

面向低轨卫星互联网的智能准入与联合调度-协作定位一体化系统

作者姓名

2025 年 9 月 2 日

目录

1	摘要 (中文)	3
2	Abstract (English)	3
3	第 1 章 绪论	3
3.1	1.1 研究背景与意义	3
3.2	1.2 问题提出	4
3.3	1.3 研究内容与贡献	4
3.4	1.4 论文组织结构	4
3.5	附：图表清单 (建议)	4
3.6	附：关键定义占位	4
4	第 2 章 相关工作	4
4.1	2.1 LEO 卫星网络资源管理	4
4.2	2.2 网络准入控制技术	5
4.3	2.3 深度强化学习在网络中的应用	5
4.4	2.4 QoE 优化与评估指标	5
4.5	2.5 研究空白与本文定位	5
4.6	2.6 LEO 融合定位最新进展 (会议论文, 近年)	5
5	第 3 章 系统总体设计	5
5.1	3.1 问题建模与分析	5
5.2	3.2 分层决策架构设计	5
5.3	3.3 系统功能模块	6
5.4	3.4 技术路线选择	6
5.5	3.5 本章小结	6
5.6	附：图表清单 (建议)	6
5.7	附：关键公式 (MDP/目标)	6
5.8	参考公式对齐	6
5.9	附：定位协同与特征注入 (新增)	7
6	第 4 章 基于 DRL 的智能准入控制	7
6.1	4.1 马尔可夫决策过程建模	7
6.2	4.2 时间感知状态空间设计	7
6.3	4.3 细粒度动作空间	7
6.4	4.4 QoE 驱动奖励函数 (公式)	7
6.5	4.5 PPO 算法实现与训练细节 (伪代码)	7
6.6	4.6 DSROQ 执行与接口 (合并)	8
6.7	附：图表清单 (建议)	8

6.8	参考公式对齐	8
7	第 5 章 协作定位 (匿名化参考 B)	8
7.1	5.1 概述与目标	8
7.2	5.2 指标与度量 (参考 B)	8
7.3	5.3 协作提示与波束调度 (匿名化)	8
7.4	5.4 与策略层 (第 4 章) 的融合	9
7.5	5.5 与执行层 (第 4 章-DSROQ 合并节) 的融合	9
7.6	5.6 实验设计要点	9
7.7	5.7 图表建议	9
7.8	5.8 匿名化公式片段 (通用形态)	9
7.9	参考公式对齐	9
8	第 6 章 实验设计与结果分析	9
8.1	6.1 实验环境搭建	9
8.2	6.2 数据集与参数设置	10
8.3	6.3 基线方法对比	10
8.4	6.4 性能指标与统计方法	10
8.5	6.5 多场景评估结果	10
8.6	6.6 消融与敏感性	10
8.7	6.7 本章小结	10
8.8	附: 指标定义 (公式)	10
8.9	附: 统计方法	11
8.10	附: 图表清单 (建议)	11
8.11	6.8 端到端演示流程 (含可视化联动)	11
9	第 7 章 系统实现与部署	12
9.1	7.1 后端核心实现	12
9.2	7.2 前端可视化	12
9.3	7.3 部署与运维	12
9.4	7.4 测试与性能	12
9.5	7.5 本章小结	12
9.6	附: 实现伪代码	12
9.7	附: 图表清单 (建议)	13
9.8	附: 定位 API 示例 (Flask)	13
9.9	附: 前端联动示例 (Vue 风格伪代码)	13
9.10	参考公式对齐	14
10	第 8 章 消融实验与参数分析	14
10.1	8.1 状态空间消融	14
10.2	8.2 动作空间消融	14
10.3	8.3 奖励函数消融	14
10.4	8.4 超参数敏感性	14
10.5	8.5 本章小结	14
10.6	附: 消融实验流程 (伪代码)	14
10.7	附: 图表清单 (建议)	15
10.8	附: 图表生成命令示例 (定位消融)	15
11	第 9 章 结论与展望	15
11.1	9.1 研究工作总结	15
11.2	9.2 主要贡献	15
11.3	9.3 不足与局限	15
11.4	9.4 未来工作方向	15
11.5	附: 图表清单 (建议)	16
11.6	附: 未来路线图 (占位)	16

1 摘要 (中文)

