l]$  的 FIM 可以表示为：

$$\mathbf{J}_l = \mathbb{E}_{\mathbf{r}_l | \mathbf{u}_l} \left\{ \frac{\partial}{\partial \mathbf{u}_l} \ln f(\mathbf{r}_l | \mathbf{u}_l) \left( \frac{\partial}{\partial \mathbf{u}_l} \ln f(\mathbf{r}_l | \mathbf{u}_l) \right)^T \right\}, \quad (\text{A.15})$$

其中， $f(\mathbf{r}_l | \mathbf{u}_l)$  为条件概率密度函数，可以表示为

$$f(\mathbf{r}_l | \mathbf{u}_l) \propto \exp \left\{ -\frac{1}{\sigma_w^2} \sum_{m=1}^M \sum_{n=1}^N \int_{\mathcal{T}_m} \left| r_{m,l,n}(t) - \sum_{k=1}^K \sqrt{p_m} h_{m,l,n,k} s_m(t - \tau_{m,l,n}) \right|^2 dt \right\}. \quad (\text{A.16})$$

由于式 (A.16) 是传播时延  $\tau_{m,l,n}$  的函数，可以先计算关于时延向量  $\boldsymbol{\tau}_l = [\tau_{1,l,1}, \dots, \tau_{M,1,N}]^T$  的 FIM。接着，利用链式法则得到关于向量  $\mathbf{u}_l$  的 FIM<sup>[121, 142]</sup>，即

$$\mathbf{J}_l = \mathbf{Q}_l \mathbf{J}_l(\boldsymbol{\tau}_l) \mathbf{Q}_l^T, \quad (\text{A.17})$$

其中，雅可比矩阵  $\mathbf{Q}_l = \frac{\partial \boldsymbol{\tau}_l}{\partial \mathbf{u}_l}$  由下式给出：

$$\mathbf{Q}_l = \frac{1}{c} \begin{bmatrix} \sin \psi_{1,l}^T \cos \varphi_{1,l}^T + \sin \psi_{1,l}^R \cos \varphi_{1,l}^R & \dots & \sin \psi_{M,l}^T \cos \varphi_{M,l}^T + \sin \psi_{N,l}^R \cos \varphi_{N,l}^R \\ \sin \psi_{1,l}^T \sin \varphi_{1,l}^T + \sin \psi_{1,l}^R \sin \varphi_{1,l}^R & \dots & \sin \psi_{M,l}^T \sin \varphi_{M,l}^T + \sin \psi_{N,l}^R \sin \varphi_{N,l}^R \\ \cos \psi_{1,l}^T + \cos \psi_{1,l}^R & \dots & \cos \psi_{M,l}^T + \cos \psi_{N,l}^R \end{bmatrix}, \quad (\text{A.18})$$

以及 FIM  $\mathbf{J}_l(\boldsymbol{\tau}_l)$  的元素可以表示为

$$\mathbf{J}_l(\boldsymbol{\tau}_l)[i, j] = \frac{\partial^2 [\ln f(\mathbf{r}_l | \mathbf{u}_l)]}{\partial \tau_{m,l,n} \partial \tau_{z,l,q}}, \quad (\text{A.19})$$

其中  $i = (n-1) + m$ ， $j = (q-1) + z$ 。因此，FIM  $\mathbf{J}_l(\boldsymbol{\tau}_l)$  可以表示为

$$\mathbf{J}_l(\boldsymbol{\tau}_l) = \text{diag}[\Upsilon_{1,l,1}, \dots, \Upsilon_{M,l,N}]^{MN \times MN}, \quad (\text{A.20})$$

其中  $\Upsilon_{m,l,n} = 8\pi^2 B^2 p_m \sum_{k=1}^K |h_{m,l,n,k}|^2 / \sigma_w^2$  是对角矩阵  $\mathbf{J}_l(\boldsymbol{\tau}_l)$  主对角线上的元素<sup>[143]</sup>。将式 (A.18)-(A.20) 代入式 (A.17) 中，可以得到最终的  $\mathbf{J}_l$ 。至此，FIM 的推导完毕。

