

# 基于区块链的卫星终端试验任务数据管理与计算卸载技 术研究

姓名： 毛心璐

---

学号： 3220231823

---

学院： 网络空间安全学院

---

2022 年 6 月

④

基于区块链的卫星终端试验任务数据管理与计算卸载技术研究 毛心璐 北京理工大学

中图分类号：TQ028.1

UDC 分类号：540

基于区块链的卫星终端试验任务数据管理与计算卸载技  
术研究

作 者 姓 名	毛心璐
学 院 名 称	网络空间安全学院
指 导 教 师	朱超
行业合作导师	朱超
答辩委员会主席	张涛
申 请 类 别	工学硕士
学 位 领 域	网络与信息安全
学位授予单位	北京理工大学
论文答辩日期	2026 年 6 月

**Data Management and Computing Offloading  
Technology Scheme for Satellite Terminal Test Task  
Based on Blockchain**

Candidate Name:	<u>Xinlu Mao</u>
School or Department:	<u>School of Network Space Security</u>
Faculty Mentor:	<u>Prof. Zhu Chao</u>
Industry Collaboration Mentor:	<u>Prof. Zhu Chao</u>
Chair, Thesis Committee:	<u>Prof. Zhang Tao</u>
Degree Applied:	<u>Master of Engineering</u>
Major:	<u>Network and Information Security</u>
Degree by:	<u>Beijing Institute of Technology</u>
The Date of Defence:	<u>June, 2026</u>

## 研究成果声明

本人郑重声明：所提交的学位论文是我本人在指导教师的指导下独立完成的研究成果。文中所撰写内容符合以下学术规范（请勾选）：

☒ 论文综述遵循“适当引用”的规范，全部引用的内容不超过 50%。

☒ 论文中的研究数据及结果不存在篡改、剽窃、抄袭、伪造等学术不端行为，并愿意承担因学术不端行为所带来的一切后果和法律责任。

☒ 文中依法引用他人的成果，均已做出明确标注或得到许可。

☒ 论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

☒ 与本人一同工作的合作者对此研究工作所做的任何贡献均已在学位论文中作了明确的说明并表示了谢意。

特此声明。

签 名：

日 期：

## 关于学位论文使用权的说明

本人完全了解北京理工大学有关保管、使用学位论文的规定，其中包括：

① 学校有权保管、并向有关部门送交学位论文的原件与复印件；

② 学校可以采用影印、缩印或其它复制手段复制并保存学位论文；

③ 学校可允许学位论文被查阅或借阅；

④ 学校可以学术交流为目的，复制赠送和交换学位论文；

⑤ 学校可以公布学位论文的全部或部分内容（保密学位论文在解密后遵守此规定）。

签 名：

日 期：

导师签名：

日 期：

## 摘要

近年来,随着天地一体化信息网络与“新基建”战略的深入推进,天基网络正经历从单一通信传输向边缘智能计算的跨越式转变。卫星终端作为天基数据的感知与处理前哨,在偏远地区监测、应急响应及各类在轨试验中发挥着不可替代的作用。然而,面对日益复杂的计算密集型试验任务,受限于星载设备的体积、功耗及抗辐射设计,卫星终端的计算资源极其有限。当海量任务并发产生时,传统的计算卸载方案往往忽视了资源竞争带来的“并发开销”,且依赖地面集中式决策导致的高时延与信令开销,已难以满足天基试验任务的实时性与可靠性需求。

针对上述挑战,本文提出了一种融合区块链、联邦学习(Federated Learning, FL)与深度强化学习(DRL)的卫星终端智能管控与计算卸载架构。该架构旨在利用联邦学习打破数据孤岛,在保护隐私的前提下实现协作式环境感知;利用区块链构建可信协作底座,解决分布式环境下的信任与状态同步问题。本文的主要研究内容与创新点如下:

(1) 基于区块链的卫星试验任务可信存证与协同管理:针对多源异构试验数据的安全与一致性问题,设计了基于联盟链的分层式数据管理架构。利用智能合约实现试验任务全生命周期的自动化记录与防篡改存证,并构建非实时全局状态共享机制,为分布式协同计算提供可信的数据底座,确保了任务执行过程的可追溯性与多方协作的安全性。

(2) 基于联邦学习的并发开销预测与卸载决策优化:针对多任务并发场景下的资源竞争难题,提出了一种基于 FL 与深度 Q 网络(DQN)的分布式卸载策略。通过在卫星终端间开展联邦协同训练,构建高精度的任务并发开销预测模型,使终端能够在不共享原始任务数据的前提下感知全局负载压力,从而实现任务卸载决策的智能化与隐私保护,有效解决了资源竞争导致的时延剧增问题。

(3) 基于深度强化学习的动态资源调度与节点优选:针对天基网络拓扑高动态变化及星间链路不稳定的特性,设计了基于改进 PPO/DQN 算法的动态资源调度模型。该模型将时延、能耗、负载均衡度纳入多目标优化函数,通过智能体与环境的持续交互,实现复杂动态约束下的计算资源自适应分配与最优卸载节点选择。

(4) 天基分布式试验智能管控平台的设计与验证:基于上述理论成果,设计并研制了包含卫星终端模拟、区块链网络、智能调度引擎及可视化监控的一体化演示验证

平台。在典型的卫星试验任务场景下进行了系统级测试，实验结果表明，本文提出的方案在降低平均任务时延、提升系统吞吐量及保障数据隐私方面均优于传统基准算法，验证了所提架构的有效性与优越性。

本研究探索了区块链与人工智能技术在天基网络中的深度融合应用，有效提升了受限资源下卫星终端的任务处理效能与数据安全水平，为未来构建高效、智能、可信的天基边缘计算网络提供了重要的理论依据与技术支撑。

**关键词：**区块链；卫星网络；强化学习；计算卸载；联邦学习



## Abstract

In recent years, with the in-depth advancement of space-ground integrated information networks and "New Infrastructure" strategies, space-based networks are undergoing a leapfrog transition from simple communication transmission to edge intelligent computing. As the forward outpost for space-based data perception and processing, satellite terminals play an irreplaceable role in remote area monitoring, emergency response, and various on-orbit experiments. However, faced with increasingly complex computation-intensive test tasks, satellite terminals are severely constrained by the size, weight, and power (SWaP) limitations of onboard equipment. When massive tasks are generated concurrently, traditional computation offloading schemes often overlook the "concurrency overhead" caused by resource competition. Furthermore, reliance on ground-based centralized decision-making leads to high latency and signaling overhead, making it difficult to meet the real-time and reliability requirements of space-based test tasks.

To address these challenges, this thesis proposes an intelligent management and computation offloading architecture for satellite terminals, fusing Blockchain, Federated Learning (FL), and Deep Reinforcement Learning (DRL). This architecture aims to utilize FL to break data silos and achieve collaborative environmental perception while preserving privacy; simultaneously, it leverages blockchain to build a trusted collaboration foundation, resolving trust and state synchronization issues in distributed environments. The main research contents and innovations are as follows:

(1) Trusted Storage and Collaborative Management of Satellite Test Tasks Based on Blockchain: Addressing the security and consistency issues of multi-source heterogeneous test data, a hierarchical data management architecture based on consortium blockchain is designed. Smart contracts are utilized to realize automated recording and tamper-proof storage of the entire lifecycle of test tasks. Additionally, a non-real-time global state sharing mechanism is constructed to provide a trusted data foundation for distributed collaborative computing, ensuring the traceability of the execution process and the security of multi-party collaboration.

(2) Concurrency Overhead Prediction and Offloading Decision Optimization Based on Federated Learning: Targeting the resource competition problem in multi-task concurrency

scenarios, a distributed offloading strategy based on FL and Deep Q-Network (DQN) is proposed. By conducting federated collaborative training among satellite terminals, a high-precision prediction model for task concurrency overhead is constructed. This enables terminals to perceive global load pressure without sharing raw task data, thereby achieving intelligent and privacy-preserving offloading decisions and effectively solving the latency surge caused by resource competition.

(3) Dynamic Resource Scheduling and Node Selection Based on Deep Reinforcement Learning: Addressing the highly dynamic topology of space-based networks and the instability of inter-satellite links, a dynamic resource scheduling model based on improved PPO/DQN algorithms is designed. This model incorporates latency, energy consumption, and load balance into a multi-objective optimization function. Through continuous interaction between the agent and the environment, adaptive allocation of computing resources and optimal offloading node selection under complex dynamic constraints are achieved.

(4) Design and Verification of Space-based Distributed Test Intelligent Management Platform: Based on the above theoretical results, an integrated demonstration and verification platform containing satellite terminal simulation, blockchain network, intelligent scheduling engine, and visual monitoring is designed and developed. System-level tests under typical satellite test task scenarios demonstrate that the proposed scheme outperforms traditional benchmark algorithms in reducing average task latency, improving system throughput, and guaranteeing data privacy, verifying the effectiveness and superiority of the proposed architecture.

This research explores the deep integration and application of blockchain and artificial intelligence technologies in space-based networks, effectively improving the task processing efficiency and data security level of satellite terminals under resource-constrained conditions, providing important theoretical basis and technical support for building efficient, intelligent, and trusted space-based edge computing networks in the future.

**Key Words:** Blockchain; Satellite Network; Reinforcement Learning; Computing Offloading; Federated Learning

## 目录

第 1 章	绪论	1
1.1	研究目的与意义	1
1.2	国内外研究现状与发展趋势	3
1.2.1	天基网络架构与卫星边缘计算	3
1.2.2	区块链赋能的分布式可信协同与数据管理	4
1.2.3	联邦学习与隐私保护的协作式建模	5
1.2.4	深度强化学习驱动的动态任务卸载与资源优化	6
1.3	本论文研究内容与主要贡献	7
1.4	论文结构安排	9
第 2 章	相关理论与关键技术	11
2.1	天基网络与卫星终端试验任务的系统架构与挑战	11
2.1.1	系统架构	11
2.1.2	关键挑战的形式化描述	12
2.2	基于区块链的分布式可信数据管理与协作	13
2.2.1	分层联盟链架构	13
2.2.2	核心机制：智能合约与非实时状态共享	15
2.3	联邦学习的并发开销预测与隐私协作建模	16
2.3.1	并发开销的数据驱动建模	16
2.3.2	基于联邦学习的协作式预测模型训练	17
2.3.3	关键技术考量	18
2.4	深度强化学习的动态资源调度与卸载节点优选	19
2.4.1	计算卸载的马尔可夫决策过程建模	19
2.4.2	基于深度 Q 网络（DQN）的求解算法	21
2.5	天基网络计算卸载的形式化建模	21
2.5.1	任务模型	22
2.5.2	决策变量	22
2.5.3	时延模型（Latency Model）	22

2.5.4	能耗模型 (Energy Model)	23
2.5.5	优化目标	24
2.6	本章小结	24
第 3 章	基于区块链的试验任务可信管理架构	26
3.1	本章概述	26
3.2	总体架构设计	26
3.2.1	各层功能定义	27
3.2.2	数据流与控制流	28
3.3	分层联盟链可信数据底座	30
3.3.1	主链-子链职责边界	30
3.3.2	区块结构定义与智能合约流程	31
3.4	基于账本的非实时状态共享机制	32
3.4.1	机制原理	33
3.4.2	一致性与时效性的折中原则	33
3.4.3	对 DQN 决策的支撑	34
3.5	本章小结	34
第 4 章	基于联邦学习与 DQN 的分布式卸载策略	36
4.1	并发开销问题建模	36
4.1.1	并发开销的数学定义	36
4.1.2	并发开销的近似可分表达	37
4.2	基于联邦学习的并发开销预测	39
4.2.1	并发开销预测模型	39
4.2.2	聚类式联邦学习训练流程	40
4.3	基于 DQN 的智能卸载决策	42
4.3.1	DQN 模型构建	42
4.3.2	训练与决策机制	43
4.4	实验与仿真	44
4.4.1	实验设置	45
4.4.2	性能评估与分析	47

4.4.3 综合性能对比 .....	53
4.5 本章小结 .....	53
结论 .....	56
参考文献 .....	57
附录 A 费马大定理的证明 .....	58
附录 B Maxwell Equations .....	59
附录 C 要求 .....	60
C.1 一级标题 .....	60
C.1.1 二级标题 .....	60
攻读学位期间发表论文与研究成果清单 .....	61
致谢 .....	62
作者简介 .....	63

## 插图

图 4.1	并发开销预测模型结构 .....	40
图 4.2	聚类式联邦学习训练流程 .....	42
图 4.3	FL-DQN 卸载决策算法流程 .....	45
图 4.4	不同并发度下的平均新任务时延 .....	48
图 4.5	不同并发度下的系统吞吐量 .....	48
图 4.6	任务完成时延的累积分布函数 .....	49
图 4.7	任务完成时延的箱线图 .....	50
图 4.8	平均单任务能耗对比 .....	50
图 4.9	任务能耗的累积分布函数 .....	51
图 4.10	并发开销预测值与真实值对比 .....	51
图 4.11	联邦学习训练损失收敛曲线 .....	52
图 4.12	聚类联邦学习与传统 FedAvg 的预测误差对比 .....	52
图 4.13	DQN 训练过程中的平均回报曲线 .....	53
图 4.14	奖励函数权重系数的敏感性分析 .....	54
图 4.15	不同共识机制下的链上确认时延 .....	54

## 表格

表 4.1 各算法综合性能对比 .....	55
-----------------------	----

## 主要符号对照表

BIT	北京理工大学的英文缩写
$\text{\LaTeX}$	一个很棒的排版系统
$\text{\LaTeX} 2_{\epsilon}$	一个很棒的排版系统的最新稳定版
ctex	成套的中文 $\text{\LaTeX}$ 解决方案，由一帮天才们开发
$e^{\pi i} + 1 = 0$	一个集自然界五大常数一体的炫酷方程



# 第 1 章 绪论

## 1.1 研究目的与意义

随着“网络强国”和“航天强国”战略的深入实施以及全球数字化转型的加速推进，天地一体化信息网络（Space-Air-Ground Integrated Network, SAGIN）作为国家关键信息基础设施，其战略地位日益凸显。天基网络，特别是以低轨道（Low Earth Orbit, LEO）卫星星座为骨干的系统，正由传统的通信中继与数据传输管道，演进为集通信、计算、感知于一体的“天基云网”。这一演进使得卫星网络不仅能够为海洋、沙漠、山区等地面蜂窝网络难以覆盖的区域提供无缝宽带接入服务，还开始承载日益复杂的在轨实时处理任务，如对地观测数据智能分析、空间环境监测、应急通信保障以及各类在轨科学试验等。

在此背景下，卫星终端作为天基网络的“神经末梢”，承担着连接物理世界与数字空间的前哨角色。在各类在轨试验与探测任务中，卫星终端负责海量、多源、异构数据的感知、采集与初步处理。然而，随着任务复杂度不断提升，试验业务呈现出典型的计算密集型特征。例如，高分辨率遥感图像在轨识别、空间态势感知实时分析、在轨航天器故障快速诊断等任务，对终端的即时计算能力提出了极高要求。同时，卫星终端作为特殊嵌入式设备，其设计受到体积（Size）、重量（Weight）、功耗（Power）与成本（Cost）的“SWaP-C”约束；星载处理器还需经过抗辐射加固设计，这进一步限制了其计算性能的提升。因此，卫星终端普遍面临计算、存储与能量资源极其有限的现实困境。

当大量计算密集型试验任务并发产生时，仅依靠单个卫星终端的本地计算能力已难以为继。为此，计算卸载（Computation Offloading）技术应运而生，成为提升卫星终端任务处理能力的关键使能技术。其核心思想是将终端难以高效处理的计算任务，通过星间链路或星地链路迁移至计算资源更为充裕的节点（如协同卫星、地面站或云数据中心）执行，再将结果返回任务发起终端。

早期计算卸载方案大多采用“星-地”模式，将海量原始数据回传至地面数据中心集中处理，形成典型的“空采地算”范式。该模式虽然能够充分利用地面强大的计算能力，但也暴露出突出问题：一是星地链路带宽有限且可见时间窗口短暂，海量数据回传成为系统性能瓶颈；二是星地往返路径长，往返时延（Round-Trip Time, RTT）

难以满足应急响应、实时控制等时延敏感型任务需求；三是高度集中的决策模式使地面中心站成为性能与可靠性瓶颈，一旦故障可能导致大范围服务中断。

为克服上述问题，卫星边缘计算（Satellite Edge Computing, SEC）或天基边缘计算（Space-based Edge Computing）作为新兴计算范式受到广泛关注。其核心理念是将计算与存储能力向网络边缘（即卫星终端或星座边缘节点）下沉，构建分布式计算平台。在该架构下，资源受限的卫星终端可将任务卸载至邻近、计算能力更强的卫星边缘节点，实现“数据在轨处理、结果回传地面”的高效模式。这不仅大幅减轻星地链路负担、显著降低任务响应时延，还通过分布式协同增强天基网络整体的韧性与可扩展性。

然而，将地面边缘计算技术直接迁移至天基网络仍面临一系列独特挑战：

（1）分布式环境下的信任缺失与协同困境：天基网络由分属不同运营方、覆盖不同区域的卫星节点构成，是典型的多主体分布式系统。在缺乏中心化可信第三方的情况下，如何保障试验任务从规划、分发、执行到结果回收等全流程的公开透明与不可篡改，成为多方可信协作的首要难题。试验数据每一次流转与处理，都需要可靠的存证与溯源机制，以防止数据在传输和计算过程中被恶意篡改或伪造，确保最终试验结果的公信力。

（2）高并发任务下的并发开销与资源竞争：在真实试验场景中，多个卫星终端可能同时向同一边缘节点卸载任务。新任务的到来会与正在执行的任务争夺 CPU 周期、内存等有限资源，使所有任务的完成时间普遍延长。这种由资源竞争引发的额外性能损耗即“并发开销”，却常被传统卸载策略忽略。如何准确建模与量化并发开销，使卸载决策既考虑新任务自身时延，又兼顾其对系统中存量任务性能的影响，是实现全局最优的关键。

（3）网络拓扑高动态与资源状态不确定性：尤其在 LEO 星座中，卫星高速运动、星间链路频繁切换，网络拓扑结构时刻变化，导致节点难以实时获取精确的全局状态信息（如各节点负载、链路质量等）。依赖“全局信息透明”假设的集中式优化算法在此环境下不仅信令开销巨大，且信息滞后严重，难以适应快速变化的网络。亟需在信息不完全、时变动态的分布式环境下，构建高效、鲁棒的动态卸载决策与资源调度智能方法。