本研究面向动态拓扑与资源受限的低轨 (LEO) 卫星网络, 提出一种以用户体验 (QoE) 为目标的智能准入控制与系统化实现框架。方法层面, 上层采用深度强化学习 (DRL) 作为“守门员”, 设计时间感知状态空间与细粒度动作 (接受/拒绝/降级/延迟/部分), 构建由 QoE 变化、公平性与网络效率等组成的复合奖励以优化长期收益; 下层集成 DSROQ, 实现对路由、带宽与调度的联合分配与执行。系统层面, 构建基于 Hypatia 的端到端工程平台, 并引入融合/协作定位模块: 以 CRLB/GDOP 等定位质量指标为特征, 结合波束调度提示 (Beam Hint) 与可见波束/协作卫星统计, 将定位质量约束与网络决策协同, 使准入-分配与定位可用性形成闭环优化。我们在多星座规模、负载水平、动态与故障场景下, 对比阈值、负载、启发式、传统机器学习及其他 DRL 方法。结果表明, 所提方法在平均 QoE、准入率、Jain 公平性、网络利用率与定位可用性等指标上取得显著提升, 同时保持决策延迟可控; 消融与敏感性分析进一步验证了时间信息、奖励设计与定位特征对性能的关键贡献。研究成果具备可复现性与工程落地性, 为 LEO 网络的 QoE 优化、定位协同与资源一体化管理提供了新思路。

关键词: LEO 卫星网络; 准入控制; 深度强化学习; QoE; DSROQ; Hypatia; 协作定位; CRLB; GDOP

2 Abstract (English)

We propose a QoE-driven admission control and system framework for low Earth orbit (LEO) satellite networks featuring dynamic topology and constrained resources. A deep reinforcement learning (DRL) agent serves as a top-layer gatekeeper with a time-aware state space and a fine-grained action set (accept/reject/degrade/delay/partial). A composite reward captures QoE improvement, fairness, and network efficiency for long-term optimization, while the lower layer integrates DSROQ to jointly perform routing, bandwidth allocation, and scheduling. At the system level, we build an end-to-end platform on Hypatia and incorporate a fusion/cooperative positioning module: positioning quality (CRLB/GDOP), beam scheduling hints, and visible beams/cooperative satellites are injected as features and constraints to jointly optimize admission-allocation decisions and positioning availability. Across diverse constellations, loads, dynamics, and failures, our method outperforms threshold-, load-, heuristic-, classical ML-, and other DRL-based baselines in average QoE, admission rate, Jain's fairness, network utilization, and positioning availability, with controllable decision latency. Ablation and sensitivity analyses confirm the key roles of temporal information, reward design, and positioning features. The framework is reproducible and engineering-ready, offering a unified perspective on QoE optimization, positioning cooperation, and resource management in LEO networks.

Keywords: LEO satellite networks; Admission control; Deep reinforcement learning; QoE; DSROQ; Hypatia; Cooperative positioning; CRLB; GDOP

3 第 1 章 绪论

3.1 1.1 研究背景与意义

低轨卫星 (LEO) 网络具备覆盖广、时延低、容量高等优势, 但其动态拓扑与有限资源使端到端服务质量与用户体验 (QoE) 保障极具挑战。现有研究多聚焦于既有业务的路由、带宽与调度优化, 缺少对“是否接纳新流量”的系统性决策, 容易在高负载下诱发拥塞并损害全局 QoE。

本文面向 LEO 网络提出一套以 QoE 为目标的智能准入与联合调度-协作定位一体化系统: 上层基于 DRL 实现准入决策与联合调度策略学习, 下层由 DSROQ 完成路由/带宽/调度的联合分配, 同时引入协作定位模

块 (CRLB/GDOP、可见波束/协作卫星、Beam Hint) 作为状态与约束, 使“定位质量—准入/分配—网络运行”形成闭环优化。

3.2 1.2 问题提出

- 传统准入控制 (阈值/负载/启发式) 难以适应动态拓扑与复杂业务需求, 缺乏对全局 QoE 的长期最优考虑。
- DSROQ 仅对已接纳流量执行联合路由/带宽/调度优化, 未显式建模新流量的准入决策, 可能导致“过度接纳”。
- 需要一个上层“守门员”在新流到达时依据网络状态与 QoE 影响做出接纳/拒绝/降级/延迟等细粒度决策。