### A.3.2 命题 5.1 的证明

为了方便表示，重写式 (5.4) 为

$$\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) = \frac{\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}}{\bar{\mathbf{p}}^T \left( \sum_{i=1}^3 \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \right) \bar{\mathbf{p}}}, \quad (\text{A.21})$$

其中  $\mathbf{g}_l^1 = \mathbf{g}_l^x$ ,  $\mathbf{B}_l^1 = \mathbf{B}_l^x$ ,  $\mathbf{g}_l^2 = -\mathbf{g}_l^{xy}$ ,  $\mathbf{B}_l^2 = \mathbf{B}_l^{xy}$ ,  $\mathbf{g}_l^3 = \mathbf{g}_l^{xz}$ ,  $\mathbf{B}_l^3 = \mathbf{B}_l^{xz}$ 。接着，对  $\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}))$  进行一阶求导，得到

$$\begin{aligned} \nabla_{\bar{\mathbf{p}}} \text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) &= \frac{\mathbf{A}_l \bar{\mathbf{p}}}{\sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}}} + \frac{\mathbf{A}_l^T \bar{\mathbf{p}}}{\sum_{i=1}^3 \left[ (\mathbf{g}_l^i)^T \bar{\mathbf{p}} \right] \left[ \bar{\mathbf{p}}^T (\mathbf{B}_l^i)^T \bar{\mathbf{p}} \right]} - \frac{(\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \mathbf{g}_l^i}{\left[ \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \right]^2} \\ &- \frac{(\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \mathbf{B}_l^i \bar{\mathbf{p}}}{\left[ \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \right]^2} - \frac{(\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 (\mathbf{g}_l^i)^T \bar{\mathbf{p}} (\mathbf{B}_l^i)^T \bar{\mathbf{p}}}{\left( \sum_{i=1}^3 \left[ (\mathbf{g}_l^i)^T \bar{\mathbf{p}} \right] \left[ \bar{\mathbf{p}}^T (\mathbf{B}_l^i)^T \bar{\mathbf{p}} \right] \right)^2}. \end{aligned} \quad (\text{A.22})$$

进一步整理式 (A.22)，可以得到

$$\begin{aligned} \nabla_{\bar{\mathbf{p}}}\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) &= \frac{(\mathbf{A}_l + \mathbf{A}_l^T) \bar{\mathbf{p}} \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}}}{\left[ \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \right]^2} - \frac{(\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 (\bar{\mathbf{p}}^T \mathbf{g}_l^i) [\mathbf{B}_l^i + (\mathbf{B}_l^i)^T] \bar{\mathbf{p}}}{\left[ \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \right]^2} \\ &\quad - \frac{(\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 (\bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}}) \mathbf{g}_l^i}{\left[ \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} \right]^2}. \end{aligned} \quad (\text{A.23})$$

从式 (A.23) 可以看出，对于非零向量  $\bar{\mathbf{p}}$ ，一阶导数  $\nabla_{\bar{\mathbf{p}}}\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}))$  的分母总是大于等于零。因此，只需要检查一阶导数的分子  $\nabla_{\bar{\mathbf{p}}}^{\text{Numerator}}$  的正负性，即

$$\begin{aligned} \nabla_{\bar{\mathbf{p}}}^{\text{Numerator}} &= (\mathbf{A}_l + \mathbf{A}_l^T) \bar{\mathbf{p}} \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}} - (\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 (\bar{\mathbf{p}}^T \mathbf{g}_l^i) [\mathbf{B}_l^i + (\mathbf{B}_l^i)^T] \bar{\mathbf{p}} \\ &\quad - \left( \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \right) (\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 [\mathbf{B}_l^i + (\mathbf{B}_l^i)^T] \bar{\mathbf{p}}. \end{aligned} \quad (\text{A.24})$$

然后，在分子  $\nabla_{\bar{\mathbf{p}}}^{\text{Numerator}}$  上左乘一个非零向量  $\bar{\mathbf{p}}^T$ ，得到

$$\bar{\mathbf{p}}^T \nabla_{\bar{\mathbf{p}}}^{\text{Numerator}} = - (\bar{\mathbf{p}}^T \mathbf{A}_l \bar{\mathbf{p}}) \sum_{i=1}^3 \bar{\mathbf{p}}^T \mathbf{g}_l^i \bar{\mathbf{p}}^T \mathbf{B}_l^i \bar{\mathbf{p}}. \quad (\text{A.25})$$