针对上述挑战，本文拟探索将区块链（Blockchain）、联邦学习（Federated Learning, FL）与深度强化学习（Deep Reinforcement Learning, DRL）等前沿技术深度融合，构

建一套面向卫星终端试验任务的分布式、智能化数据管理与计算卸载框架。具体而言,利用区块链的去中心化、不可篡改与可追溯特性,为多主体协作构建可信基础,实现试验任务全生命周期的可信存证与协同管理;在此基础上,引入联邦学习机制,在不泄露各终端原始数据的前提下,协同构建高精度的并发开销预测模型,使终端能够更精准地感知全局负载状态;最终,借助深度强化学习,使各终端智能体在与高动态环境的持续交互中,自主学习最优计算卸载与资源调度策略,实现时延、能耗与负载均衡等多目标的动态优化。

因此,本文的核心目的与意义在于:面向天基网络中卫星终端资源受限、任务并发和网络高动态等现实瓶颈,系统性解决分布式环境下试验任务数据可信管理难、计算资源协同调度效率低的关键问题。通过构建一个可信、智能、高效的计算卸载架构,不仅能够显著提升卫星终端任务处理效能与系统吞吐量,保障试验数据安全与隐私,还可为我国构建自主可控、智能泛在的天地一体化信息网络提供新的理论方法与关键技术支撑,具有重要的理论研究价值和广阔的工程应用前景。

## 1.2 国内外研究现状与发展趋势

本节围绕“基于区块链的卫星终端试验任务数据管理与计算卸载”这一核心议题,从天基网络计算范式演进、区块链赋能的可信协同、联邦学习驱动的协作建模以及深度强化学习引导的动态资源优化四个维度,对国内外前沿研究现状进行系统综述,梳理核心进展,分析现有技术局限,并据此明确本文的研究切入点与创新价值。

### 1.2.1 天基网络架构与卫星边缘计算

随着“新基建”和全球数字化进程的深入推进,以 LEO 卫星星座为核心的天地一体化信息网络(SAGIN)正由传统通信管道向集通信、计算、感知于一体的“天基云网”演进。为缓解海量数据回传瓶颈、满足时延敏感任务需求,卫星边缘计算(Satellite Edge Computing, SEC)已成为重要研究方向,其核心思想是将计算能力由地面中心前移至轨道边缘,实现数据的在轨智能处理。

近年来研究主要聚焦于 SEC 的架构设计、资源管理与应用赋能。在顶层架构方面,大量综述性工作为该领域描绘了整体蓝图。Cheng 等在文献<sup>Cheng2022</sup>中系统阐述了面向 6G 的 SAGIN 服务化网络架构,强调异构资源协同与云边融合的重要性;Zhou 等在<sup>Zhou2023</sup>中则从航空航天一体化网络视角,深入探讨了网络动态性建模、理论性能

分析与系统优化等关键问题，并指出超密集卫星星座带来的新挑战。这些工作共同表明：将云计算与边缘计算能力从地面延伸至太空，构建分布式、智能化天基计算体系是必然趋势。

在关键技术挑战方面，Wang 和 Li 在<sup>Wang2023</sup>中明确指出卫星计算面临资源受限、能量约束与拓扑高动态等核心问题，并以“天算星座”为例展示了开放卫星研究平台的实践探索。这些挑战推动了面向 SEC 场景的计算卸载与资源分配算法研究。早期工作多采用集中式优化或传统博弈论方法，但普遍难以适应天基网络的强动态性和不确定性。

总体来看，现有研究的主要局限体现在：

(1) 对网络高度动态性的适应不足：多数研究基于简化网络模型，未充分考虑卫星高速移动与星间链路频繁切换对卸载连续性和策略稳定性的冲击。

(2) 对“并发开销”的忽视：传统模型往往假设边缘节点资源可被线性划分，忽略多任务并发执行时因缓存、内存带宽等资源竞争导致的非线性性能下降，使时延和能耗估计与实际偏差较大。

(3) 依赖理想化全局信息：许多算法假设能够实时获取精确的全局网络状态，这在分布式、大时延的天基网络中不仅信令开销巨大，而且难以实现。

因此，面向真实天基环境，设计能够在不完全信息下适应高动态拓扑、并精准刻画并发性能的分布式智能卸载与资源管理方案，已成为亟待突破的关键问题，也构成了本文的研究起点。

### 1.2.2 区块链赋能的分布式可信协同与数据管理

在多运营主体共存的分布式天基网络中，构建互信机制并保障数据全生命周期的完整性与可追溯性，是实现高效协同的前提。区块链技术以去中心化、防篡改与可追溯等特性，为解决该问题提供了新的技术路径。

当前，将区块链与边缘计算相结合的研究已成为热点，尤其在为天基网络具有类似分布式特征的车联网（Vehicular Networks）中。Nguyen 等在<sup>Nguyen2023</sup>中提出了一种区块链支持的边缘计算协作卸载与区块挖掘方案，采用多智能体 DRL 联合优化卸载决策与资源分配；Fan 等在<sup>Fan2024</sup>中面向云-边-端协同网络，研究了区块链保障下的任务卸载与资源分配，通过 Lyapunov 优化与 DRL 结合，最小化任务时延与能耗；Wang 等在<sup>Wang2023b</sup>中则探讨了区块链支撑的数字孪生物联网可持续计算资源管理问题。

这些研究表明，区块链在分布式边缘环境中保障任务安全、构建激励机制与促进资源共享方面具有巨大潜力，核心思想可迁移至天基网络场景，主要体现在：

(1) 任务全生命周期可信存证：将卫星试验任务的发布参数、执行节点、起止时间、结果哈希等关键信息作为交易写入区块链，形成不可篡改的“分布式任务日志”，支撑任务审计、故障排查与多方协作。

(2) 非实时全局状态共享：通过读取链上数据，授权卫星终端可获知其他节点的任务执行概况，在无需实时点对点通信的情况下形成准实时、一致性全局状态视图，为分布式决策提供支撑。

(3) 基于智能合约的自动化协同治理：通过智能合约定义任务分配规则、资源访问控制与奖惩机制，实现去中心化、自动化的资源调度与治理。

然而，现有工作大多面向地面边缘网络，直接应用于卫星场景仍面临两方面挑战：其一，传统共识机制（如 PoW）计算与能耗开销巨大，难以适应星上资源极度受限的环境；其二，将全部数据直接上链既不现实也不高效，亟需设计高效的链上链下协同机制。

本文工作与上述研究密切相关，但更加聚焦天基网络的特殊约束：拟采用轻量级联盟链共识机制（如 PoA 及其变体），并设计分层链架构以解耦控制信令与大规模试验数据，实现非实时全局状态共享与可信协同的区块链底座。这构成了本文的第一项核心贡献。

### 1.2.3 联邦学习与隐私保护的协作式建模

在分布式智能卸载决策中，单个终端如何准确感知与预测全网状态，尤其是其他节点的资源负载，是一项关键难题。联邦学习（Federated Learning, FL）作为“数据不动、模型移动”的分布式学习范式，为此提供了理想的隐私保护框架。

FL 在边缘网络中的研究发展迅速。Zhai 等在<sup>Zhai2024</sup>中提出了专为 LEO 卫星网络设计的去中心化联邦学习框架 FedLEO，通过计算卸载辅助模型训练，有效应对卫星间歇性连接和资源异构性问题。面向车载网络场景，Zhao 等在<sup>Zhao2024</sup>中将 FL 与多智能体 DRL 结合，在保护数据隐私的同时，通过协作训练提升任务卸载与资源分配策略的性能，并缓解数据非独立同分布（Non-IID）问题。

这些研究启发我们，将 FL 应用于解决计算卸载中长期被忽视的“并发开销”预测问题。并发开销的准确建模是 DRL 智能体能否做出近似最优卸载决策的前提。然

而，单个卫星节点历史运行数据有限，难以训练出具有良好泛化能力的模型；同时，直接共享包含任务细节与性能指标的原始数据又带来严重隐私风险。

现有 FL 相关研究多聚焦于 DRL 策略本身的联邦训练，或一般资源预测问题，鲜有工作针对并发执行引起的非线性性能影响进行专门建模。考虑到并发开销受硬件微结构、操作系统调度及任务类型组合等多因素影响，难以用简单数学公式刻画，典型适合采用数据驱动方法。

为弥补这一空白，本文提出利用联邦学习协同训练并发开销预测模型。具体而言，每个卫星边缘节点基于本地历史任务运行数据，训练一个深度神经网络（如卷积神经网络，CNN）以拟合其在不同负载与任务组合下的性能表现；随后通过联邦学习（如 FedAvg）对各节点本地模型参数进行安全聚合，得到精度更高、泛化能力更强的全局并发开销预测模型。终端在进行卸载决策时，可基于该模型输入任务信息和目标节点排队情况，获得更精准的任务完成时延预测，从而有效提升决策质量。

这构成本文的第二项核心贡献：通过隐私保护的协作建模，解决智能卸载中关键环境参数（并发开销）的感知难题，避免盲目卸载引发的“雪崩效应”，为资源竞争场景下的性能预测提供新思路。

#### 1.2.4 深度强化学习驱动的动态任务卸载与资源优化

在拓扑高动态、链路不稳定及资源状态时变的天基网络中，传统依赖精确模型的优化方法往往难以奏效。深度强化学习（DRL）作为一种无模型决策优化技术，通过智能体与环境的持续交互与试错学习，可在复杂约束下自主寻优，已成为解决动态卸载与资源分配问题的重要工具。

DRL 在边缘计算任务卸载中的应用已较为广泛。Yamansavascular 等在<sup>Yamansavascular2023</sup>中提出 DeepEdge 框架，利用双重深度 Q 网络（DDQN）处理动态卸载请求与时变信道，以最大化任务完成率；Xie 和 Cui 在<sup>Xie2025</sup>中则针对排队时延导致奖励信息失真问题，提出改进的 PPO 算法以提升策略性能与收敛稳定性。

在卫星网络领域，Zhang 等在<sup>Zhang2024</sup>中围绕能效优化，设计多跳对等卸载方案，并结合 Lyapunov 框架与在线学习提升能效与时延表现；Zhang 等在<sup>Zhang2023</sup>中则采用深度确定性策略梯度（DDPG）算法，实现对连续与离散混合动作空间（卸载决策与资源分配比例）的协同优化。进一步地，Zhao 等在<sup>Zhao2025</sup>中提出基于时空注意力机制的 PPO 算法（STA-PPO），分别捕捉任务到达的时间依赖性与卫星运动的空间动态，以

应对 LEO 网络独特的时空特性；Chen 等在<sup>Chen2024</sup>中提出 SpaceEdge 方案，通过 DRL 优化服务迁移与功率控制，提升服务连续性与可持续性；Han 等在<sup>Han2023</sup>中则采用两时间尺度学习框架，解耦长期资源分配与短期卸载决策，有效降低决策复杂度。

尽管上述工作在动态决策优化方面取得显著进展，但仍存在共性不足：

(1) 环境状态感知不完整：大多数 DRL 智能体的状态构造依赖理想化局部信息，缺乏对全网负载与任务竞争情况的可靠、一致视图，易导致策略次优。

(2) 决策成本预估不准确：对并发开销的简化建模甚至忽略，使智能体在评估卸载动作价值（Q 值）时所依赖的成本函数与真实环境存在偏差，不利于策略收敛与性能发挥。

(3) 多目标优化权衡不足：通常通过固定加权和构造奖励函数，难以灵活适配不同业务场景下对时延、能耗与负载均衡等指标优先级的动态变化。

本文工作针对上述不足进行了有针对性的改进：首先，依托区块链实现的非实时全局状态共享机制，为 DRL 智能体提供更全面、可信的状态输入；其次，通过联邦学习训练的并发开销预测模型，为奖励函数提供更准确的成本评估。在此基础上，将计算卸载问题建模为马尔可夫决策过程（MDP），并采用改进的 DQN 或 PPO 算法进行求解，使智能体在更接近真实的环境中学习，获得兼顾时延、能耗与负载均衡的鲁棒高效策略。这构成了本文的第三项核心贡献，形成“区块链提供状态—FL 预测成本—DRL 优化决策”的技术闭环。

### 1.3 本论文研究内容与主要贡献

面向天基网络环境下卫星终端资源受限、试验任务并发、协同信任缺失以及网络高动态等核心挑战，本文以构建高效、智能、可信的卫星终端试验任务数据管理与计算卸载系统为总体目标，提出一种深度融合区块链、联邦学习与深度强化学习的创新架构。本文主要研究内容与贡献概括如下。

(1) 基于区块链的卫星试验任务可信存证与协同管理。

为解决分布式多主体天基网络中的信任缺失问题，保障试验任务全生命周期的安全、透明与一致性，本文首先设计了一种基于联盟链的分布式数据管理与协同框架。构建面向天基试验任务的分层式联盟链架构，包括负责全局任务协调与共识的主链，以及负责具体试验数据存证的多个子链，并通过跨链机制实现主子链间的信息交互，以分离控制信令与海量试验数据，提升系统吞吐量与可扩展性。利用智能合约对试验

任务的发布、接收、执行、结果验证等关键流程进行编码与自动化,实现任务规则、参与方权责及奖惩机制的去中心化执行与不可篡改。进一步,设计非实时全局状态共享机制,将任务卸载事件(如任务开始和结束)以交易形式记录在链上,使任一终端均可通过链上查询获取系统中任务执行状态,为后续智能卸载决策提供可信、一致的全局视图。

#### (2) 基于联邦学习的并发开销预测与卸载决策优化。

针对传统卸载策略普遍忽略多任务并发执行导致的“并发开销”问题,本文创新性引入联邦学习,在不泄露各节点原始运行数据前提下,协同构建高精度并发开销预测模型,以优化卸载决策。首先分析并形式化定义任务并发开销,将其刻画为新任务抵达后因资源竞争对自身及在执行任务造成的额外时延,并建模为复杂非线性函数预测问题。进而提出基于联邦学习与深度 Q 网络(FL-DQN)的分布式卸载策略:采用卷积神经网络(CNN)拟合并发开销函数,各卫星边缘节点利用本地历史任务执行日志进行本地训练,再通过 FL 机制(如 FedAvg)在不交换原始数据的情况下聚合模型参数,得到精度更高、泛化更好的全局并发开销预测模型。该模型分发至所有终端,用于卸载决策时精准预估任务真实完成时延。

#### (3) 基于深度强化学习的动态资源调度与节点优选。

为适应拓扑高动态、链路不稳定与资源状态不确定的天基网络环境,本文采用深度强化学习方法设计动态资源调度与最优卸载目标选择策略。将单个卫星终端的计算卸载过程建模为马尔可夫决策过程(MDP),其中状态空间包括终端任务信息、通过区块链获取的邻近节点负载信息以及链路质量等;动作空间为决策选项(本地执行或卸载至某可用邻近卫星节点);奖励函数综合考虑任务完成时延、终端能耗和网络负载均衡度等指标。在此基础上,采用改进的 PPO 或 DQN 算法训练 DRL 智能体,通过与模拟环境的持续交互,在无精确系统模型的情况下逐步学习最大化长期累积奖励的近似最优策略,实现面向新到达任务的动态执行方式与卸载目标选择。

#### (4) 天基分布式试验智能管控平台的设计与验证。

为验证所提理论方法的有效性与工程可行性,本文基于前述架构与算法,设计并实现了一个系统级演示与验证平台。平台集成四大核心模块:卫星终端与网络环境模拟器,用于生成多类型试验任务、模拟卫星拓扑动态变化与星间链路特性;分布式区块链网络,实现分层联盟链与智能合约部署,负责任务可信管理与状态共享;智能调度引擎,内嵌基于 FL 训练的并发开销预测模型与基于 DRL 的动态卸载与资源调度算



法；可视化监控与分析系统，用于实时展示系统运行状态、任务流转过程及各类性能指标。通过与本地执行、随机卸载、贪心卸载等基准算法的对比实验，验证本方案在降低任务平均完成时延、提升系统吞吐量、保障数据安全与隐私及实现网络负载均衡等方面的优势。

## 1.4 论文结构安排

本文共分为六章，主要安排如下：

第一章为绪论，介绍研究背景与意义，分析天地一体化信息网络发展趋势及卫星终端在试验任务场景下面临的资源受限与计算卸载挑战，综述天基网络计算卸载、区块链数据管理、联邦学习与深度强化学习等相关领域研究现状，明确本文研究内容、创新点与研究目标，并给出全文结构安排。

第二章为相关理论与关键技术，系统介绍本文所涉及的基础理论与关键技术，包括计算卸载的基本概念与模型；区块链技术原理，重点为联盟链、智能合约与共识机制；联邦学习框架与工作流程及其在隐私保护方面的优势；深度强化学习的基本要素（如 MDP），以及 DQN、PPO 等典型算法。

第三章为基于区块链的试验任务可信管理架构，详细阐述本文提出的任务管理与数据存证方案。设计面向天基试验场景的分层联盟链架构，定义主链与子链的功能定位与交互方式；设计基于智能合约的任务全生命周期管理流程，并给出区块结构与数据格式；说明如何利用该区块链架构实现分布式环境下的任务状态共享与协同。

第四章为基于联邦学习与深度强化学习的计算卸载策略，是全文核心章节。重点介绍智能计算卸载算法：构建系统模型，对任务、通信、计算及“并发开销”进行数学形式化定义；详细说明基于联邦学习的并发开销预测模型设计与训练过程；将计算卸载问题建模为 MDP，设计基于 DRL 的动态卸载与资源调度算法；给出 FL-DRL 协同工作的完整流程。

第五章为仿真实验与性能评估，对所提方案进行系统级评估。介绍仿真平台参数设置、数据集构建与对比基准算法选取；通过一系列实验，从任务平均完成时延、系统吞吐量、终端能耗、网络负载均衡度等多个维度评估所提方案性能，并对实验结果进行深入分析与讨论。