3.3 1.3 研究内容与贡献

- 构建“准入—分配”的分层决策架构: 上层 DRL 准入与联合调度, 下层 DSROQ 资源分配; 引入协作定位模块并将其指标注入状态/奖励与资源约束。
- 设计时间感知状态空间、细粒度动作空间与 QoE 驱动奖励函数, 同时新增“定位可用性/质量”项, 实现 QoE 与定位双目标的长期优化。
- 基于 Hypatia 搭建高保真仿真与可视化平台, 完成端到端工程实现; 提供图表与脚本, 保证可复现。
- 在 QoE、准入率、公平性、网络利用率与定位可用性等指标上取得显著改进, 并通过消融与敏感性分析验证关键组件的必要性。

3.4 1.4 论文组织结构

全文共九章: 第 2 章综述; 第 3 章系统设计; 第 4 章 DRL 准入; 第 5 章 DSROQ 集成; 第 6 章实验; 第 7 章实现与部署; 第 8 章消融与参数分析; 第 9 章总结与展望。

[图占位] 研究路线图

[参考占位] 近 5 年核心文献与数据来源

3.5 附: 图表清单 (建议)

- 研究动机与问题示意图
- 本文技术路线图 (问题 → 方法 → 实现 → 实验 → 结论)
- 贡献对比表 (与代表性工作在目标/方法/系统/实验维度对比)

3.6 附: 关键定义占位

- QoE/QoS、EF/AF/BE 分类与含义
- LEO 动态拓扑、ISL/GL 链路、星地/星间参数
- 本文术语表 (统一符号)

4 第 2 章 相关工作

4.1 2.1 LEO 卫星网络资源管理

综述近 5 年在星座建模、路由、带宽分配与调度方面的关键进展, 强调动态拓扑与链路可用性变化下的资源协调难题。

4.2 2.2 网络准入控制技术

- 传统：固定/动态阈值、负载感知、带宽预留与启发式策略。
- 机器学习：SVM/随机森林/浅层 NN 等分类式准入。
- 强化学习：DQN/A3C/DDPG/PPO 等在无线与有线网络的应用与迁移性。

4.3 2.3 深度强化学习在网络中的应用

总结 DRL 在路由、调度、资源分配与拥塞控制的成功案例与挑战（样本效率、稳定性、泛化、约束处理）。

4.4 2.4 QoE 优化与评估指标

QoE 建模方法、QoS 与 QoE 的映射关系、Jain 公平性、端到端时延/丢包/吞吐等综合指标体系。

4.5 2.5 研究空白与本文定位

LEO 网络实时准入控制仍缺少面向 QoE 的分层优化框架与可复现实验。本文以 DRL 为上层“守门员”，结合 DSROQ 形成端到端的准入-分配协同。

[表占位] 方法对比表（类别/代表/优势/局限/适用场景）

[参考占位] 参考文献与数据集链接

4.6 2.6 LEO 融合定位最新进展（会议论文，近年）

- 近年会议工作普遍采用“定位导向的联合设计”范式：以 CRLB/TDOA 为目标，联合优化波束调度与波束成形，分层求解（外层启发式调度、内层凸松弛/SDR 成形），并通过协作与干扰抑制提升定位几何与测量质量；在中等功率条件下，平均定位精度相较分离式方案有显著提升（约 10%—60% 量级，依具体设置而定）。
- 启示：联合优化与分层架构可推广至通信-定位一体化与网络资源决策问题。

[引用原则] 引用允许在参考列表中以匿名化条目或编号出现，但正文不直接出现具体论文题名与作者名称，仅描述方法脉络与共性结论。

5 第 3 章 系统总体设计

5.1 3.1 问题建模与分析

- 决策时刻：新流到达事件驱动；时间步与窗口协同。
- MDP 抽象：状态 S （时间感知 + 网络与 QoE 特征）、动作 A （接纳/拒绝/降级/延迟/部分）、奖励 R （QoE 变化 + 公平与违规惩罚 + 效率奖励），折扣 γ 。
- 目标：最大化长期加权 QoE，满足关键 QoS 约束并兼顾公平性。

5.2 3.2 分层决策架构设计

上层 DRL 准入控制为“是否/如何接纳”做出策略性决策，下层 DSROQ 执行联合路由、带宽分配与调度，实现资源层面的最优化与时序同步。

flowchart TD

A[新流到达] --> B[状态提取/预测]

B --> C{DRL 准入控制}

C -- 接受/降级/延迟/部分 --> D[DSROQ 联合优化]