根据式 (A.21) 中的  $\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}})) > 0$ ，可以知道式 (A.25) 中的  $\bar{\mathbf{p}}^T \nabla_{\bar{\mathbf{p}}}^{\text{Numerator}} < 0$ ，这意味着对于非零向量  $\bar{\mathbf{p}}$ ，一阶导数  $\nabla_{\bar{\mathbf{p}}}\text{tr}(\mathbf{C}_l^{x,y,z}(\bar{\mathbf{p}}))$  总是小于等于零。因此CRB矩阵的迹是  $\bar{\mathbf{p}}$  的单调递减函数。至此，命题 5.1 证明完毕。

### A.3.3 问题 (5.23) 的凸性证明

对于问题 (5.23)，定义其目标函数为

$$\mathcal{U}(\mathbf{q}, \mathbf{f}, \mathbf{F}) = \sum_{m=1}^M \lambda_m (E_m^{\text{loc}} + E_m^{\text{tra}}(q_m) + E_m^{\text{ser}}). \quad (\text{A.26})$$



则计算  $\mathcal{U}(\mathbf{q}, \mathbf{f}, \mathbf{F})$  对  $\mathbf{q}$ ,  $\mathbf{f}$  和  $\mathbf{F}$  的二阶导数, 得到

$$\nabla_{\mathbf{q}}^2 \mathcal{U} = \frac{\lambda_m D_m (\ln 2)^2 (1 - \vartheta_{m,0})^2}{B \chi_m} \cdot \frac{2^{\frac{1-\vartheta_{m,0}}{q_m}}}{q_m^3} \geq 0, \quad (\text{A.27a})$$

$$\nabla_{\mathbf{f}}^2 \mathcal{U} = 2\lambda_m \vartheta_{m,0} \kappa_m^U D_m C_m f_m^{\text{loc}} \geq 0, \quad (\text{A.27b})$$

$$\nabla_{\mathbf{F}}^2 \mathcal{U} = 2\lambda_m \vartheta_{m,n} \kappa_n^M D_m C_m f_{m,n}^{\text{ser}} \geq 0. \quad (\text{A.27c})$$

由式(A.27a)-(A.27c)可知, 目标函数  $\mathcal{U}(\mathbf{q}, \mathbf{f}, \mathbf{F})$  是变量  $\mathbf{q}$ ,  $\mathbf{f}$  和  $\mathbf{F}$  的联合凸函数。同时, 约束条件 (5.9f)-(5.9i) 因其线性结构均为凸集。此外, 就  $\mathbf{q}$  和  $\Theta$  而言, 约束条件 (5.14) 是一个LMI。对于约束(5.19), 定义  $\mathcal{L}_m(q_m, f_m^{\text{loc}}, f_{m,n}^{\text{ser}})$  为

$$\mathcal{L}_m(q_m, f_m^{\text{loc}}, f_{m,n}^{\text{ser}}) = T_m^{\text{loc}} + \frac{q_m D_m}{B(1 - \vartheta_{m,0})} + T_m^{\text{ser}} - \tau_m. \quad (\text{A.28})$$

相似地,  $\mathcal{L}_m(q_m, f_m^{\text{loc}}, f_{m,n}^{\text{ser}})$  就  $q_m$ ,  $f_m^{\text{loc}}$  和  $f_{m,n}^{\text{ser}}$  的二阶导数均大于等于零, 即

$$\nabla_{q_m}^2 \mathcal{L}_m = 0, \quad (\text{A.29a})$$

$$\nabla_{f_m^{\text{loc}}}^2 \mathcal{L}_m = \frac{2\vartheta_{m,0} D_m C_m}{(f_m^{\text{loc}})^3} \geq 0, \quad (\text{A.29b})$$

$$\nabla_{f_{m,n}^{\text{ser}}}^2 \mathcal{L}_m = \frac{2\vartheta_{m,n} D_m C_m}{(f_{m,n}^{\text{ser}})^3} \geq 0. \quad (\text{A.29c})$$