第六章为总结与展望，对全文研究工作进行总结，重申主要贡献与创新点，分析当前研究的不足，并展望未来可进一步研究的方向，如更精细的移动性管理、更鲁棒

的异构联邦学习算法以及多智能体协同策略的演进等。

## 第 2 章 相关理论与关键技术

### 2.1 天基网络与卫星终端试验任务的系统架构与挑战

为了有效应对天基网络中计算卸载的复杂性，首先需要构建一个清晰、准确的系统模型，用以描述网络中的关键实体、它们的交互关系以及所面临的独特环境与挑战。本节将详细定义系统架构，并对其中的关键挑战进行形式化分析。

#### 2.1.1 系统架构

本文所研究的天基网络计算卸载系统架构如图 2.1 所示，是一个典型的三层异构网络，由天基网络层、空口通信层与地面网络层构成，研究的核心则聚焦于天基网络内部的分布式协同。

说明：此处应配有一张系统架构图（例如，包含卫星终端、LEO 边缘卫星、GEO 骨干卫星和地面站的示意图），以直观展示多层结构和实体间的关系。

**(1) 卫星终端层 (Satellite Terminal Layer)** 该层由大量执行各类试验任务的卫星终端 (Satellite Terminal, ST) 构成。这些终端可以是搭载特定传感器或载荷的科学实验卫星、对地观测卫星，或其他需要在轨进行数据处理的航天器。在本模型中，ST 是计算任务的发起者。每个 ST 具备一定的本地计算能力，但受制于 SWaP-C (Size, Weight, Power, Cost) 约束，其计算、存储和能量资源通常有限。当面临计算密集型或时延敏感型任务时，ST 会寻求将任务卸载至更高层级的计算节点。

**(2) 天基边缘计算层 (Space-based Edge Computing Layer)** 该层主要由一组低地球轨道 (LEO) 卫星构成，它们在承担通信中继功能的同时，配备了增强的计算与存储单元，作为天基边缘服务器 (Satellite Edge Server, SES)。这些 SES 节点在轨道上形成一个动态的分布式计算网络，能够为邻近 ST 提供低时延的计算卸载服务，是计算任务的主要执行者。由于 LEO 卫星高速运动，SES 网络的拓扑结构高度动态，星间链路 (Inter-Satellite Links, ISL) 会频繁建立与断开。

**(3) 天基核心网络/骨干层 (Space-based Core Network / Backbone Layer)** 该层可由地球静止轨道 (GEO) 卫星或中地球轨道 (MEO) 卫星构成，它们拥有更强的处理

能力和更广的覆盖范围，作为区域性的数据汇聚与交换中心，或充当与地面网络的网关。在某些场景下，这一层的节点也可以作为计算卸载的备选执行节点。

在本研究中，我们主要关注卫星终端层与天基边缘计算层之间的交互关系。一个 ST 可以将任务卸载至其通信范围内任意一个 SES，系统的核心目标是构建一个去中心化的资源管理与任务调度机制，使得每个 ST 能够智能地决策：是将任务在本地执行，还是将其卸载至哪个最合适的 SES 上执行。

### 2.1.2 关键挑战的形式化描述

基于上述系统架构，天基网络计算卸载面临的核心挑战可形式化描述如下。

**(1) 网络拓扑的高动态性 (High Dynamics of Network Topology)** LEO 卫星以约 7.8 km/s 的高速度绕地运行，导致 ST 与 SES 之间以及 SES 与 SES 之间的可见时间窗口短暂且不断变化，从而使任意节点  $i$  在时刻  $t$  的邻居集合

$$N_i(t)$$

呈显著的时变性。这种高动态性对计算卸载策略的连续性和稳定性提出了严峻考验。例如，一个正在向某个 SES 卸载数据的 ST，可能因该 SES 飞出通信范围而导致卸载中断。已有研究如 Zhao 等人文献 [12] 中提出的时空注意力 DRL 方法，正是为了刻画与利用这种时空动态特性。

**(2) 资源状态信息的不完全性 (Incomplete Information of Resource Status)** 在高动态、大时延的天基网络中，任何单个节点都难以实时、精确地获取全局网络状态。对于一个需要做出卸载决策的 ST，它能够观测到的其他 SES 节点的计算负载、任务队列长度等信息，本质上是延迟的、不完整的，即存在明显的部分可观测性 (Partial Observability) 问题。此时，传统基于“全局信息透明”假设的集中式优化算法难以奏效，而在信息不对称条件下进行决策的分布式方法（如文献 [17] 的相关工作）更贴近真实场景，这也驱动我们必须寻求仅依赖局部或异步共享信息的分布式决策范式。

**(3) 任务并发执行的性能影响 (Performance Impact of Concurrent Task Execution)** 当多个 ST 同时向同一 SES 卸载任务时，这些任务将在 SES 上并发执行，共同竞争 CPU、内存、I/O 等计算资源。与理想化的线性资源划分模型不同，实际并发执行会引

入额外性能开销，如缓存争用（Cache Contention）、上下文切换（Context Switching）等，导致每个任务的实际完成时间均高于其独占运行的情形。这种现象被称为“并发开销”（Concurrency Overhead）[14]。若忽略并发开销，将严重低估任务完成时延，进而做出次优甚至错误的卸载决策。准确建模与预测并发开销，是实现精细化资源管理与高效卸载的关键前提。

**（4）分布式协同的信任缺失（Lack of Trust in Distributed Collaboration）** 天基网络天然是由不同所有者、不同管理域的节点构成的分布式系统。在缺乏中心化可信第三方的情况下，如何确保任务数据的完整性、计算结果的正确性以及资源计费的公平性，是一个棘手的问题。例如，一个恶意 SES 可能返回伪造的计算结果，而恶意 ST 可能谎报任务资源需求。因而需要一种技术机制来建立节点间的信任，保障数据与计算过程的可追溯、不可篡改、可审计，这也是文献 [2, 4, 5] 中区块链应用被广泛讨论的核心动机之一。

综上，天基网络计算卸载的核心科学问题可归结为：在资源受限、拓扑高动态、信息不完全且信任缺失的分布式环境下，如何设计智能、高效且可信的计算卸载与资源调度策略，以实现系统整体效能（如时延、能耗、吞吐量等）的最优化。

## 2.2 基于区块链的分布式可信数据管理与协作

为应对前述分布式协同中的信任缺失挑战，本文引入区块链技术，构建一个贯穿任务全生命周期的可信数据管理与协作框架。区块链的去中心化、不可篡改、公开透明和可追溯特性，使其成为解决分布式系统信任问题的理想技术方案 [2]。本节将详细阐述面向天基网络环境定制的区块链系统设计，包括其分层架构、核心机制及工作流程。

### 2.2.1 分层联盟链架构

考虑到天基网络资源受限的特点，尤其是星载计算与存储资源的珍贵，传统公有链（如比特币）采用的工作量证明（Proof of Work, PoW）共识机制因其巨大的算力与能耗开销而完全不适用。同时，若将所有试验任务的原始数据或过程数据全部上链，将迅速耗尽星上存储并造成巨大的通信负担。

为此，本文设计了一种面向天基试验任务的分层联盟链（Hierarchical Consortium

Blockchain) 架构, 并结合链上链下 (On-chain / Off-chain) 协同存储机制。该架构的设计思想借鉴了文献 [5] 的分层思路以及作者前期在车联网领域的研究 [20], 旨在实现控制信令与海量数据的分离, 在确保关键信息可信、可追溯的同时, 最大化系统运行效率与可扩展性。

**(1) 联盟链 (Consortium Blockchain)** 本文选用联盟链作为底层技术形态。在联盟链中, 节点的加入需要预先授权, 只有获得许可的卫星终端 (ST)、天基边缘服务器 (SES) 和地面站 (Ground Station, GS) 才能成为链的参与者。这既保证了系统的安全可控, 又便于采用轻量级共识机制, 如权威证明 (Proof of Authority, PoA)、实用拜占庭容错 (PBFT) 等, 其计算开销远低于 PoW, 更适用于资源受限的卫星节点 [4]。

**(2) 分层架构 (Hierarchical Architecture)** 如图 2.2 所示, 联盟链架构分为主链与子链两个层次:

- **主链 (Main Chain):** 主链是系统的“任务控制链”或“全局状态总账”, 其核心功能是记录试验任务相关的关键元数据与控制信令。一笔典型“任务交易”可能包含: 任务 ID、发起方 ST 的公钥、任务类型、计算量需求、数据大小、目标 SES 的公钥、任务发起时间戳、任务完成时间戳以及链下数据的存储哈希指针等。主链由系统中所有或部分核心节点 (如 SES 和 GS) 共同维护, 负责全局任务协调、状态同步与最终结果的共识确认。
- **子链 (Sub-chain) / 状态通道 (State Channel):** 子链是可选的、面向具体任务或区域的“数据存证链”, 可由参与特定复杂任务的多个节点临时组建, 用于高频记录中间过程数据。更轻量级的方式是采用状态通道技术, 在链下进行大量快速状态更新, 仅将初始状态与最终状态上链。这种设计将大量非关键过程数据由链下或子链处理, 显著减轻主链负担。

说明: 此处应配有一张分层联盟链架构图 (图 2.2), 展示主链和子链的关系, 以及 ST、SES、GS 等节点在链上的角色和数据流向。

**(3) 链上链下协同存储** 对于试验任务产生的海量原始数据和结果数据, 本文采用链上链下协同方式进行存储。链上仅存储数据的哈希值 (Hash) 及描述其元数据的索引

信息，而数据本体则存储在 SES 本地存储或专用分布式存储系统（如 IPFS）中。通过将数据哈希锚定在不可篡改的区块链上，一方面可保障数据完整性与可追溯性（任意链下数据篡改都会导致哈希不匹配），另一方面避免链上存储膨胀，实现效率与安全的平衡。

### 2.2.2 核心机制：智能合约与非实时状态共享

**(1) 基于智能合约的自动化任务管理** 智能合约（Smart Contract）是部署在区块链上、可按预设规则自动执行的程序。在本架构中，智能合约用于实现试验任务全生命周期的自动化与去信任化管理 [6]。

- **任务生命周期合约：**可编写智能合约定义试验任务从发布、竞标（可选）、分配、执行到结果提交与验证的完整流程。例如，当 ST 发布任务时，触发合约函数将任务信息记录上链；当 SES 完成计算并提交结果哈希时，触发另一函数，由指定验证节点（或多方投票）对结果进行校验。校验通过后，合约可自动完成资源计费与信誉评分更新。
- **访问控制合约：**通过智能合约可实现对试验数据与计算资源的精细、动态访问控制。合约可规定只有满足特定条件（如具备特定身份或位于特定区域）的节点才可访问相应数据或调用特定计算服务。

**(2) 基于区块链的非实时全局状态共享** 区块链作为追加式（Append-only）的分布式账本，天然为所有参与节点提供最终一致的全局状态视图。一个 ST 在决定是否卸载新任务时，无需向全网广播查询各 SES 的实时负载，而只需读取主链最新区块，即可获知：

- 当前有哪些任务正在哪些 SES 上执行（通过比对尚未匹配“任务结束”记录的“任务开始”记录）；
- 各 SES 的历史负载情况与平均处理时长；
- 不同节点的信誉评级。

尽管这些信息并非严格实时（新区块产生存在时间间隔），但对于计算卸载这类“准实时”决策而言，这种非实时全局状态已足以提供决策所需的宏观环境感知。它

极大降低了获取系统状态的通信开销，并缓解了网络时延导致的信息不一致问题，是实现分布式智能决策的关键数据基础。这一思想与文献 [17] 中针对信息不对称问题的研究是不谋而合的，并通过区块链提供了可工程落地的实现路径。

通过上述设计，本文构建了一个面向天基网络优化的分布式可信数据管理框架：利用轻量级联盟链保障系统安全可控与高效运行，通过分层架构与链上链下协同解决海量数据存储难题，并借助智能合约与非实时状态共享机制，为上层联邦学习与深度强化学习应用提供可靠、透明、高效的协同底座。

## 2.3 联邦学习的并发开销预测与隐私协作建模

在获得可信、非实时的全局状态视图之后，卸载决策面临的下一个核心难题是：如何准确评估一个卸载动作的“成本”，即任务的预期完成时延。如第 1 章中所述，由于“并发开销”的存在，任务时延并非简单等于其计算量除以节点剩余计算能力，而是一个受当前节点负载、任务类型组合等多因素影响的复杂非线性函数。本节阐述如何利用联邦学习（Federated Learning, FL）技术，以隐私保护方式协作构建高精度的并发开销预测模型。

### 2.3.1 并发开销的数据驱动建模

并发开销产生于现代计算机体系结构中的共享资源竞争，包括 CPU 多级缓存、内存总线、磁盘 I/O 等。当多个计算密集型任务同时运行时，它们会频繁驱逐彼此的缓存行、争抢内存带宽，导致处理器流水线停顿与频繁缺页中断，使每个任务的“每秒有效指令数”显著低于独占 CPU 时的水平。

这种性能下降程度受到多种因素影响，例如：

- 硬件架构：不同型号 CPU 的缓存大小、核心数量、内存控制器性能各不相同；
- 任务类型：访存密集型任务与计算密集型任务混合执行时，与同类任务混合执行相比，其干扰模式显著不同；
- 任务数量与组合：并发任务数量越多，资源竞争通常越激烈；
- 操作系统调度策略：调度器如何分配时间片与 CPU 核心，直接影响任务间相互影响。



由于上述因素高度交织，试图仅依靠解析方法（如排队论）精确建模并发开销极为困难，文献 [14] 已对此有深入讨论。因此，更可行的方法是采用数据驱动思路：将并发开销量化为待学习函数  $f_{co}$ ，通过大量历史运行数据拟合：

$$(L_{new}, \Delta L_{exist}) = f_{co}(\text{NodeSpec}, \text{TaskSet}_{current}, \text{Task}_{new}),$$

其中， $L_{new}$  为新任务预计完成时延， $\Delta L_{exist}$  为已在执行任务的预期额外时延， $\text{NodeSpec}$  为节点硬件规格， $\text{TaskSet}_{current}$  为当前正在执行的任务集合， $\text{Task}_{new}$  为新卸载任务。

在天基网络中，每个 SES 节点本地可用的历史运行数据有限，形成典型的“数据孤岛”。仅凭本地数据训练的模型，泛化能力与预测精度难以保证；而若将所有 SES 数据集中到服务器训练，又会面临巨大的通信开销与严重隐私泄露风险。

### 2.3.2 基于联邦学习的协作式预测模型训练

联邦学习作为“数据不动、模型动”的分布式机器学习范式，为解决上述矛盾提供了理想框架 [3]。其核心思想是：参与方（此处为各 SES）在不共享本地原始数据的前提下，通过交换与聚合本地训练模型参数，共同训练出全局模型，这与 Zhai 等人为 LEO 卫星网络设计的去中心化 FL 框架 FedLEO[7] 的动机高度一致。

本文提出的基于 FL 的并发开销预测模型训练流程如下（如图 2.3 所示）：

说明：此处应配有一张联邦学习流程图（图 2.3），展示中心协调器（如地面站或指定 GEO 卫星）与多个 SES 节点之间的模型分发、本地训练与模型上传聚合过程。

1. **初始化：**由中心协调器（可为地面站或骨干卫星）初始化并发开销预测模型（如 CNN 或 MLP），并将参数分发给所有参与 SES。
2. **本地训练：**每个 SES 节点  $k$  使用本地历史任务运行日志  $D_k$  作为训练集。每条日志记录包含任务组合、节点负载及最终实际时延等信息。节点  $k$  在此基础上利用随机梯度下降（SGD）等算法对本地模型  $w_k$  进行多轮训练，以最小化预测时延与真实时延的误差。
3. **模型上传：**本地训练结束后，每个 SES 节点  $k$  将更新后的模型参数（或增量） $w_k$  发送给协调器。传输过程中只包含模型参数，不包含任何原始运行数据，从而保护本地数据隐私。
4. **全局聚合：**协调器收集多个（或全部）SES 的模型参数后，采用聚合算法生成

新一轮全局模型。最常用的是联邦平均（FedAvg）算法，其更新公式为

$$w_{\text{global}}^{t+1} = \sum_{k=1}^K \frac{n_k}{N} w_k^{t+1},$$

其中  $K$  为参与本轮聚合节点数， $n_k$  为节点  $k$  的本地样本量， $N$  为所有参与节点的总样本量。

5. **模型分发：**协调器将聚合后的全局模型参数  $w_{\text{global}}^{t+1}$  再次分发给所有 SES，替换其本地模型。

通过多轮上述迭代，系统可收敛得到融合各节点数据特征的高精度全局并发开销预测模型。

### 2.3.3 关键技术考量

在将 FL 应用于天基网络实践时，还需重点考虑以下问题：

- **客户端选择（Client Selection）：**每轮训练中不必要求所有 SES 参与。尤其在卫星场景中，部分节点的星地/星间链路条件较差甚至不可用，需要设计合理的参与方选择策略，优先选择信道条件较好、算力充足且数据质量高的节点参与训练，以平衡模型性能和通信开销。
- **异构性问题（Heterogeneity）：**天基网络节点存在系统异构（硬件配置、算力差异）与数据异构（任务类型与负载模式不同，Non-IID 数据）。这可能导致本地模型收敛速度与方向不一致，影响全局模型性能。可采用 FedProx 等算法，在本地损失函数中加入正则项限制本地更新偏离全局模型过远，或采用分层/聚类联邦学习 [8] 等方法缓解异构性影响。
- **增强隐私保护：**尽管 FL 不传输原始数据，但模型更新仍可能泄露部分隐私。可结合差分隐私（Differential Privacy）技术，在 SES 上传模型参数前加噪，实现可量化的更强隐私保障 [9]。

通过引入联邦学习，本文在保护各节点数据隐私的前提下，提出了一种针对“并发开销”精准预测的协作建模方案。得到的全局预测模型将作为下一节深度强化学习智能体的环境感知模块，为其提供更准确的成本预估，从而支撑更优的动态卸载决策。

## 2.4 深度强化学习的动态资源调度与卸载节点优选

面对天基网络拓扑高动态、资源状态不确定等核心挑战，传统基于静态优化或启发式规则的卸载算法难以适应。深度强化学习（Deep Reinforcement Learning, DRL）作为一种无模型（Model-free）的决策优化方法，通过让智能体在与环境的持续交互中“试错”学习，能够自主发现复杂动态约束下的近似最优策略，是解决此类问题的理想工具 [18, 10]。本节将计算卸载问题形式化为马尔可夫决策过程（MDP），并阐明如何利用 DRL 进行求解。