因此, 函数  $\mathcal{L}_m(q_m, f_m^{\text{loc}}, f_{m,n}^{\text{ser}})$  是  $q_m$ ,  $f_m^{\text{loc}}$  和  $f_{m,n}^{\text{ser}}$  的联合凸函数, 这意味着约束条件 (5.19) 也是凸集。再者, 由于特征值的和的线性结构, 约束条件 (5.22) 也是凸的。综上所述, 由于目标函数和约束条件均是凸的, 问题 (5.23) 是一个标准凸优化问题。至此, 问题 (5.23) 的凸性证明完毕。

## A.4 第 6 章相关证明

### A.4.1 定理 6.1 的证明

对于基站接收的通信信号  $y_{k,j}^{\text{comm}}$ , 关于来自第  $k$  个终端的第  $j$  个通信信号的 MSE 可以表示为

$$\text{MSE}_{k,j}^{\text{comm}} = \mathbb{E} \left\{ (y_{k,j}^{\text{comm}} - s_{k,j}^{\text{comm}}) (y_{k,j}^{\text{comm}} - s_{k,j}^{\text{comm}})^H \right\}$$

$$\begin{aligned}
 & + \sum_{i=1}^K \sum_{m=1}^J \mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{a}_{i,m} \mathbf{a}_{i,m}^H \mathbf{H}_i^H \mathbf{u}_{k,j} + \sum_{i=1}^K \sum_{l=1}^L \mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{b}_{i,l} \mathbf{b}_{i,l}^H \mathbf{H}_i^H \mathbf{u}_{k,j} \\
 & + \sum_{i=1}^K \sum_{n=1}^I R_n^2 \mathbf{u}_{k,j}^H \mathbf{G}_{i,n} \mathbf{c}_{i,n} \mathbf{c}_{i,n}^H \mathbf{G}_{i,n}^H \mathbf{u}_{k,j} + \sum_{i=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 \mathbf{u}_{k,j}^H \mathbf{F}_{i,o} \mathbf{b}_{i,m} \mathbf{b}_{i,m}^H \mathbf{F}_{i,o}^H \mathbf{u}_{k,j} \\
 & + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2 - \mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{a}_{k,j} - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \mathbf{u}_{k,j} + 1.
 \end{aligned} \tag{A.30}$$

通过从上式中提取公因式，并定义  $\Xi = \sum_{i=1}^K \sum_{m=1}^J \mathbf{H}_i \mathbf{c}_{i,m} \mathbf{c}_{i,m}^H \mathbf{H}_i^H + \sum_{i=1}^K \sum_{l=1}^L \mathbf{H}_i \mathbf{b}_{i,l} \mathbf{b}_{i,l}^H \mathbf{H}_i^H + \sum_{i=1}^K \sum_{n=1}^I R_n^2 \mathbf{G}_{i,n} \mathbf{a}_{i,n} \mathbf{a}_{i,n}^H \mathbf{G}_{i,n}^H + \sum_{i=1}^K \sum_{o=1}^O \sum_{m=1}^I \bar{R}_o^2 \mathbf{F}_{i,o} \mathbf{a}_{i,m} \mathbf{a}_{i,m}^H \mathbf{F}_{i,o}^H + \sigma_n^2 \mathbf{I}$ ，式 (A.30) 可以重新表示为

$$\begin{aligned}
 \text{MSE}_{k,j}^{\text{comm}} & = \mathbf{u}_{k,j}^H \Xi \mathbf{u}_{k,j} - \mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{a}_{k,j} - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \mathbf{u}_{k,j} + 1 \\
 & = (\mathbf{u}_{k,j}^H - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \Xi^{-1}) \Xi (\mathbf{u}_{k,j}^H - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \Xi^{-1})^H \\
 & \quad - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \Xi^{-H} \mathbf{H}_k \mathbf{a}_{k,j} + 1.
 \end{aligned} \tag{A.31}$$

可以发现，当  $\mathbf{u}_{k,j} = \Xi^{-1} \mathbf{H}_k \mathbf{c}_{k,j}$  时， $\text{MSE}_{k,j}^{\text{comm}}$  可以被最小化。因此，关于通信信号  $s_{k,j}^{\text{comm}}$  的 MMSE 可以表示为

$$\begin{aligned}
 e_{k,j}^{\text{comm}} & = 1 - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \Xi^{-H} \mathbf{H}_k \mathbf{a}_{k,j} = \frac{\Xi^H - \mathbf{a}_{k,j}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{a}_{k,j}}{\Xi^H} \\
 & = \frac{\mathbf{u}_{k,j}^H \Xi^H \mathbf{u}_{k,j} - \mathbf{u}_{k,j}^H \mathbf{a}_{k,j}^H \mathbf{H}_k^H \mathbf{H}_k \mathbf{a}_{k,j} \mathbf{u}_{k,j}}{\mathbf{u}_{k,j}^H \Xi^H \mathbf{u}_{k,j}} = \frac{1}{1 + \Gamma_{k,j}}.
 \end{aligned} \tag{A.32}$$

可见，通信信号的 MMSE 等价于其 SINR 加 1 的结果的倒数，这有助于问题 S-3' 的目标函数进行等价转换。至此，定理 6.1 证明完毕。

#### A.4.2 定理 6.2 的证明

证明之前，需要先介绍下面两个有用的引理。

**引理A.2.** 若存在矩阵  $\mathbf{Q} \in \mathbb{C}^{t \times n}$  和矩阵  $\mathbf{P} \in \mathbb{C}^{n \times s}$ ,  $\forall t, n, s$ ，满足  $\mathbf{QP} = \mathbf{0}$ ，基于西尔维斯特不等式<sup>[140]</sup>，可以推出  $\text{Rank}(\mathbf{Q}) + \text{Rank}(\mathbf{P}) \leq n$ 。

**引理A.3.** 对于两个具有相同尺寸的矩阵  $\mathbf{Q}$  和  $\mathbf{P}$ ，不等式  $\text{Rank}(\mathbf{Q} + \mathbf{P}) \leq \text{Rank}(\mathbf{Q}) + \text{Rank}(\mathbf{P})$  成立。

证明:

$$\begin{aligned} \text{Rank}(\mathbf{Q} + \mathbf{P}) &= \text{Rank} \begin{bmatrix} \mathbf{Q} + \mathbf{P} \\ \mathbf{0} \end{bmatrix} \leq \text{Rank} \begin{bmatrix} \mathbf{Q} + \mathbf{P} \\ \mathbf{P} \end{bmatrix} \\ &= \text{Rank} \begin{bmatrix} \mathbf{Q} \\ \mathbf{P} \end{bmatrix} \leq \text{Rank}(\mathbf{Q}) + \text{Rank}(\mathbf{P}). \end{aligned} \quad (\text{A.33})$$

□

接着, 建立问题 M-2” 关于  $\mathbf{A}_{k,j}$  的拉格朗日函数,

$$\begin{aligned} \mathcal{L}(\mathbf{A}_{k,j}) &= \sum_{k=1}^K \left( \sum_{j=1}^J \text{tr}(\mathbf{A}_{k,j}) + \sum_{i=1}^I \|\mathbf{b}_{k,i}\|^2 + \sum_{l=1}^L \|\mathbf{c}_{k,l}\|^2 \right) + \sum_{k=1}^K \sum_{j=1}^J \mu_{k,j} T_{k,j} \\ &+ \sum_{i=1}^I \lambda_i [\overline{\text{MSE}}_i^{\text{sens}} - \delta_i] + \sum_{l=1}^L \beta_l [\overline{\text{MSE}}_l^{\text{comp}} - \chi_l] - \sum_{k=1}^K \sum_{j=1}^J \Theta_{k,j} \mathbf{A}_{k,j}, \end{aligned} \quad (\text{A.34})$$

其中  $T_{k,j} = \sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{A}_{i,n} \mathbf{H}_i^H \mathbf{u}_{k,j}) + X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2 - \frac{1}{\gamma_{k,j}} \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{A}_{k,j} \mathbf{H}_k^H \mathbf{u}_{k,j})$ 。