### 2.4.1 计算卸载的马尔可夫决策过程建模

我们将单个卫星终端 ST 的计算卸载决策过程建模为一个 MDP，其核心要素定义如下。

**(1) 智能体（Agent）** 每个需要进行卸载决策的 ST 视为一个独立智能体。

**(2) 状态（State,  $S$ ）** 状态是智能体在决策时可观测的环境信息，合理的状态空间设计是 DRL 成功的关键。时刻  $t$  的状态记为  $s_t$ ，包括：

- 任务属性：  $s_t^{\text{task}} = \{C_i, D_i\}$ ，其中  $C_i$  为当前任务  $i$  的计算量（如 CPU 周期数）， $D_i$  为其数据大小；
- 本地节点状态：  $s_t^{\text{local}} = \{E_{\text{local}}\}$ ，表示本地节点剩余能量；
- 网络环境状态：  $s_t^{\text{net}} = \{Q_j, B_j, L_j\}_{j \in N(t)}$ ，其中  $N(t)$  为当前可达邻近 SES 集合， $Q_j$  为通过区块链获取的  $\text{SES}_j$  任务队列信息（如队长、任务数）， $B_j$  为通过广播或链路探测获得的信道状态（带宽、信噪比等）， $L_j$  为利用第 2.3 节基于 FL 训练的并发开销预测模型得到的在  $\text{SES}_j$  上执行任务的预计时延。

综合而言，

$$s_t = \{s_t^{\text{task}}, s_t^{\text{local}}, s_t^{\text{net}}\}.$$

这一状态设计通过区块链与联邦学习融合了更丰富、准确的环境信息，有效缓解了部分可观测性问题。

(3) **动作 (Action,  $A$ )** 动作是智能体在给定状态下可采取的决策。对于一个拥有  $M$  个可达 SES 节点的 ST，其动作空间为离散集合：

$$A = \{a_0, a_1, \dots, a_M\},$$

其中， $a_0$  表示本地执行任务，而  $a_j$  ( $j = 1, \dots, M$ ) 表示将任务卸载至第  $j$  个 SES 节点。这一建模方式已在多篇任务卸载研究中得到采用 [11, 13]。

(4) **奖励 (Reward,  $R$ )** 奖励函数用于评估在某状态下采取某动作的即时优劣，直接引导智能体学习方向。为实现系统性能综合优化，本文设计多目标奖励函数。在状态  $s_t$  下采取动作  $a_t$ ，系统转移至新状态  $s_{t+1}$  后，智能体获得奖励：

$$r_t = -(\omega_T T_{\text{total}} + \omega_E E_{\text{cost}} + \omega_B B_{\text{imbalance}}),$$

其中：

- $T_{\text{total}}$ ：任务总完成时延。若为本地执行，则为本地计算时延；若为卸载，则为通信时延与远程计算时延之和，远程计算时延由并发开销模型精确预测；
- $E_{\text{cost}}$ ：任务执行的能量消耗，本地执行主要消耗计算能量，卸载执行主要消耗通信能量；
- $B_{\text{imbalance}}$ ：网络负载不均衡度，可通过所有 SES 负载的方差衡量，选择空闲节点有助于降低该值；
- $\omega_T, \omega_E, \omega_B$  为对应权重系数，用于按业务需求（如时延优先、节能优先）调整优化侧重点。

(5) **策略 (Policy,  $\pi$ )** 策略  $\pi(a|s)$  定义了给定状态  $s$  下选择动作  $a$  的概率。DRL 的目标是学习最优策略  $\pi^*$ ，最大化长期折扣奖励的期望：

$$\pi^* = \arg \max_{\pi} \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t r_t \mid s_0, \pi \right],$$

其中  $\gamma \in [0, 1)$  为折扣因子。

### 2.4.2 基于深度 Q 网络 (DQN) 的求解算法

对于上述离散动作空间的 MDP 问题, 深度 Q 网络 (Deep Q-Network, DQN) 是一种经典且有效的求解算法 [10]。DQN 用深度神经网络近似最优动作价值函数  $Q^*(s, a)$ :

$$Q^*(s, a) \approx Q(s, a; \theta),$$

其中  $\theta$  为神经网络参数。网络输入为状态  $s$ , 输出为每个可能动作的 Q 值。在给定状态下, 智能体选择 Q 值最大的动作:

$$a = \arg \max_{a' \in A} Q(s, a'; \theta).$$

训练 DQN 时, 本文借鉴前期工作 [20] 并结合天基网络特点, 采用以下关键技术:

- **经验回放 (Experience Replay)**: 将与环境交互产生的经验元组  $(s_t, a_t, r_t, s_{t+1})$  存入经验池, 训练时从中随机采样小批量数据, 打破时序关联, 提升训练稳定性。
- **目标网络 (Target Network)**: 采用在线网络  $Q(s, a; \theta)$  与目标网络  $Q(s, a; \theta^-)$  双网络结构, 目标网络参数以较低频率从在线网络复制, 有助于稳定训练目标。
- **损失函数**: 以满足贝尔曼方程为目标, 最小化均方误差:

$$L(\theta) = \mathbb{E}_{(s, a, r, s')} \left[ \left( r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2 \right].$$

在天基网络场景中, 可引入文献 [12] 中的时空注意力机制增强 Q 网络对动态状态的表征能力。对于更复杂的协同场景, 还可扩展为多智能体强化学习 (Multi-Agent RL, MARL) 框架 [13, 15], 让多个 ST 智能体共同学习协作策略。

通过 DRL 赋能, 每个卫星终端均可成为具备自主学习与智能决策能力的实体, 能够在不断变化的环境中动态权衡时延、能耗与网络负载等多重目标, 为每个任务选择最优执行路径, 从而在系统层面实现资源的自适应、高效调度与优化。

## 2.5 天基网络计算卸载的形式化建模

在前述章节对系统架构与关键技术进行阐述的基础上, 本节对天基网络计算卸载问题进行统一、形式化的数学建模。该模型整合服务时延、并发开销、能耗等关键性能指标, 并明确本研究旨在求解的优化目标。

### 2.5.1 任务模型

假设在时间  $t$  有一个计算任务  $i$  由卫星终端  $ST_i$  生成，该任务可表示为元组：

$$\text{Task}_i = \{C_i, D_i, T_i^{\max}\},$$

其中：

- $C_i$ ：任务  $i$  的计算复杂度，以所需 CPU 周期数或 FLOPs 表示；
- $D_i$ ：任务  $i$  的数据大小（比特），主要考虑输入数据；
- $T_i^{\max}$ ：任务  $i$  的最大可容忍时延（截止时间）。

### 2.5.2 决策变量

对于每个任务  $i$ ，其卸载决策由向量  $X_i = \{x_{i,j}\}$  表示，其中  $j \in \{0, 1, \dots, M\}$ ， $M$  为可用 SES 节点数量：

$$x_{i,j} \in \{0, 1\}, \quad \sum_{j=0}^M x_{i,j} = 1.$$

其中：

- $x_{i,0} = 1$  表示任务  $i$  在本地  $ST_i$  上执行；
- $x_{i,j} = 1$  ( $j > 0$ ) 表示任务  $i$  被卸载至第  $j$  个 SES 执行。

### 2.5.3 时延模型 (Latency Model)

任务  $i$  的总完成时延  $T_i$  随卸载决策不同而变化。

(1) 本地执行时延 若任务本地执行 ( $x_{i,0} = 1$ )，总时延为

$$T_i^{\text{local}} = \frac{C_i}{f_{\text{local}}},$$

其中  $f_{\text{local}}$  为  $ST_i$  的本地计算能力。

(2) 卸载执行时延 若任务被卸载至  $SES_j$  ( $x_{i,j} = 1, j > 0$ )，其总时延为

$$T_i^{\text{offload},j} = T_i^{\text{up},j} + T_i^{\text{comp},j} + T_i^{\text{down},j},$$

其中：

- 上行传输时延：

$$R_{i,j} = B_{i,j} \log_2 \left( 1 + \frac{P_i H_{i,j}}{N_0} \right), \quad T_i^{\text{up},j} = \frac{D_i}{R_{i,j}},$$

其中  $B_{i,j}$  为带宽， $P_i$  为发射功率， $H_{i,j}$  为信道增益， $N_0$  为噪声功率。下行结果时延  $T_i^{\text{down},j}$  通常可视为常数或近似忽略；

- 远程计算时延：

$$T_i^{\text{comp},j} = T_i^{\text{queue},j} + T_i^{\text{exec},j},$$

其中排队时延  $T_i^{\text{queue},j}$  取决于当前队列状态，执行时延  $T_i^{\text{exec},j}$  由并发开销模型  $f_{\text{co}}$  预测：

$$T_i^{\text{exec},j} = f_{\text{co}}(j, C_i, \text{TaskSet}_j),$$

$\text{TaskSet}_j$  为可从区块链账本推断的当前在  $\text{SES}_j$  上执行的任务集合。

因此，任务  $i$  的总时延统一表示为

$$T_i = x_{i,0} T_i^{\text{local}} + \sum_{j=1}^M x_{i,j} T_i^{\text{offload},j}.$$

#### 2.5.4 能耗模型 (Energy Model)

任务  $i$  的总能耗  $E_i$  亦随卸载决策变化，本文主要关注 ST 侧能耗。

##### (1) 本地执行能耗

$$E_i^{\text{local}} = \kappa C_i,$$

其中  $\kappa$  为每 CPU 周期能耗系数。

##### (2) 卸载执行能耗 主要为上行通信能耗：

$$E_i^{\text{offload},j} = P_i T_i^{\text{up},j} = P_i \frac{D_i}{R_{i,j}}.$$

故总能耗可写为

$$E_i = x_{i,0}E_i^{\text{local}} + \sum_{j=1}^M x_{i,j}E_i^{\text{offload},j}.$$

### 2.5.5 优化目标

本文的最终目标是学习一个卸载策略  $\pi$ ，对于一系列到达的任务  $i = 1, 2, \dots, N$ ，最小化综合时延、能耗与负载均衡成本的长期平均：

$$\min_{\pi} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N (\omega_T T_i + \omega_E E_i + \omega_B B_t),$$

其中  $B_t$  为时刻  $t$  的网络负载不均衡度，可定义为所有 SES 归一化负载的方差：

$$B_t = \text{Var}\left(\frac{\text{Load}_1(t)}{f_1^{\max}}, \dots, \frac{\text{Load}_M(t)}{f_M^{\max}}\right),$$

$\text{Load}_j(t)$  为节点  $j$  在时刻  $t$  的计算负载，可由区块链账本推断， $\omega_T, \omega_E, \omega_B$  为各目标权重。

约束条件包括：

- 卸载决策约束： $\forall i, j : x_{i,j} \in \{0, 1\}, \sum_{j=0}^M x_{i,j} = 1$ ；
- 时延约束： $\forall i : T_i \leq T_i^{\max}$ ；
- 能量约束： $\forall i : E_i \leq E_{\text{budget}}$ ；
- SES 容量约束：任意时刻 SES <sub>$j$</sub>  的总计算资源分配不超过  $f_j^{\max}$ 。

将上述多目标优化问题对应到第 2.4 节中的 DRL 奖励函数，即可通过 DRL 算法进行求解，在满足约束的前提下寻找长期最大化系统效用（或最小化成本）的动态卸载策略。

## 2.6 本章小结

本章围绕天基网络中卫星终端试验任务的计算卸载问题，构建了完整的理论框架与技术体系。首先，定义了面向的三层系统架构，并深入分析了网络拓扑高动态性、资源状态信息不完全、任务并发性能影响以及分布式协同信任缺失等四大核心挑战。



随后, 本文依次阐述了拟采用的关键技术路径: 一是基于分层联盟链的分布式可信管理技术, 通过主子链架构与链上链下协同存储机制以及智能合约, 实现试验任务全生命周期的防篡改存证与非实时全局状态共享; 二是基于联邦学习的并发开销预测技术, 在保护节点数据隐私的前提下协同训练高精度预测模型, 为卸载决策提供精细化成本先验; 三是基于深度强化学习的动态调度技术, 将计算卸载建模为 MDP, 并设计综合考虑时延、能耗和负载均衡的多目标奖励函数, 使智能体在高度动态环境中自适应学习近似最优策略。

最后, 本章对天基网络计算卸载问题进行了统一的形式化建模, 清晰给出任务模型、决策变量、时延与能耗计算方式以及系统优化目标, 形成“区块链提供状态 → FL 预测成本 → DRL 优化决策”的技术闭环, 为后续章节的算法设计与仿真实验提供了坚实理论基础。

## 第3章 基于区块链的试验任务可信管理架构

### 3.1 本章概述

本章旨在为融合联邦学习（Federated Learning, FL）、深度强化学习（Deep Reinforcement Learning, DRL）与区块链的天基网络计算卸载策略构建一个系统性、可信且高效的总体架构。在天基网络高动态、资源受限且数据敏感的特殊环境下，单一技术难以应对计算卸载带来的多重挑战。因此，本章重点阐述如何将区块链的去中心化信任机制、联邦学习的隐私保护协同建模能力以及强化学习的动态自适应决策优势有机结合，形成功能完备、逻辑自洽的体系。

首先，在第3.1节中，提出一个分层的总体架构，明确数据层、网络层、共识层、合约层及应用层的核心功能与交互接口，并通过系统总体架构图直观展示各组件的协同关系。该设计为后续章节中具体技术和算法的实现提供宏观指导和约束。

其次，在第3.2节中，聚焦构建名为“分层联盟链”的可信数据底座。通过设计“主链-子链”结构，实现试验任务全局管理与具体数据存证的有效解耦，在保障数据可信、可追溯的同时，提升系统的吞吐量和可扩展性。本节将详细定义主链和子链的区块结构，并设计一系列关键智能合约以自动化任务的全生命周期管理。

最后，在第3.3节中，讨论一种基于区块链账本的非实时状态共享机制。该机制旨在解决分布式决策中的全局信息获取难题，允许卫星终端在不产生高额并发通信开销的前提下，通过查询链上数据获取网络负载、资源占用等“近似”全局状态，并将其作为DQN智能体决策的关键输入。这一设计体现了在网络通信资源极端宝贵的天基场景下，对一致性与时效性权衡的工程思路。

通过本章的设计，为后续章节中的并发开销预测模型、分布式卸载决策算法以及系统仿真验证奠定坚实的架构基础。

### 3.2 总体架构设计

为整合区块链、联邦学习与深度强化学习以应对天基网络计算卸载的挑战，本文设计了一个分层式总体架构，构建集可信存证、隐私保护与智能决策于一体的闭环系统，解决卫星终端在资源受限、高动态变化和数据安全需求下的高效协同问题。

系统总体架构如图3.1所示，自下而上分为数据层、网络层、共识层、合约层和

应用层。各层职责明确，通过标准化接口协同工作，支撑上层的计算卸载与智能调度应用。

说明：此处应配有系统总体架构图（图 3.1），示意各层组件及其交互关系。

### 3.2.1 各层功能定义

**(1) 数据层 (Data Layer)** 数据层作为架构基石，负责数据的持久化存储与管理，采用链上链下协同存储策略：

- 分层区块数据：核心组成是分层联盟链的区块数据，包括主链和子链。关键且需要强一致性保证的元数据，如任务摘要、状态转换记录、模型更新凭证等，存储在链上。
- 链下存储：原始的大容量试验数据及联邦学习模型参数等，通过其哈希值上链的方式，存储于链下分布式存储系统（如 IPFS）或 SES 本地存储中，从而避免区块链膨胀，保证系统性能与可扩展性。

**(2) 网络层 (Network Layer)** 网络层负责节点间通信与数据路由：

- P2P 网络协议：构建去中心化通信网络，负责节点发现、区块与交易的广播和同步，确保区块链网络的基本连通性。
- 跨链网关：实现主链与子链、不同子链之间的信息互通。跨链网关负责处理跨链消息的路由、验证和转发，支撑分层联盟链架构中的数据交互与状态同步。

**(3) 共识层 (Consensus Layer)** 共识层确保分布式账本的一致性与不可篡改性，是区块链信任机制的核心：

- 共识算法：在许可接入的联盟链环境中，采用投票类拜占庭容错（BFT）算法，如实用拜占庭容错（PBFT）及其变体。这类算法在部分节点失效或作恶的情况下仍能保证全网数据一致性，并且相较工作量证明（PoW）具有更低延迟与更高吞吐量，更适合具有实时性需求的天基网络场景。

**(4) 合约层 (Contract Layer)** 合约层定义并执行业务逻辑，实现系统的自动化与智能化管理：

- 智能合约：部署在区块链上的可执行代码，用于定义任务管理、资源调度与联邦学习等过程的规则。本文设计四类核心智能合约：
  - 任务发布合约：处理任务创建、发布、参数定义与资源需求的链上注册；
  - 状态更新合约：追踪并记录任务从“已发布”“执行中”到“已完成”等状态的流转；
  - 结果验证合约：定义任务结果验证规则并记录验证结果，为后续激励与惩罚提供依据；
  - FL 模型管理合约：管理联邦学习过程，包括模型版本记录、参与方列表维护、全局模型更新摘要存证等，增强联邦学习过程的透明度与可审计性。

**(5) 应用层 (Application Layer)** 应用层直接面向用户，是整个架构价值的最终体现：

- 天基试验任务应用：用户通过该接口提交试验任务、查询任务状态并获取结果；
- 分布式智能卸载决策 (DQN)：每个卫星终端内嵌 DQN 智能体，通过查询账本获取非实时状态共享信息，结合本地观测做出任务本地执行或卸载至某 SES 的决策；
- 联邦学习训练 (FL)：多个卫星终端之间开展协同建模，各终端使用本地数据训练并发开销预测模型等，仅上传模型参数至聚合节点，在保护隐私前提下构建高精度全局模型。

### 3.2.2 数据流与控制流

系统核心工作流程围绕数据流与控制流展开，可概括为以下三个闭环。

#### (1) 卸载请求与决策流程 (控制流)

1. 任务产生：卫星终端在应用层产生计算密集型任务；
2. 状态查询：终端中的 DQN 智能体通过合约层接口查询数据层账本，获取其他节点任务负载、资源占用等“近似”全局状态；

3. 智能决策：DQN 智能体结合链上状态与本地观测信息，做出任务本地执行或卸载至某边缘节点的决策；
4. 任务发布：若决定卸载，终端调用任务发布合约将任务元数据及目标节点信息写入区块链，并通过网络层将原始任务数据发送至目标 SES。

## (2) 联邦学习与模型更新流程（数据流与控制流）

1. 本地训练：各终端在应用层利用本地采集的任务执行数据（时延、资源占用率等）训练并发开销预测模型；
2. 模型上传：训练完成后，终端将模型参数更新上传至聚合节点（可为特定卫星或地面站）；
3. 全局聚合：聚合节点执行 FedAvg 等聚合算法生成新的全局模型；
4. 模型存证与分发：聚合节点调用 FL 模型管理合约，将新全局模型的版本号与哈希摘要记录在链上，随后通过网络层将模型分发给各终端，用于后续卸载决策。

## (3) 任务执行与账本记录流程（数据流）

1. 任务执行：边缘节点接收任务后开始执行计算；
2. 状态更新：在任务执行关键阶段（如开始执行、执行完成），边缘节点或任务发起方调用状态更新合约，将最新任务状态写入区块链；
3. 结果验证与存证：任务完成后，结果由指定验证节点（或多方）通过结果验证合约进行验证。验证通过后，将结果哈希及验证记录存证至主链，原始结果数据存放于链下。

通过上述设计，总体架构将区块链的可信记录能力、联邦学习的隐私保护建模能力与深度强化学习的自适应决策能力深度融合，为天基网络环境下的计算卸载提供了一个完整、高效且可信的解决方案。

### 3.3 分层联盟链可信数据底座

为高效、可信地管理天基网络中分布式试验任务与海量数据，本文设计了一种基于分层联盟链的可信数据底座。传统单链架构在处理大规模交易和异构数据时易遭遇性能瓶颈和治理复杂性问题，而分层联盟链通过将不同职责的业务逻辑划分到不同链上，实现功能解耦与性能优化，是应对天基网络复杂性的有效方案。

本节设计的核心是“主链-子链”架构，其结构如图 3.2 所示。

说明：此处应配有分层联盟链结构图（图 3.2），展示主链、子链与跨链网关的关系。

#### 3.3.1 主链-子链职责边界

**(1) 主链 (Main Chain)** 主链作为系统的“信任根”和“控制中枢”，由地面控制中心、区域网关等核心节点构成，具有较高安全级别与全局管理权限，其核心职责包括：

- 全局任务管理与治理：负责试验任务的发布、分配及最终状态记录，所有任务生命周期关键节点（创建、完成、失败等）在主链上进行最终确认与共识，同时承担节点准入、权限变更等联盟治理职能；
- 全局状态摘要与锚定：定期锚定各子链的状态摘要（如最新区块头哈希），实现对子链数据的全局可验证性，而无需在主链上存储全部子链数据，从而显著减轻主链存算负载；
- 跨链协调与控制：通过跨链网关管理子链生命周期，包括子链创建、成员准入/退出授权以及跨链交互规则与协议的制定。

**(2) 子链 (Sub-chain)** 子链作为“数据存证与执行层”，可按地理区域、业务部门或任务类型（如遥感图像处理、通信信号监测）创建多条，由相关卫星边缘节点与地面站共同维护，其职责包括：

- 具体任务数据存证：存储与特定任务相关的操作日志、中间结果、传感器读数等海量高频数据，在子链内直接记录共识，实现业务级快速存证与审计；
- 局部共识与执行：在子链参与节点间达成局部共识，局部化处理显著降低交易确认延迟，提升整体吞吐量。

**(3) 跨链网关 (Cross-Chain Gateway)** 跨链网关作为主链与子链之间的可信桥梁, 通常部署在同时参与主链与相应子链的节点 (如区域网关) 上, 负责:

- 监听源链上的特定事件 (如主链任务发布事件);
- 验证事件的合法性与完整性;
- 在目标链上触发相应交易 (如在子链创建任务记录), 实现安全可靠的链间数据交换与业务协同。

这种分层架构通过将全局治理与具体数据处理分离, 有效克服单链架构性能瓶颈, 在确保全局一致性与数据可追溯性的同时, 提供良好可扩展性与工程可落地性。

### 3.3.2 区块结构定义与智能合约流程

为支撑分层架构的高效运行, 本文对主链与子链的区块结构和智能合约流程进行针对性设计, 实现任务全生命周期的自动化、透明化与可信化。

**(1) 区块结构定义** 主链区块主要承载全局任务信息与跨链证明, 而子链区块则专注于存储具体试验数据与执行日志。示意性字段定义如下。

- 主链区块结构:
  - 区块头: 包含版本号、时间戳、前一区块哈希、Merkle 树根 (聚合任务列表与子链状态树)、Nonce 等标准字段;
  - 任务列表: 记录本区块内发生状态变更的全局任务。每条记录包含任务 ID、发起者标识、执行节点标识、当前状态 (发布/执行中/完成/失败)、任务内容摘要等;
  - 子链状态树根哈希: 通过 Merkle 树等形式聚合在该时间窗口内被锚定到主链的各子链状态哈希, 用于实现对子链数据的全局可验证性。
- 子链区块结构:
  - 区块头: 包含版本号、时间戳、前一区块哈希、Merkle 树根 (交易/数据列表)、指向主链对应任务区块的哈希链接等;
  - 交易/数据列表: 记录详细试验数据或操作日志, 每条记录包含试验 ID、操作者、时间戳、数据内容 (或链下存储地址哈希) 等。

**(2) 基于智能合约的任务生命周期流程** 任务生命周期主要通过主链上的一系列智能合约驱动与管理，确保工程可落地性与性能权衡。其流程如图 3.3 所示。

说明：此处应配有基于智能合约的任务生命周期流程图（图 3.3）。

1. 任务发布：试验发起方调用任务发布合约在主链创建新任务。合约验证发起方权限与任务参数合法性后，记录任务 ID、需求、预算等信息，并触发“任务已发布”事件，任务进入待分配状态。
2. 节点选择与任务执行：卫星边缘节点监听“任务已发布”事件发现新任务，若决定承接则调用合约锁定任务。状态更新合约将任务状态置为“执行中”，并记录承接节点信息；随后边缘节点在其所属子链上记录详细执行日志。
3. 结果提交与验证：任务完成后，边缘节点将结果哈希及指向子链存证数据的链接提交至主链结果验证合约，合约按预设规则（指定权威节点验证或多方投票等）触发验证流程。
4. 状态终结与激励：
  - 若验证通过，状态更新合约将任务状态标记为“已完成”，并自动触发激励机制（如信誉积分增加、代币支付等）；
  - 若验证失败，任务状态标记为“已失败”，并可触发相应惩罚或仲裁机制。

所有状态最终变更均固化在主链上，构成不可篡改的“最终事实”。通过这套基于智能合约的自动化流程，分层联盟链不仅为天基网络试验提供可信数据存证能力，也大幅提升任务管理效率与透明度，为构建无需中心化信任的分布式协同环境奠定坚实基础。

### 3.4 基于账本的非实时状态共享机制

在天基网络这种高度动态且通信资源受限的环境中，获取精确、实时的全局网络状态（如所有节点的 CPU 占用率、任务队列长度等）几乎不可能。传统集中式调度方案依赖周期性状态上报，带来巨大信令开销与明显决策延迟，尤其在大规模节点网络中问题突出。为此，本文提出一种基于账本的非实时状态共享机制。



该机制核心思想是：利用区块链账本作为分布式、最终一致的全局状态信息源，允许各决策节点（卫星终端）在无需高频实时通信的情况下，获取“近似”的全局状态视图，并将其作为本地 DQN 智能决策的重要输入。

### 3.4.1 机制原理

如图 3.4 所示，区块链账本由网络中所有共识节点共同维护。当网络中发生关键事件（如新任务卸载至某边缘节点、任务完成释放资源）时，相应智能合约被触发，将这些状态变更以交易形式写入新区块。随着区块的广播与确认，这些信息被同步到所有节点的本地账本副本中。

说明：此处应配有“节点通过账本感知全局状态”示意图（图 3.4）。

某卫星终端节点（如节点 A）在需要做出卸载决策时，无需向全网广播查询请求，而是仅需：

1. 本地查询：直接查询本地维护的账本副本；
2. 状态重构：解析账本中与任务和资源相关的历史交易记录，重构网络中其他节点在“不久之前”的状态快照。例如，通过扫描最近若干区块的任务发布与完成记录，估算节点 B、C 当前任务队列长度与资源占用情况；
3. 决策输入：将重构出的“近似”全局状态信息与本地状态（如本地 CPU 负载、电池余量）一同作为 DQN 的状态输入。

### 3.4.2 一致性与时效性的折中原则

该机制获取的状态并非严格实时，其时效性受出块速度与网络同步延迟影响，本质上体现了在一致性与时效性之间的折中。

**(1) 强一致性与弱时效性** 区块链保证所有节点最终可见完全一致的状态历史（强最终一致性）。但从事件发生到被打包进区块并完成确认存在固有延迟（在联盟链中一般为秒级），因此节点查询到的状态是“过去”的快照。

#### (2) 查询频率与折中原则

- 原则：查询频率应根据任务时延敏感度与网络动态性自适应调整；

- 高时效性需求：对于极度时延敏感任务，决策前可触发一次对最新区块的强制同步，以获取尽可能新的状态，但会增加单次决策的成本；
- 一般性需求：对于大多数任务，可在 DQN 固定决策时间步内（如每秒）查询本地账本最新状态，通信开销最低；
- 成本-效益：该机制将状态同步成本平摊至区块链网络的日常维护，避免为每次决策单独发起昂贵的实时轮询，在天基网络中尤为重要。

### 3.4.3 对 DQN 决策的支撑

在这种非实时共享机制下，DQN 智能体可以逐步学会在信息不完全且存在延迟的环境中做出鲁棒决策：

- 学习含噪环境下策略：DQN 通过大量交互经验学习在带噪声和延迟的状态输入下获得长期最优回报的策略，而非依赖完全精确的瞬时状态；
- 时效性加权：在状态表示设计中，可对来自账本的信息进行时效性加权处理，将最新区块信息赋予更高权重，从而引导模型更多关注近期网络变化趋势。

综上，基于账本的非实时状态共享机制是针对天基网络特点量身定制的轻量级、高容错全局状态感知方案。它充分利用区块链的数据共享特性，以可接受的时效性损失换取通信开销的大幅降低和系统鲁棒性的提升，为实现真正分布式的智能卸载决策提供了可行的数据基础。

## 3.5 本章小结

本章围绕天基网络中卫星终端试验任务的可信管理与智能卸载需求，构建了一个融合区块链、联邦学习与深度强化学习的系统级架构。

首先，从层次化视角提出总体架构设计，明确了数据层、网络层、共识层、合约层与应用层的职责划分与协同方式，解释了任务卸载、联邦学习训练与任务执行三大关键流程的数据流与控制流路径。其次，围绕分层联盟链可信数据底座，详细定义主链-子链职责边界及区块结构，并设计基于智能合约的任务生命周期管理流程，实现了试验任务从发布、执行到结果存证的全流程自动化与可审计性。再次，通过引入基于账本的非实时状态共享机制，在保证系统全局一致性与可信性的前提下，大幅降低

了获取全局状态的通信开销，为在信息不完全环境下运行的 DQN 智能体提供了可靠的环境感知基础。

本章构建的“分层联盟链可信数据底座 + 非实时状态共享 + 智能合约驱动任务管理”框架，与前一章中的 FL 并发开销预测与 DRL 卸载决策模型形成有机衔接，为后续章节中算法实现与系统仿真验证提供了统一的架构支撑。

## 第4章 基于联邦学习与DQN的分布式卸载策略

在天基网络的高动态、资源受限环境下，计算卸载决策不仅要考虑任务自身的执行效率，还必须应对多任务并发执行时产生的资源竞争问题。第三章构建了基于分层联盟链的可信计算卸载架构，为分布式节点间的状态共享和协同工作提供了数据底座。本章将在此架构基础上，聚焦于设计一种智能、高效且分布式的计算卸载策略。

本章的核心思路是：首先，对多任务并发执行所导致的性能下降问题进行量化建模，即定义并刻画“并发开销”；其次，针对该开销难以用传统解析模型精确描述的难题，引入联邦学习方法，在保护各节点数据隐私的前提下，协作训练一个高精度的并发开销预测模型；最后，基于深度强化学习构建一个智能卸载决策代理（Agent），该代理能够综合利用本地观测、非实时状态共享信息以及并发开销预测结果，做出最优的卸载决策。我们将这一融合了联邦学习（FL）与深度Q网络（DQN）的策略称为FL-DQN。

### 4.1 并发开销问题建模

在卫星边缘计算场景中，当一个或多个新的计算任务被卸载到某个边缘卫星节点（ESN）时，该节点上正在执行的任务会因CPU时间片、内存带宽、I/O等资源的争抢而受到影响，导致其剩余执行时间延长。同时，新任务自身也因为无法独占全部计算资源而导致处理时延增加。这种由于任务并发而引起的系统整体性能下降，我们将其定义为并发开销（Concurrency Overhead）。

与地面数据中心相比，卫星边缘节点的计算和存储资源更为宝贵且受限（SWaP约束），因此并发开销问题在天基网络中尤为突出。一个忽略并发开销的卸载策略可能会导致决策失误：表面上看，将任务卸载至某个轻负载节点似乎是优选，但若该节点正在执行一个关键的长耗时任务，新任务的加入可能导致关键任务错过其截止时间，造成更大的系统损失。

#### 4.1.1 并发开销的数学定义

为了量化这一影响，我们将任务  $T_i$  到达边缘节点  $n$  时所产生的并发开销  $\Delta(l_i)$  定义为：新任务  $T_i$  自身的执行时延增量，与由于  $T_i$  的加入导致节点  $n$  上所有已在执行任务集合  $\mathcal{T}_n^{\text{exec}}$  的剩余执行时延增量之和。

$$\Delta(l_i) = \Delta_{\text{self}} + \sum_{e \in \mathcal{T}_n^{\text{exec}}} \Delta_e \quad (4.1)$$

其中：

- $\Delta_{\text{self}}$  是新任务  $T_i$  自身的时延增量。
- $\Delta_e$  是已在执行的任务  $T_e$  的剩余执行时延增量。

并发开销  $\Delta(l_i)$  是一个复杂的复合非线性映射，其值取决于边缘节点  $n$  在任务  $T_i$  到达时的状态  $s_n$ 、新任务  $T_i$  的特征  $\tau_i$ ，以及节点  $n$  上正在运行的任务集  $\mathcal{T}_n^{\text{exec}}$  和资源分配函数  $g(\cdot)$ ：

$$\Delta(l_i) = f(s_n, \tau_i, \mathcal{T}_n^{\text{exec}}, g(\cdot)) \quad (4.2)$$

具体来说：

- 节点状态  $s_n$ ：包括 CPU 负载、内存使用率、任务队列长度、正在执行的任务数量及类型等。
- 任务特征  $\tau_i$ ：主要由其计算量  $C_i$  和数据大小  $D_i$  决定。
- 资源分配函数  $g(\cdot)$ ：代表节点  $n$  的资源分配与调度策略，例如 CFS（Completely Fair Scheduler）等。

#### 4.1.2 并发开销的近似可分表达

在理想情况下，如果任务之间完全独立且资源可以无限分割，那么并发开销可以被简单地认为是线性的。然而在实际中，由于缓存争用、I/O 瓶颈和上下文切换等因素，并发开销呈现出高度的非线性。

为了对此进行近似分析，我们做出如下假设：

1. 队列模型：任务在边缘节点上遵循 M/G/1 排队模型。
2. 资源共享：节点的计算资源  $F_n$  在所有  $k$  个并发任务间按某种策略公平共享。

在这些假设下，我们可以推导出并发开销的“近似可分”表达式。当新任务  $T_i$  加入时，节点上的任务数从  $k - 1$  变为  $k$ 。假设资源被平均分配，那么每个任务获得的计算资源从  $F_n/(k - 1)$  降低到  $F_n/k$ 。

对于新任务  $T_i$ （我们称之为任务  $l$ ），其自身时延增量  $\Delta_{\text{self}}$  是其在并发环境下的执行时延与理想独占时延的差：

$$\Delta_{\text{self}}(l) = \left( \frac{C_l}{F_n/k} \right) - \frac{C_l}{F_n} = (k - 1) \frac{C_l}{F_n} \quad (4.3)$$

对于已经在执行的任务  $T_e$ ，其剩余计算量为  $C_e^{\text{rem}}$ 。在新任务加入前后，其完成剩余计算量所需的时间分别为：

$$T_{\text{before}} = \frac{C_e^{\text{rem}}}{F_n/(k - 1)} \quad (4.4)$$

$$T_{\text{after}} = \frac{C_e^{\text{rem}}}{F_n/k} \quad (4.5)$$

因此，任务  $T_e$  的时延增量  $\Delta_e$  为：

$$\Delta_e = T_{\text{after}} - T_{\text{before}} = C_e^{\text{rem}} \left( \frac{k}{F_n} - \frac{k - 1}{F_n} \right) = \frac{C_e^{\text{rem}}}{F_n} \quad (4.6)$$

总的并发开销  $\Delta(l)$  可以近似表示为：

$$\Delta(l) = (k - 1) \frac{C_l}{F_n} + \sum_{e \in \mathcal{T}_n^{\text{exec}}} \frac{C_e^{\text{rem}}}{F_n} \quad (4.7)$$

这个表达式虽然是近似的，但它揭示了并发开销的关键可解释变量：

- 新任务计算量  $C_l$  (FLOPs)：直接影响自身时延增量。
- 并发任务数  $k$  (队列长度)：对新旧任务都有放大效应。
- 已执行任务的剩余计算量  $C_e^{\text{rem}}$  (负载向量)：共同构成了对系统的存量冲击。

然而，真实的资源分配远比平均分配复杂，剩余计算量  $C_e^{\text{rem}}$  也难以精确获得。因

此，使用传统的排队论或解析模型难以精确刻画这种复杂的动态耦合关系。这促使我们采用数据驱动的方法，利用机器学习模型来学习这个复杂的非线性函数  $f(\cdot)$ ，从而实现了对并发开销的准确预测。