此外,  $\mu_{k,j}$ ,  $\lambda_i$ ,  $\beta_l$  和  $\Theta_{k,j}$  分别为关于约束条件 C5, C6, C7 和 C8 的拉格朗日乘子。为了探索满足斯莱特条件的最优解  $\mathbf{A}_{k,j}^*$ , 利用下面的 KKT 条件:

$$\sum_{i=1, i \neq k}^K \sum_{n=1, n \neq j}^J \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_i \mathbf{A}_{i,n}^* \mathbf{H}_i^H \mathbf{u}_{k,j}) + X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2 - \frac{1}{\gamma_{k,j}} \text{tr}(\mathbf{u}_{k,j}^H \mathbf{H}_k \mathbf{A}_{k,j}^* \mathbf{H}_k^H \mathbf{u}_{k,j}) = 0, \quad (\text{A.35a})$$

$$\Theta_{k,j}^* \mathbf{A}_{k,j}^* = \mathbf{0}, \quad (\text{A.35b})$$

$$\nabla_{\mathbf{A}_{k,j}^*} \mathcal{L} = \mathbf{I}_M + \sum_{i=1}^I \lambda_i^* \mathbf{H}_k^H \mathbf{v}_i \mathbf{v}_i^H \mathbf{H}_k + \sum_{l=1}^L \beta_l^* \mathbf{H}_k^H \mathbf{z}_l \mathbf{z}_l^H \mathbf{H}_k - \frac{\mu_{k,j}^*}{\gamma_{k,j}} \mathbf{H}_k^H \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H \mathbf{H}_k - \Theta_{k,j}^* = \mathbf{0}, \quad (\text{A.35c})$$

$$\mu_{k,j}^* \geq 0, \lambda_i^* \geq 0, \beta_l^* \geq 0, \Theta_{k,j}^* \succeq \mathbf{0}. \quad (\text{A.35d})$$

在式 (A.35a) 中, 由于  $X_{k,j} + \sigma_n^2 \|\mathbf{u}_{k,j}\|^2 > 0$  可以推出  $\mathbf{A}_{k,j}^* \neq \mathbf{0}$ , 这意味着

$$\text{Rank}(\mathbf{A}_{k,j}^*) \geq 1. \quad (\text{A.36})$$

接着, 利用式 (A.35b) 中  $\Theta_{k,j}^*$  和  $\mathbf{A}_{k,j}^*$  的关系, 基于引理 A.2 可得

$$\text{Rank}(\Theta_{k,j}^*) + \text{Rank}(\mathbf{A}_{k,j}^*) \leq M. \quad (\text{A.37})$$

然后，将式 (A.36) 带入式 (A.37) 中，得到

$$\text{Rank}(\Theta_{k,j}^*) \leq M - 1. \quad (\text{A.38})$$

基于引理 A.3，可以从式 (A.35c) 中得出

$$\text{Rank}(\Upsilon_{k,j}) + \text{Rank}(\Theta_{k,j}^*) \geq \text{Rank}(\mathbf{I}_M), \quad (\text{A.39})$$

其中  $\Upsilon_{k,j} = \mathbf{H}_k^H \left( \frac{\mu_{k,j}^*}{\gamma_{k,j}} \mathbf{u}_{k,j} \mathbf{u}_{k,j}^H - \sum_{i=1}^I \lambda_i^* \mathbf{v}_i \mathbf{v}_i^H - \sum_{l=1}^L \beta_l^* \mathbf{z}_l \mathbf{z}_l^H \right) \mathbf{H}_k$ 。由于  $\Upsilon_{k,j} \neq \mathbf{0}$  且  $\text{Rank}(\mathbf{I}_M) = M$ ，可推出

$$\text{Rank}(\Theta_{k,j}^*) \geq M - 1. \quad (\text{A.40})$$

基于式 (A.38) 和式 (A.40)，显然， $\text{Rank}(\Theta_{k,j}^*) = M - 1$ 。最后，将其带入式 (A.37) 中，再结合式 (A.36)，可以总结出

$$\text{Rank}(\mathbf{A}_{k,j}^*) = 1, \quad (\text{A.41})$$

这证明了采取的 SDR 是紧的，即在问题 M-2' 中丢弃约束条件 C9。至此，定理 6.2 证明完毕。

## 攻读博士学位期间参与项目及研究成果

### 参与的科研项目：

1. 国家自然科学基金区域创新重点项目 (No. U21A20443)：低轨卫星物联网的移动覆盖和巨址接入理论与关键技术，2022.01-2025.12.
2. 浙江省自然科学基金杰青项目 (No. LR20F010002)：面向 B5G 蜂窝物联网的大规模接入理论和技术研究，2020.01-2023.12.
3. 华为合作项目 (No. TC20210521010)：有限带宽均衡算法研究，2021.06-2022.06.
4. 华为合作项目 (No. YBN2018115223)：感知、通信、计算协同融合，2019.01-2020.12.