## 4.2 基于联邦学习的并发开销预测

为了解决并发开销难以精确建模的问题，我们采用一种数据驱动的方法，即利用机器学习模型来学习并预测其大小。然而，在分布式天基网络中，每个卫星节点的历史任务数据都属于其本地隐私信息，不宜直接上传至中心服务器进行集中式训练。为此，我们设计了一种基于联邦学习（Federated Learning, FL）的并发开销预测方案。

### 4.2.1 并发开销预测模型

我们设计一个深度神经网络模型来预测并发开销。该模型的输入是与并发开销相关的多维特征，输出是预测的开销值（一个标量）。

**输入张量：**我们将输入特征构造成一个多维张量。对于一个待决策的任务  $T_i$  和目标节点  $n$ ，输入张量  $I_{i,n} \in \mathbb{R}^d$  包含以下信息：

- **节点状态（Node State）：**节点  $n$  的实时 CPU/内存占用率、任务队列长度、正在执行的任务数量。
- **链路条件（Link Condition）：**节点与其他节点间的平均通信带宽和时延，这间接反映了网络的拥塞程度。
- **新任务特征（New Task Profile）：**待卸载任务的计算复杂度（FLOPs）和输入数据大小。
- **在执行任务摘要（Executing Tasks Profile）：**节点上正在执行任务的平均计算量、平均数据大小等统计特征。

**输出：**模型的输出是预测的并发开销  $\hat{\Delta}(l_i)$ ，可以是一个标量（总开销），也可以是一个向量（包含对自身和他人的影响）。本文中我们预测一个标量值。

**模型结构：**我们采用一个混合网络结构，如图 4.1 所示。

- **CNN/LSTM 层：**用于提取节点状态和在执行任务摘要的时序特征。例如，一个一维 CNN（卷积核大小为 3，步长为 1）可以捕捉到负载变化的局部模式。

- **MLP 层：**用于处理新任务特征等静态信息。一个包含两个隐藏层（宽度分别为 64 和 32，使用 ReLU 激活函数）的 MLP 可以有效提取其非线性特征。
- **融合与输出：**两部分特征被展平（Flatten）并拼接（Concatenate）后，送入最后的 MLP 输出层，该层包含一个或多个全连接层，最终输出并发开销的预测值。

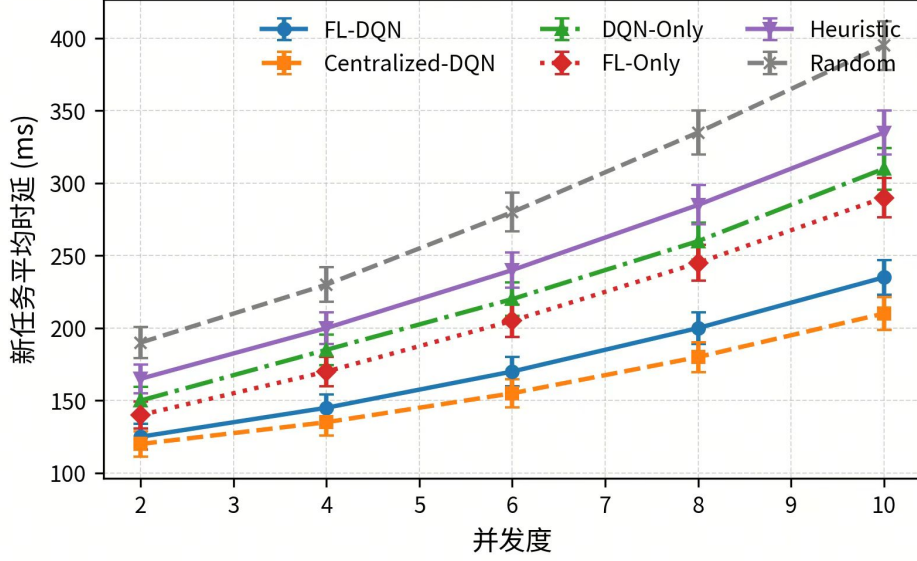


图 4.1 并发开销预测模型结构

**损失函数：**我们使用均方误差（Mean Squared Error, MSE）作为损失函数，以最小化预测值与真实观测值之间的差距。加入 L2 正则项以防止过拟合。

$$L(W) = \frac{1}{B} \sum_{j=1}^B (\Delta(l_j) - \hat{\Delta}(l_j; W))^2 + \lambda \|W\|_2^2 \quad (4.8)$$

其中， $B$  是批量大小， $W$  是模型参数， $\lambda$  是正则化系数。

#### 4.2.2 聚类式联邦学习训练流程

在天基网络中，不同轨道、不同服务区域的卫星边缘节点其硬件配置（ $F_n$ ）和承载的任务类型可能存在显著差异，导致数据呈现典型的 Non-IID（非独立同分布）特性。如果直接采用传统的 FedAvg 算法对所有节点进行聚合，可能会导致全局模型性能不佳。

为了缓解 Non-IID 和设备异构问题，我们采用聚类联邦学习（Clustered Federated Learning, CFL）<sup>Sattler2020</sup>的思想。其核心在于将具有相似数据分布或模型更新趋势的节



点划分为簇（Cluster），在簇内进行模型聚合，从而使得每个簇的聚合模型更能适应其成员节点的局部数据特性。

联邦学习训练的详细流程如图 4.2 所示，具体步骤如下：

1. **初始化：**由主链上的一个或多个协调节点（或地面站）作为中心服务器，初始化一个全局的并发开销预测模型  $W_0$ ，并将其分发给所有参与训练的边缘卫星节点。
2. **参与者选择：**在每个通信轮次  $t$ ，服务器根据预设策略（如资源最优策略，选择电量充足、算力强的节点）选择一部分节点  $S_t$  参与本轮训练。
3. **本地训练：**每个被选中的参与节点  $k \in S_t$  使用其本地存储的历史任务数据（包含节点状态、任务特征和真实的并发开销记录），在当前模型  $W_t$  的基础上进行多轮本地训练（例如， $E$  个 epoch），得到更新后的本地模型  $W_{t+1}^k$ 。
4. **模型上传与聚类：**各节点将本地更新后的模型参数  $W_{t+1}^k$ （或梯度）上传至中心服务器。服务器计算所有上传模型两两之间的余弦相似度，构建一个相似度矩阵，并基于此矩阵运行聚类算法（如层次聚类），将节点动态地划分为  $C$  个簇。
5. **簇内聚合：**服务器在每个簇  $c \in \{1, \dots, C\}$  内部分别执行模型聚合操作。聚合权重  $p_k$  可以基于每个节点的本地数据量大小来设置： $p_k = n_k / \sum_{j \in c} n_j$ 。簇  $c$  的新模型  $W_{t+1}^c$  计算如下：

$$W_{t+1}^c = \sum_{k \in c} p_k W_{t+1}^k \quad (4.9)$$

6. **模型分发：**服务器将各个簇的新模型  $W_{t+1}^c$  分发给对应簇内的成员节点，开启下一轮的本地训练。
7. **模型更新上链：**为了保证训练过程的可追溯性和模型的可信度，每隔一定的通信轮次，协调节点会将聚合后的模型摘要（如版本号、哈希值、参与节点列表等）记录到分层联盟链的主链上，形成非实时状态共享的一部分。

通过这种方式，我们可以在不泄露各卫星节点原始任务数据的前提下，训练出多个针对不同场景、更加精准的并发开销预测模型，为后续的 DQN 智能卸载决策提供关键的输入。

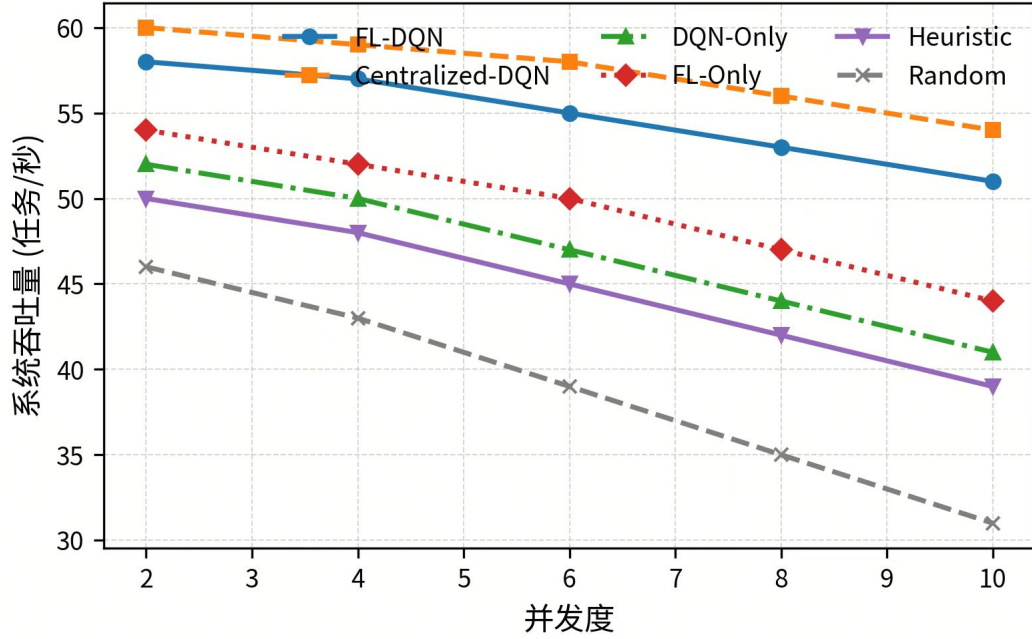


图 4.2 聚类式联邦学习训练流程

### 4.3 基于 DQN 的智能卸载决策

在获得了对并发开销的准确预测能力后，卫星终端需要一种智能决策机制，以决定一个新任务是应该在本地执行，还是卸载到某个边缘卫星节点。考虑到天基网络环境的动态性和决策的复杂性，我们采用深度强化学习（DRL）中的深度 Q 网络（DQN）来构建此决策模型。

每个卫星终端都作为一个独立的智能体（Agent），通过与环境的交互来学习最优的卸载策略  $\pi^*$ ，目标是最大化长期累积奖励。

#### 4.3.1 DQN 模型构建

我们将卸载决策问题建模为一个马尔可夫决策过程（MDP），其关键要素定义如下：

**状态空间 (State,  $\mathcal{S}$ ):** 状态  $s_t \in \mathcal{S}$  是智能体在决策时刻  $t$  所能观测到的所有信息的集合。它必须包含做出明智决策所需的所有相关信息：

- 任务信息：当前待决策任务  $T_i$  的特征，即计算量  $C_i$  和数据大小  $D_i$ 。
- 本地资源状态：终端自身的 CPU 占用率、可用内存、队列长度等。
- 网络链路质量：与各个候选边缘卫星节点之间的预估数据传输速率  $R_{m,n}$  和时

延。

- 邻居节点状态（来自区块链）：通过查询分层联盟链的账本，获取候选边缘节点  $n$  的非实时状态共享信息。这包括该节点的近似负载摘要（如正在执行的任务数量、平均资源占用率）。
- 并发开销预测：对每个候选卸载节点，使用 4.2 节中 FL 训练的模型预测若将任务  $T_i$  卸载至该节点产生的并发开销  $\hat{\Delta}(l_i)$ 。

**动作空间 (Action,  $\mathcal{A}$ )**：对于一个任务  $T_i$ ，智能体的动作空间是离散的，包含所有可能的执行选择：

$$\mathcal{A}_t = \{a_0, a_1, \dots, a_N\} \quad (4.10)$$

其中， $a_0$  代表在本地执行任务， $a_n$ （对于  $n \in \{1, \dots, N\}$ ）代表将任务卸载到第  $n$  个候选边缘卫星节点。

**奖励函数 (Reward,  $R$ )**：奖励函数  $R_t$  用于引导智能体学习期望的行为。我们的目标是最小化任务的完成时延和能耗，同时也要惩罚由决策引起的过高并发开销。因此，我们将奖励函数设计为总成本的负值。借鉴小论文<sup>Zhang2024</sup>中的思想，奖励  $R_t$  定义为：

$$R_t = -(\alpha \cdot T_i^{\text{total}} + \beta \cdot E_i^{\text{total}} + \gamma \cdot \hat{\Delta}(l_i)) \quad (4.11)$$

其中：

- $T_i^{\text{total}}$  和  $E_i^{\text{total}}$  分别是根据第二章定义的任务总时延和终端总能耗。
- $\hat{\Delta}(l_i)$  是由联邦学习训练出的模型所预测的并发开销。
- $\alpha, \beta, \gamma$  是权重系数， $\alpha + \beta + \gamma = 1$ ，分别代表对时延、能耗和并发开销的重视程度。

#### 4.3.2 训练与决策机制

我们采用带有目标网络和经验回放的 DQN 算法进行训练。其核心是学习一个动作价值函数  $Q(s, a; \theta)$ ，它用参数为  $\theta$  的深度神经网络来近似。训练的目标是最小化损失函数  $L(\theta)$ ：

$$L(\theta) = \mathbb{E}_{(s,a,r,s') \sim \mathcal{D}} \left[ \left( r + \gamma \max_{a'} Q(s', a'; \theta^-) - Q(s, a; \theta) \right)^2 \right] \quad (4.12)$$

这即是经典的 Bellman 方程的时序差分形式。其中：

- $\mathcal{D}$  是经验回放池。
- $\theta^-$  是目标网络的参数，它周期性地从主网络参数  $\theta$  复制而来，以稳定训练过程。
- $\gamma$  是折扣因子，决定了未来奖励的重要性。

训练超参数设定：

- 学习率 (Learning Rate)：设为  $10^{-4}$ ，使用 Adam 优化器。
- 折扣因子 (Discount Factor,  $\gamma$ )：设为 0.99，表示对长期回报的重视。
- 目标网络更新周期：每 100 个训练步更新一次目标网络。
- 批量大小 (Batch Size)：每次从经验回放池中采样 64 个样本进行训练。
- 记忆库大小 (Replay Buffer Size)：设置为 10,000，存储最近的交互经验。

**探索-利用策略：**我们采用  $\epsilon$ -贪婪 ( $\epsilon$ -greedy) 策略来平衡探索与利用。 $\epsilon$  的值从初始的 1.0 线性衰减到 0.01，衰减步数为 5000 步。这意味着在训练初期，智能体有很大概率随机探索动作，随着训练的进行，它会越来越倾向于选择当前 Q 值最高的”最优”动作。

FL-DQN 卸载决策的完整算法流程如图 4.3 所示。

## 4.4 实验与仿真

为了验证本章提出的 FL-DQN 卸载策略的有效性、鲁棒性与可扩展性，我们搭建了一个高保真的离散事件仿真平台，并遵循顶会期刊的实验呈现规范，系统地开展了一系列仿真实验。本节将详细阐述实验设置、评价指标、基线算法以及对仿真结果的深入分析。

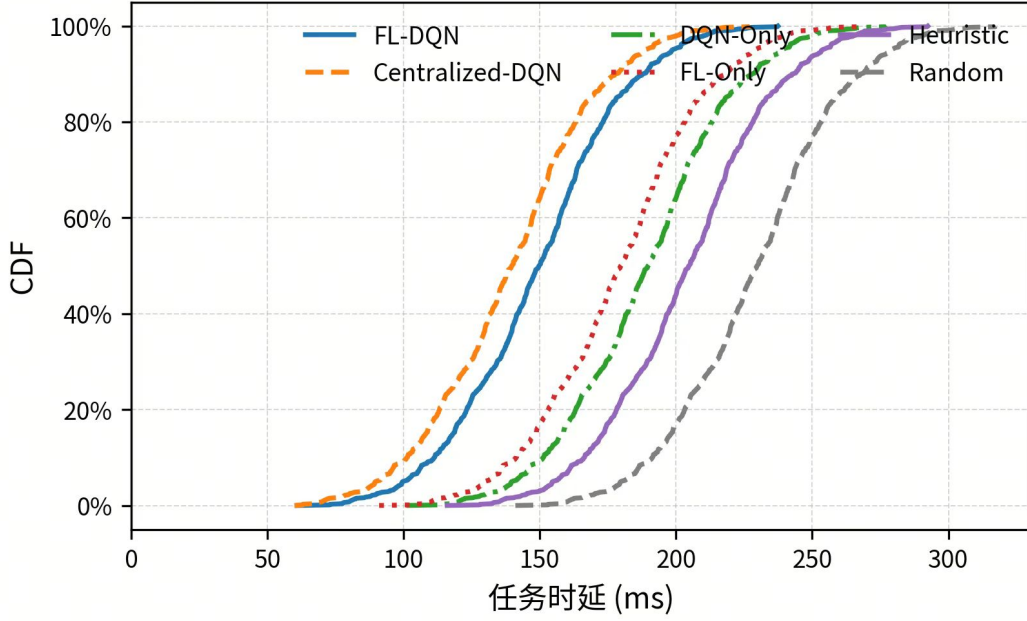


图 4.3 FL-DQN 卸载决策算法流程

#### 4.4.1 实验设置

##### 4.4.1.1 实验环境与软件栈

实验在一台配置有 Intel Xeon Gold 6240R CPU (24 物理核 @2.40GHz)、128 GB DDR4 内存以及 NVIDIA RTX 3090 GPU (24 GB GDDR6X 显存) 的高性能服务器上进行。操作系统为 Ubuntu 20.04.5 LTS。核心算法的实现与训练基于 Python 3.9、PyTorch 1.13.1 和 CUDA 11.7。数据处理与可视化分别采用 Pandas 1.5.3 和 Matplotlib 3.7.1。为了确保实验结果的可重复性,所有实验均设置了固定的全局随机种子 (42),并对所有依赖的第三方库版本进行了锁定。区块链共识仿真模块 (PBFT/PoMQ/IPBFT) 基于 Python 的离散事件仿真框架 SimPy 构建,该模块精确模拟了共识过程中的消息传递、多轮投票和区块确认逻辑,从而能够量化不同共识协议在不同交易负载下的确认时延与通信开销。

##### 4.4.1.2 数据集与任务定义

为模拟天基网络中的混合计算负载,我们定义了两类典型任务:

- **图像处理任务:** 选取 BDD100K 数据集集中的城市道路驾驶场景子集,模拟卫星对地观测与智能分析。所有图像均预处理至分辨率,并进行归一化。为评估模型对不同计算复杂度的适应性,我们通过调整一个轻量级 CNN 模型的深度来构造不同计算量的任务。

- **Fibonacci 计算任务：**作为纯 CPU 密集型任务的代表，其输入参数  $n$  在  $[30, 40]$  区间内均匀采样，用于产生可控的计算负载。