### 受资助的项目：

1. 2022.06-2023.06，受浙江大学争创优秀博士学位论文项目资助。
2. 2021.10-2022.09，受国家留学基金委资助前往新加坡科技大学进行联合培养。

### 重要获奖情况：

1. 浙江大学竺可桢奖学金
2. 浙江大学十佳大学生
3. 浙江大学学生十大学术新成果奖
4. 浙江大学优秀共产党员
5. 浙江省优秀毕业生
6. 浙江大学信电学院ISEE荣誉学子
7. 国家奖学金

## 期刊论文:

1. **Qi Q**, Chen X. Wireless powered massive access for cellular Internet of Things with imperfect SIC and non-Linear EH[J]. **IEEE Internet of Things Journal**, 2019, 6(2):3110-3120. (SCI)
2. **Qi Q**, Chen X, Lei L, Zhong C, Zhang Z. Outage-constrained robust design for sustainable B5G cellular Internet of Things[J]. **IEEE Transactions on Wireless Communications**, 2019, 18(12):5780-5790. (SCI)
3. **Qi Q**, Chen X, Ng D W K. Robust beamforming for NOMA-based cellular massive IoT with SWIPT[J]. **IEEE Transactions on Signal Processing**, 2019, 68(1):211-224. (SCI)
4. **Qi Q**, Chen X, Zhong C, Zhang Z. Integrated sensing, computation and communication in B5G cellular Internet of Things[J]. **IEEE Transactions on Wireless Communications**, 2021, 20(1):332-344. (SCI)
5. **Qi Q**, Chen X. Robust design of federated learning for edge-intelligent networks[J]. **IEEE Transactions on Communications**, 2022, 70(7):4469-4481. (SCI)
6. **Qi Q**, Chen X, Khalili A, Zhong C, Zhang Z, Ng D W K. Integrating sensing, computing, and communication in 6G wireless networks: Design and optimization[J]. **IEEE Transactions on Communications**, 2022, 70(9):6212-6227. (SCI)
7. **Qi Q**, Chen X, Zhong C, Zhang Z. Physical layer security for massive access in cellular Internet of Things[J]. **SCIENCE CHINA Information Sciences**, 2020, 63(2):1-12. (SCI)
8. **Qi Q**, Chen X, Zhong C, Zhang Z. Integration of energy, computation and communication in 6G cellular Internet of Things[J]. **IEEE Communications Letters**, 2020, 24(6):1333-1337. (SCI)
9. Jia R, Chen X, **Qi Q**. Massive beam division multiple access for B5G cellular Internet of Things[J]. **IEEE Internet of Things Journal**, 2020, 24(6):1333-1337. (SCI)
10. Wang Q, Chen X, **Qi Q**. Task-driven robust integration of communication and computation for edge-intelligent networks[J]. **IEEE Transactions on Communications**, 2022, 71(1):244-255. (SCI)

11. **Qi Q**, Chen X, Zhong C, Yuen C, Zhang Z. Deep learning-based design of uplink integrated sensing and communication[J]. Submitted to **IEEE Transactions on Wireless Communications**, 2023. (SCI, 投稿中)
12. **Qi Q**, Chen X, Yuen C. Joint offloading selection and resource allocation for integrated localization and computing in edge-intelligent networks[J]. Submitted to **IEEE Internet of Things Journal**, 2023. (SCI, 投稿中)
13. Wang Q, Chen X, **Qi Q**. Energy-efficient design of satellite-terrestrial Computing in 6G Wireless Networks[J]. Submitted to **IEEE Transactions on Communications**, 2023. (SCI, 投稿中)