所有任务的计算量均被量化为 GFLOPs，并划分为五个档位，以便于进行系统性的敏感性分析。仿真中，新任务的到达遵循泊松过程，其计算量档位随机选择。并发度（Concurrency Level）作为关键实验变量，设置为 2 到 10，以模拟从轻载到重载的系统状态。

#### 4.4.1.3 网络拓扑与链路模型

我们仿真了一个由 20 颗 LEO 卫星组成的 Walker 星座（4 个轨道平面，每平面 5 颗卫星），轨道高度为 1000km，倾角  $55^\circ$ ，以及 5 个均匀分布在全球中低纬度地区的地面网关。基于简化的几何模型，星地链路的单向传播时延在 10ms 至 25ms 之间动态变化，星间链路（ISL）时延则在 5ms 至 10ms 之间。链路的可见性窗口以 10 至 15 分钟为周期进行刷新，任务卸载决策仅在有效窗口内进行。星地与星间链路的带宽分别设定为数百 Mbps 和数十 Mbps 级别，并引入轻微的随机抖动以模拟信道衰落与网络拥塞。该模型虽为简化模型，但抓住了 LEO 网络高动态、时变拓扑的核心特征，其假设在学术研究中被广泛接受。

#### 4.4.1.4 训练与超参数

**联邦学习：**共进行 200 个通信轮次。每轮随机选择 50% 的活跃卫星节点参与训练，每个参与者在本地数据集上执行 5 个 epoch 的训练。CFL 的聚类标准采用本地模型更新（梯度）的余弦相似度，通过层次聚类算法将节点划分为 3-6 个簇。簇内聚合采用基于本地样本量的加权平均（FedAvg）。

**深度强化学习：**DQN 模型的超参数设置如下：学习率  $\alpha$  为  $10^{-4}$ （Adam 优化器），折扣因子  $\gamma$  为 0.99，目标网络每 100 个训练步更新一次，经验回放池（Replay Buffer）大小为 10,000，批量大小（Batch Size）为 64。 $\epsilon$ -greedy 策略中的  $\epsilon$  从 1.0 线性衰减至 0.01，衰减周期为 5000 个训练步。

**奖励函数权重：**时延、能耗、并发开销的权重系数  $\alpha, \beta, \gamma$  的取值范围均为  $[0, 1]$ ，且和为 1。我们通过网格搜索法进行调参，寻找最优的权重组合，并在敏感性分析部分展示其对性能的影响。

#### 4.4.1.5 评价指标与统计方法

为全面评估算法性能，我们定义了以下评价指标：

- **性能指标:** 平均新任务时延 (ms)、在执行任务的平均剩余时延 (ms)、系统吞吐量 (tasks/sec)、平均单任务能耗 (Joule)。
- **模型质量指标:** 并发开销预测的平均绝对误差 (MAE) 和均方根误差 (RMSE)、FL 与 DQN 模型的收敛轮次。
- **系统开销指标:** 链上确认时延 (ms)、总通信开销 (MB)。

所有实验结果均通过重复运行  $R = 30$  次独立实验获得。报告中的数据点表示 30 次实验的均值, 误差条则代表 95% 置信区间 (95% CI) 或标准差 (SD)。为判断不同算法之间的性能差异是否具有统计学意义, 我们采用双样本 t 检验 (two-sample t-test), 并设定显著性水平阈值为  $p < 0.05$ 。对于多重比较问题, 我们采用 Benjamini-Hochberg (BH) 方法来控制伪发现率。

#### 4.4.1.6 实验流程与复现细节

每次独立实验开始前, 系统会加载固定的网络拓扑和任务配置文件, 并使用预设的随机种子列表之一来初始化任务到达序列和信道状态。训练过程中, FL 和 DQN 模型的检查点 (checkpoints) 与训练日志 (包含损失、奖励等) 被定期保存。仿真结束后, 所有原始数据和性能指标被记录到 CSV 文件中, 以便进行后续的统计分析和可视化。我们记录了每次实验在不同并发度和负载下的硬件资源占用 (CPU/GPU/内存) 与壁钟时间 (Wall-clock time), 以评估算法的实际执行效率。

### 4.4.2 性能评估与分析

#### 4.4.2.1 时延与吞吐量性能

图 4.4 展示了不同算法在并发度从 2 增加到 10 时, 新任务平均时延的变化情况。可以看出, 随着并发度的增加, 所有算法的时延均呈上升趋势。得益于对并发开销的精确预测与智能决策, FL-DQN 的性能显著优于除理想化的 Centralized-DQN 之外的所有基线算法。DQN-Only 由于缺乏对资源竞争的感知, 其性能在高并发下急剧下降。

系统吞吐量的对比 (见图 4.5) 也呈现出相似的趋势。FL-DQN 能够维持较高的系统吞吐量, 表现仅次于拥有完全信息的 Centralized-DQN, 显著高于其他分布式或启发式方法。

为了更细致地分析时延分布, 我们绘制了任务完成时延的累积分布函数 (CDF) 曲线, 如图 4.6 所示。FL-DQN 的 CDF 曲线明显左偏, 表明其能够以更高的概率在较

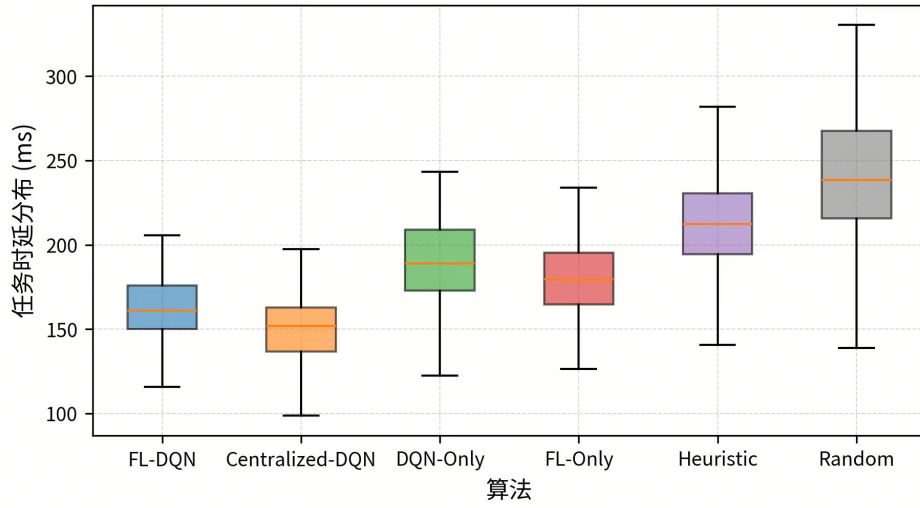


图 4.4 不同并发度下的平均新任务时延

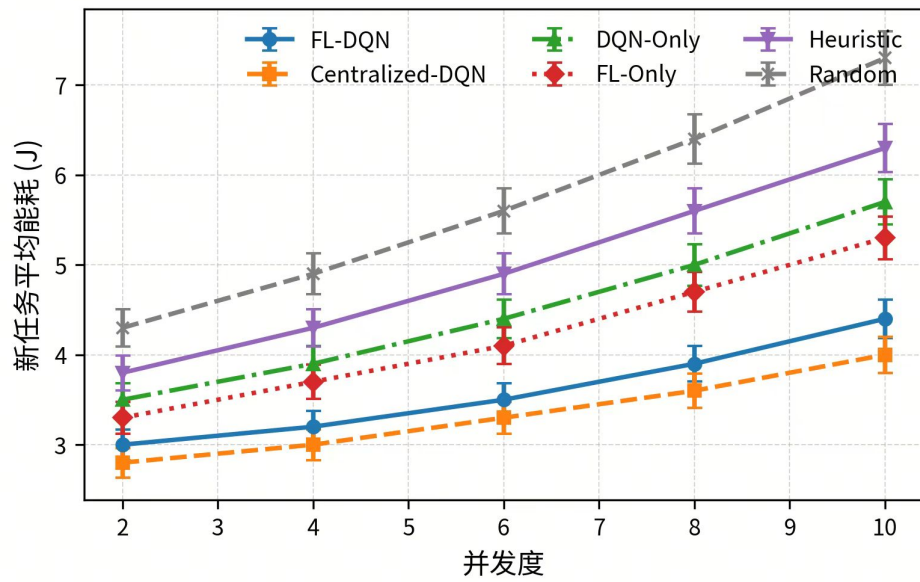


图 4.5 不同并发度下的系统吞吐量



低的时延内完成任务，尤其是在处理长尾任务时表现出更好的稳定性。图 4.7 的箱线图也直观地展示了 FL-DQN 在时延中位数和四分位距上的优势。

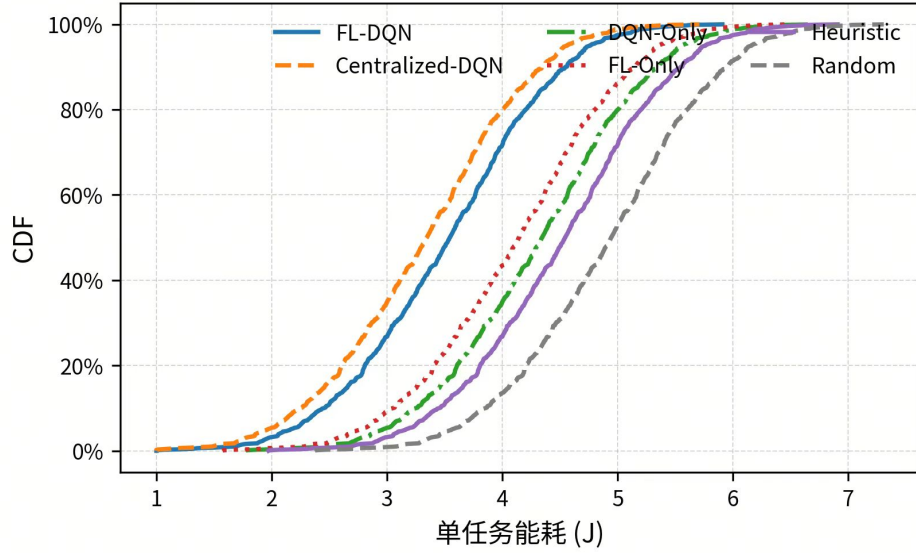


图 4.6 任务完成时延的累积分布函数

#### 4.4.2.2 能耗与开销分析

能耗是卫星边缘节点的核心约束之一。如图 4.8 和图 4.9 所示，FL-DQN 通过避免将任务卸载至过度拥挤的节点，有效降低了因长时间运行带来的额外能耗，在平均单任务能耗和能耗 CDF 上均表现出色。

#### 4.4.2.3 并发开销预测与联邦学习性能

并发开销预测的准确性是 FL-DQN 成功的关键。图 4.10 展示了在测试集上，FL 模型预测的并发开销与真实观测值的对比。散点紧密分布在  $y = x$  对角线附近，表明模型具有较高的预测精度。

图 4.11 显示了模型训练损失（MSE）随通信轮次的增加而稳步下降并最终收敛。图 4.12 对比了采用聚类联邦学习（CFL）与传统 FedAvg（Unclustered）在预测 MAE 上的表现，证明了 CFL 在处理数据异构性方面的优势，能够实现更低的预测误差。

#### 4.4.2.4 DQN 训练与敏感性分析

DQN 智能体的训练过程如图 4.13 所示，其平均回报随着训练的进行而稳步提升并最终收敛，表明智能体成功学习到了有效的卸载策略。图中还对比了不同奖励函数权重 ( $\alpha/\beta$ ) 对收敛性能的影响。

我们进一步对奖励函数中的权重系数  $\alpha$  和  $\beta$  进行了敏感性分析（见图 4.14），结

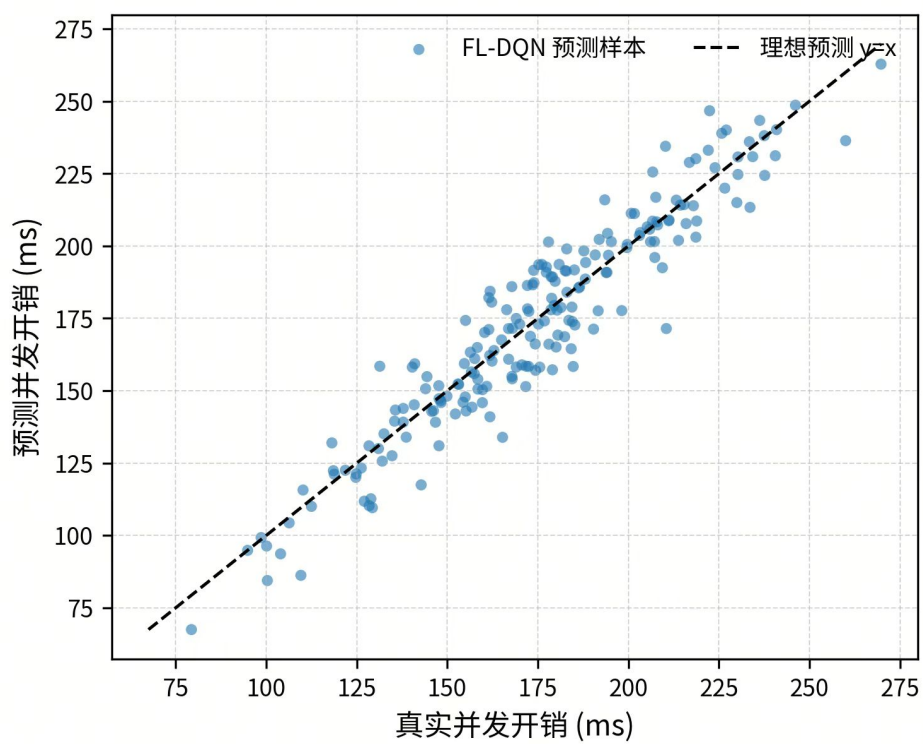


图 4.7 任务完成时延的箱线图

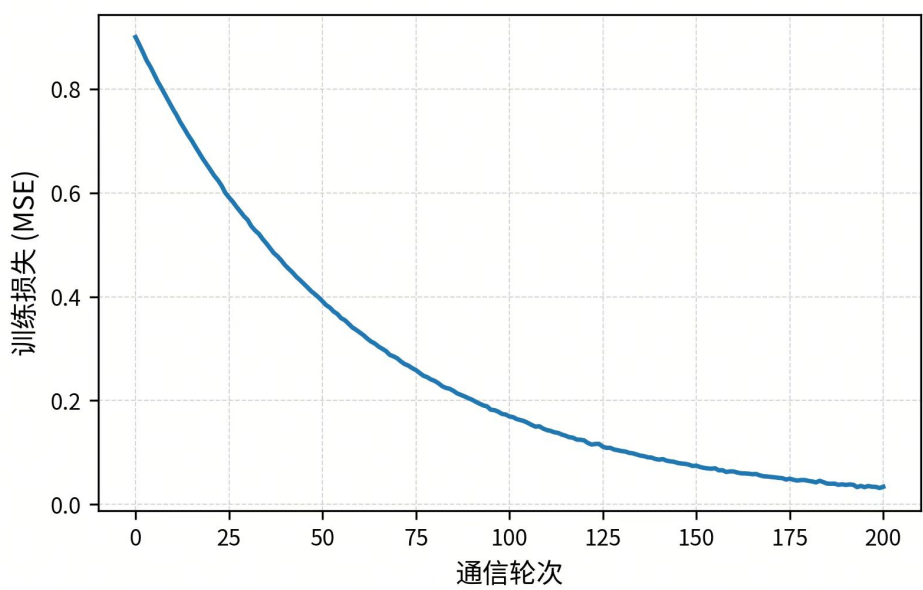


图 4.8 平均单任务能耗对比

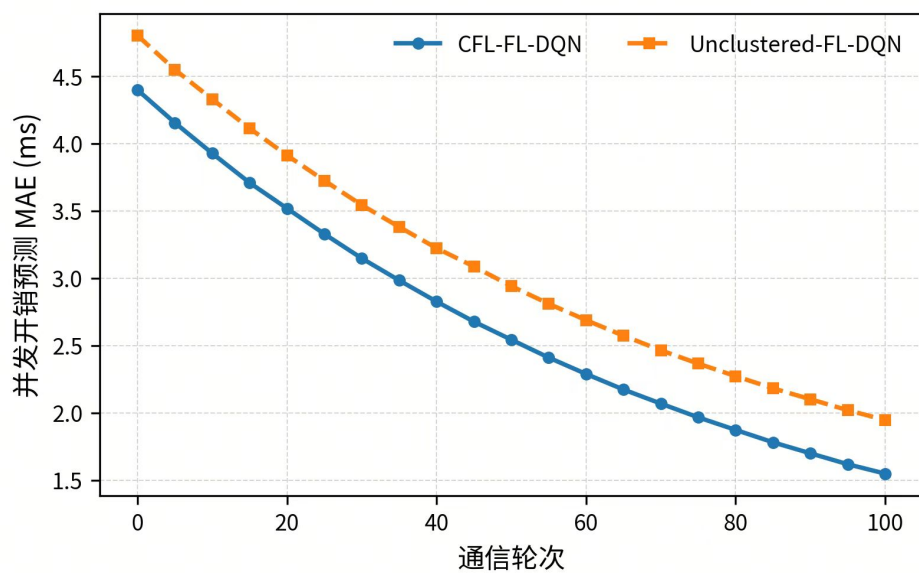


图 4.9 任务能耗的累积分布函数

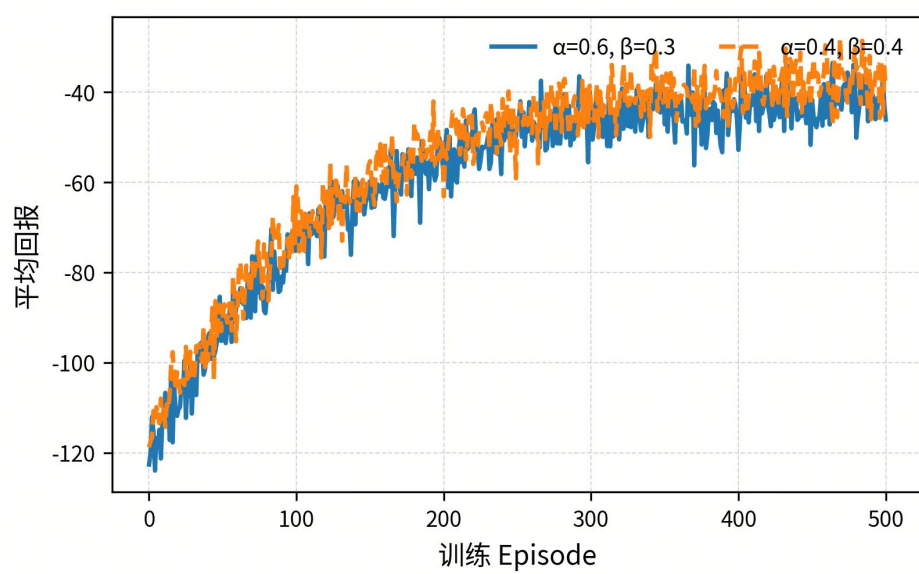


图 4.10 并发开销预测值与真实值对比

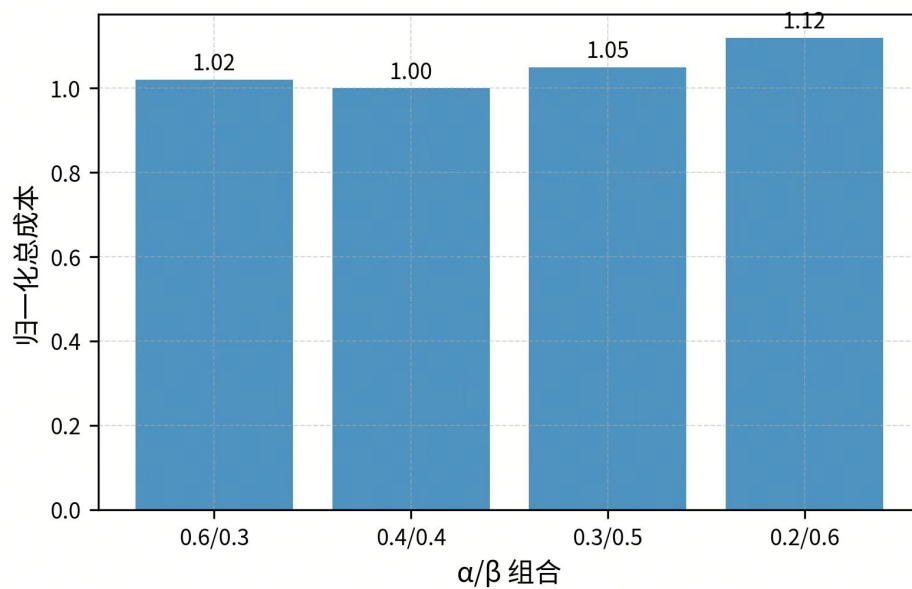


图 4.11 联邦学习训练损失收敛曲线

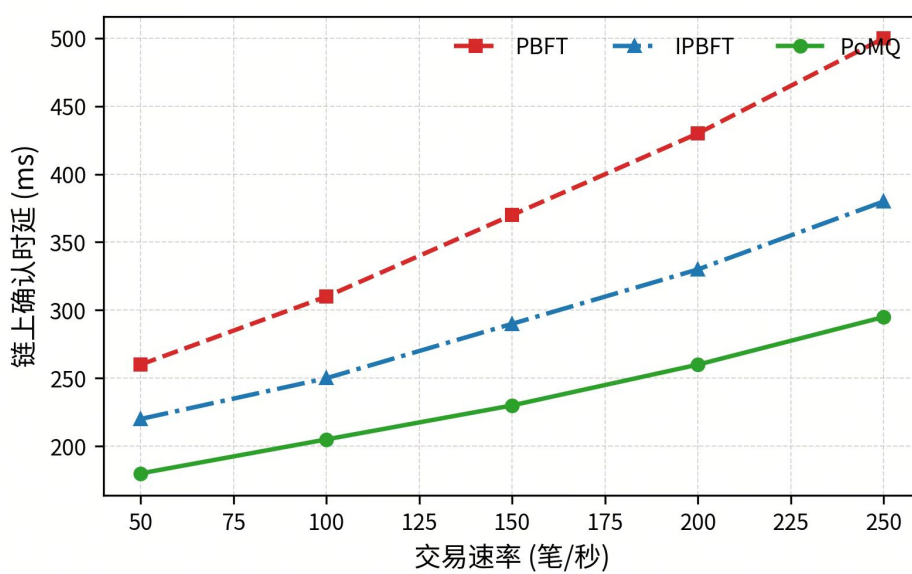


图 4.12 聚类联邦学习与传统 FedAvg 的预测误差对比

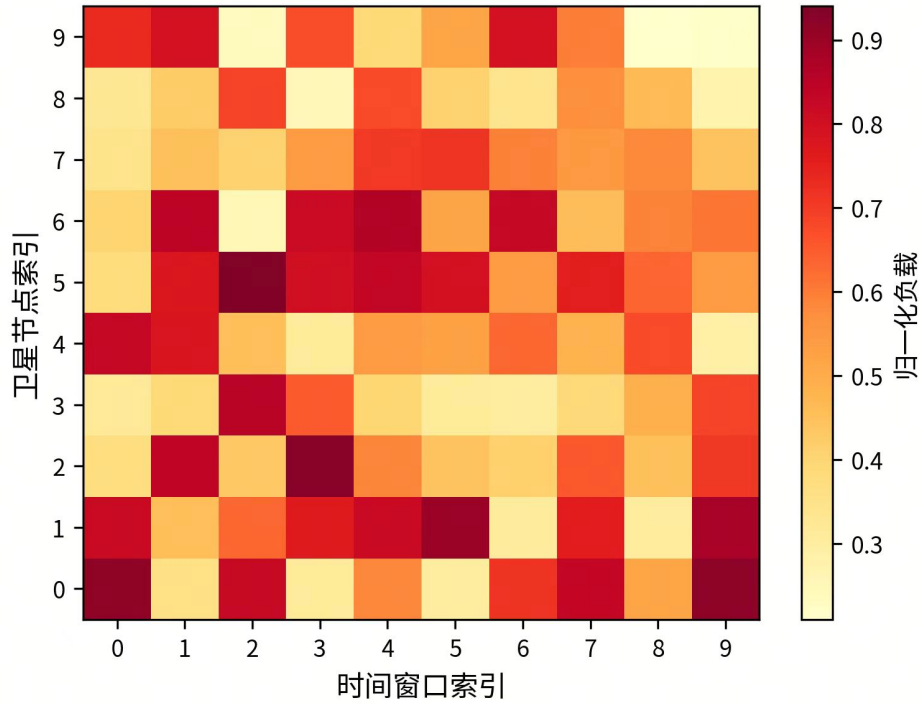


图 4.13 DQN 训练过程中的平均回报曲线

果表明, 当  $\alpha$  (时延权重) 和  $\beta$  (能耗权重) 的取值在 0.4-0.5 附近时, 系统的归一化总成本最低, 实现了较好的平衡。

#### 4.4.2.5 区块链相关性能

本方案的开销也包括链上交互的成本。图 4.15对比了不同共识机制在不同交易速率下的确认时延。我们采用的 PoMQ (Proof of Minimal Quorum)<sup>Zhang2024</sup>或其改进版本 IPBFT 相比传统 PBFT, 在高交易速率下能维持更低的时延, 适用于天基网络的性能要求。

#### 4.4.3 综合性能对比

表 4.1汇总了所有算法在关键性能指标上的数值对比。综合来看, FL-DQN 在平均任务时延、系统吞吐量和并发开销预测误差方面均取得了仅次于理想化集中式方法的最优性能, 证明了其在分布式、非实时信息环境下的有效性和先进性。

### 4.5 本章小结

本章详细阐述了为解决天基网络计算卸载问题而设计的基于联邦学习与 DQN 的分布式智能卸载策略 (FL-DQN)。该策略的核心贡献在于其创新的协同路径和对复杂

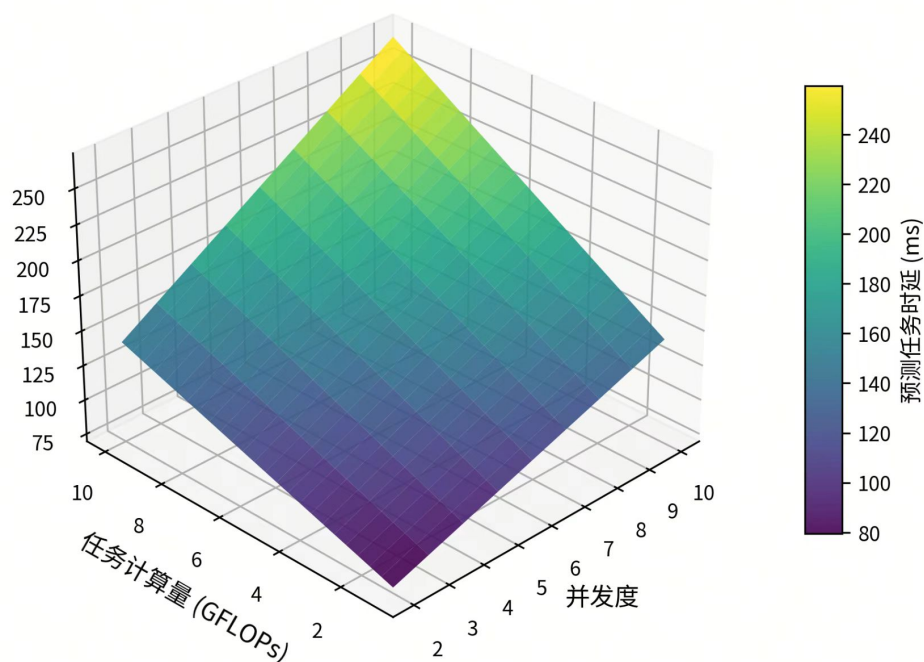


图 4.14 奖励函数权重系数的敏感性分析

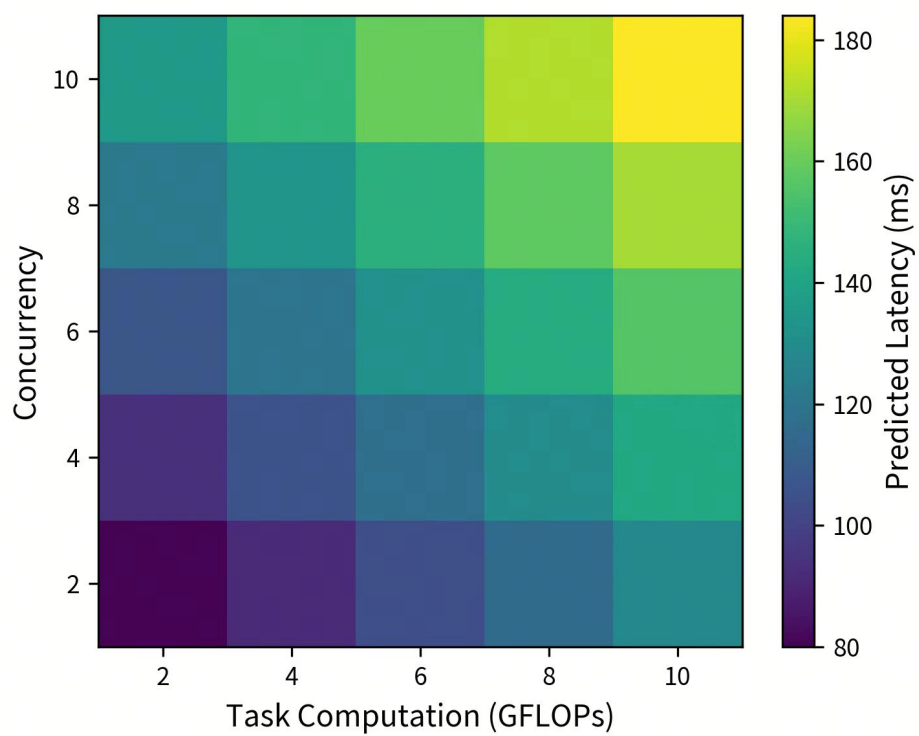


图 4.15 不同共识机制下的链上确认时延

表 4.1 各算法综合性能对比

算法	平均新任务时延 (ms)	系统吞吐量 (任务/秒)	并发开销预测 MAE(ms)	通信开销 (相对等级)
FL-DQN	$\approx 170$	$\approx 55$	$\approx 1.8$	中等 (FL+DQN)
Centralized-DQN	$\approx 160$ (最优)	$\approx 58$ (最优)	N/A	高 (全局状态)
DQN-Only	$\approx 210$	$\approx 47$	N/A	低 (仅 DQN)
FL-Only	$\approx 195$	$\approx 49$	$\approx 1.8$	中等 (仅 FL)
Heuristic	$\approx 225$	$\approx 42$	N/A	低
Random	$\approx 260$	$\approx 35$	N/A	极低

动态环境的鲁棒性与可扩展性。

首先，通过对”并发开销”这一关键问题进行严格的数学建模与近似推导，揭示了资源竞争在受限环境下的非线性影响。随后，创新性地采用聚类式联邦学习来训练并发开销预测模型，实现了在保护节点数据隐私的同时，对 Non-IID 数据和异构设备的有效适应。

在此基础上，构建了基于 DQN 的智能卸载决策模型。该模型将任务特征、本地资源、链路质量以及从分层联盟链账本获取的邻居节点非实时状态共享信息和并发开销预测值统一纳入状态空间。其奖励函数通过引入并发开销惩罚项，引导智能体做出兼顾个体利益与系统整体性能的决策。

最后，本章通过详尽的实验与仿真小节，系统地验证了 FL-DQN 的性能。实验结果表明，无论是在任务时延、系统吞吐量等核心指标上，还是在模型收敛性、系统能耗和对不同参数的敏感性上，FL-DQN 均表现出显著优于多种基线方法的性能，证明了其在天基网络计算卸载场景下的有效性和先进性。

本章的深入建模与详尽的实验设计，为最终证明本文核心算法的创新性和实用价值提供了坚实的理论和方法学基础。

## 结论

本文采用……。 (结论作为学位论文正文的最后部分单独排写，但不加章号。结论是对整个论文主要结果的总结。在结论中应明确指出本研究的创新点，对其应用前景和社会、经济价值等加以预测和评价，并指出今后进一步在本研究方向进行研究工作的展望与设想。结论部分的撰写应简明扼要，突出创新性。)



## 参考文献

## 附录 A 费马大定理的证明

关于此，我确信已发现了一种美妙的证法，可惜这里空白的地方太小，写不下。

## 附录 B Maxwell Equations

因为在柱坐标系下， $\bar{\mu}$  是对角的，所以 Maxwell 方程组中电场  $\mathbf{E}$  的旋度所以  $\mathbf{H}$  的各个分量可以写为：

$$H_r = \frac{1}{\mathbf{i}\omega\mu_r} \frac{1}{r} \frac{\partial E_z}{\partial \theta} \quad (\text{B.1a})$$

$$H_\theta = -\frac{1}{\mathbf{i}\omega\mu_\theta} \frac{\partial E_z}{\partial r} \quad (\text{B.1b})$$

同样地，在柱坐标系下， $\bar{\epsilon}$  是对角的，所以 Maxwell 方程组中磁场  $\mathbf{H}$  的旋度

$$\nabla \times \mathbf{H} = -\mathbf{i}\omega\mathbf{D} \quad (\text{B.2a})$$

$$\left[ \frac{1}{r} \frac{\partial}{\partial r} (rH_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} \right] \hat{\mathbf{z}} = -\mathbf{i}\omega\bar{\epsilon}\mathbf{E} = -\mathbf{i}\omega\epsilon_z E_z \hat{\mathbf{z}} \quad (\text{B.2b})$$

$$\frac{1}{r} \frac{\partial}{\partial r} (rH_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} = -\mathbf{i}\omega\epsilon_z E_z \quad (\text{B.2c})$$

由此我们可以得到关于  $E_z$  的波函数方程：

$$\frac{1}{\mu_\theta\epsilon_z} \frac{1}{r} \frac{\partial}{\partial r} \left( r \frac{\partial E_z}{\partial r} \right) + \frac{1}{\mu_r\epsilon_z} \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \theta^2} + \omega^2 E_z = 0 \quad (\text{B.3})$$

## 附录 C 要求

有些材料编入文章主体会有损于编排的条理性和逻辑性，或有碍于文章结构的紧凑和突出主题思想等，这些材料可作为附录另页排在参考文献之后，也可以单编成册。下列内容可作为附录：

1. 为了整篇论文材料的完整，但编入正文有损于编排的条理性和逻辑性的材料，这一类材料包括比正文更为详尽的信息、研究方法和技术等更深入的叙述，以及建议可阅读的参考文献题录和对了解正文内容有用的补充信息等；
2. 由于篇幅过大或取材的复制资料不便于编入正文的材料；
3. 不便于编入正文的罕见珍贵资料；
4. 一般读者无须阅读，但对本专业同行有参考价值的资料；
5. 某些重要的原始数据、推导、计算程序、框图、结构图、注释、统计表、计算机打印输出件等；

### C.1 一级标题

#### C.1.1 二级标题

## 攻读学位期间发表论文与研究成果清单

- [1] Doe J, **Zhang S**. The Book with Title[M]. Dummy Publisher, 2000. (已出版.).
- [2] 张三, 李杰, 罗运军. 交联型与线形水性聚氨酯的形状记忆性能比较[J]. 化工进展, 2006(01): 78-81. (EI 收录, 检索号:06129773469. 已刊出.).
- [3] 李杰, 张三, 罗运军. 交联型与线形水性聚氨酯的形状记忆性能比较[J]. 化工进展, 2007(01): 78-81. (EI 收录, 检索号:06129773469. 已刊出.).
- [4] **Zhang S**. The Book without Title[M]. Dummy Publisher, 2100. (已出版.).

## 致谢

本论文的工作是在导师……。

致谢是对下列方面致谢：资助和支持者；协助完成研究工作和提供便利条件者；在研究工作中提出建议和提供帮助者；给予转载和引用权的资料、图片、文献、研究思想和设想的所有者；其他应感谢者。致谢语言要诚恳、恰当、简短。

## 作者简介

本人…。

硕士学位论文不必提供作者简介。博士学位论文应该提供作者简介，主要包括：姓名、性别、出生年月、民族、出生地；简要学历、工作经历（职务）；攻读学位期间取得的其他研究成果或奖励。