会议论文:

1. **Qi Q**, Chen X, Ng D W K, Zhong C, Zhang Z. Robust beamforming design for SWIPT in cellular Internet of Things[C] // 2019 IEEE/CIC International Conference on Communications in China (**IEEE CIC/ICCC**) 2019:523-528. (EI)
2. **Qi Q**, Chen X, Lei L, Zhong C, Zhang Z. Robust convergence of energy and computation for B5G cellular Internet of Things[C] // 2019 IEEE Global Communications Conference (**IEEE Globecom**), 2019:1-6. (EI)
3. **Qi Q**, Chen X, Zhong C, Zhang Z. Robust integration of computation and communication in B5G cellular Internet of Things[C] // 2020 IEEE Wireless Communications and Networking Conference (**IEEE WCNC**), 2020:1-6. (EI)
4. **Qi Q**, Chen X, Yuen C. Joint resource allocation for integrated localization and computing in edge-intelligent networks[C] // 2022 IEEE Global Communications Conference (**IEEE Globecom**), 2022:1-6. (EI, 获**IEEE ComSoc Student Grant**)
5. Fu R, **Qi Q**, Zhong C, Chen X, Zhang Z. Block error rate analysis of shortPacket NOMA communications with imperfect SIC[C] // 2020 IEEE/CIC International Conference on Communications in China (**IEEE CIC/ICCC**), 2020:1-6. (EI)
6. Chu J, Chen X, **Qi Q**, Zhong C, Lin H, Zhang Z. On the design of B5G multibeam LEO satellite Internet of Things[C] // 2020 IEEE 91st Vehicular Technology Conference (**IEEE VTC-Spring**), 2020:1-6. (EI)

7. Jia R, Chen X, **Qi Q**. Low-complexity beamspace massive access for B5G cellular Internet of Things[C] // 2019 International Conference on Wireless Communications and Signal Processing (**WCSP**), 2019:1-6. (EI)
8. Wang Q, Chen X, **Qi Q**. Design of delay-optimal robust edge computing in 6G wireless networks[C] // 2022 International Conference on Wireless Communications and Signal Processing (**WCSP**), 2022:1-6. (EI)
9. **Qi Q**, Chen X, Yuen C, Zhang Z. DL-based joint waveform and beamforming design for integrated sensing and communication[C]. Submitted to 2023 IEEE Global Communications Conference (**IEEE CIC/ICCC**), 2023. (EI, 投稿中)
10. Wang Q, Chen X, **Qi Q**. Joint offloading selection, beamforming design and resource allocation for satellite-terrestrial computing[C]. Submitted to 2023 IEEE Global Communications Conference (**IEEE CIC/ICCC**), 2023. (EI, 投稿中)

学术专著:

1. Chen X, **Qi Q**. Convergence of energy, communication and computation in B5G cellular Internet of Things[M]. **Singapore: Spring**, 2020.

发明专利:

1. 齐俏, 陈晓明. 一种基于无线信息与能量协同传输的大规模接入方法[P]. 中国: CN109587088B, 2019. (授权)
2. 齐俏, 陈晓明. 一种计算与通信融合的大规模接入方法[P]. 中国: CN110380762B, 2019. (授权)
3. 齐俏, 陈晓明. 一种面向边缘智能网络的大规模接入方法[P]. 中国: CN112911608B, 2021. (授权)
4. 齐俏, 陈晓明, 张朝阳. 一种无线网络中感知、计算和通信的融合方法[P]. 中国: CN113825236B, 2021. (授权)
5. 齐俏, 陈晓明, 张朝阳. 一种定位、通信和计算的融合方法[P]. 中国: CN114928810A, 2022. (公开)



6. 齐俏, 陈晓明, 张朝阳. 一种基于深度学习的通信和感知融合方法[P]. 中国: CN115551024A, 2022. (公开)
7. 王琦, 陈晓明, 齐俏, 张朝阳. 一种计算和通信的融合方法[P]. 中国: CN114599041A, 2022. (公开)
8. 王琦, 陈晓明, 齐俏, 张朝阳. 一种星地联合的边缘计算方法[P]. 中国: CN115499874A, 2022. (公开)