C -- 拒绝 --> H[丢弃/排队]
D --> E[网络执行/调度]
E --> F[QoE 与 KPI 评估]
F --> G[策略更新]

5.3 3.3 系统功能模块

- 状态/动作/奖励子系统；预测与趋势提取。
- DSROQ 接口：MCTS 路由、李雅普诺夫调度、资源重分配触发。
- 仿真环境：Hypatia (satgenpy/ns3-sat-sim) 封装与统一 API。
- 可视化与监控：Cesium 3D、指标看板、日志与追踪。

5.4 3.4 技术路线选择

- 后端：Python + Hypatia + PyTorch + Stable-Baselines3 + SimPy
- 前端：Vue.js + CesiumJS + ECharts
- API：Flask (REST)
- 部署：Docker/K8s；监控：Prometheus/Grafana；追踪：MLflow

5.5 3.5 本章小结

阐明了以 DRL 为上层、以 DSROQ 为下层的分层协同框架，并明确了工程与实验的支撑技术路线。

5.6 附：图表清单（建议）

- 分层决策总体架构图 (Mermaid/TikZ)
- 数据流与接口时序图 (准入 → 分配 → 执行 → 评估 → 学习)
- 状态/动作/奖励设计示意图

5.7 附：关键公式 (MDP/目标)

- MDP 五元组： $\mathcal{M} = (\mathcal{S}, \mathcal{A}, P, R, \gamma)$
- 折扣回报目标：

$$\max_{\pi} \mathbb{E}_{\pi} \left[\sum_{t=0}^{\infty} \gamma^t r_t \right]$$

- 奖励一般形态 (详见第 4 章)：

$$r_t = w_1 \Delta \text{QoE}_t + w_2 \text{Fair}_t + w_3 \text{Eff}_t - w_4 \text{Viol}_t - w_5 \text{DelayPen}_t$$

- 约束与拉格朗日松弛：

$$\min_{\lambda \geq 0} \max_{\pi} \mathbb{E} \left[\sum_t \gamma^t (r_t - \sum_k \lambda_k g_k(s_t, a_t)) \right]$$

5.8 参考公式对齐

详见 docs/reference/formula_alignment.md。

5.9 附：定位协同与特征注入（新增）

- 状态特征：CRLB/GDOP、可见波束数、协作卫星数、平均/最小 SINR、波束调度提示（Beam Hint）。
- 接口对接：Hypatia→Positioning 模块产出定位质量向量，经 State Extractor 归一化后拼接进 DRL 状态；同时在 Evaluation 汇总“定位可用性”等指标。
- 协同策略：在资源紧张时优先保证高定位需求业务的可用性；当定位质量退化时，触发保守接纳或延迟接纳策略，与 DSROQ 重分配联动。

6 第 4 章 基于 DRL 的智能准入控制

6.1 4.1 马尔可夫决策过程建模

- 状态 S：链路利用率统计、EF/AF/BE 流量状态、QoE 与 QoS 违规率、轨道相位、拓扑变化率、容量/负载预测、历史趋势、新流请求属性（类型/带宽/时延/位置/时长），以及定位相关特征（CRLB/GDOP、可见波束数、协作卫星数、平均/最小 SINR、Beam Hint）。
- 动作 A：拒绝、接受、降级接受、延迟接受、部分接受。
- 奖励 R：QoE 变化为主；公平性奖励与违规惩罚；效率奖励；定位可用性与质量提升奖励；权重与归一化策略。

6.2 4.2 时间感知状态空间设计

包含当前与短期历史/预测特征，反映动态拓扑与业务演化；采用滑动窗口与轻量预测器提升稳定性与泛化。

6.3 4.3 细粒度动作空间

- 接受：直接进入下层 DSROQ。
- 降级接受：放宽部分 QoS 或降低带宽；形成退化请求再次评估。
- 延迟接受：排队等待下一个时窗（附带延迟惩罚）。
- 部分接受：按比例下调带宽/速率上限。

6.4 4.4 QoE 驱动奖励函数（公式）

- 定义： $\Delta QoE_t = QoE_t^{\text{after}} - QoE_t^{\text{before}}$
- 公平性 (Jain)： $J(\mathbf{x}) = \frac{(\sum_i x_i)^2}{n \sum_i x_i^2}$
- 定位可用性： $A_t^{\text{pos}} \in [0, 1]$ （基于可见波束/协作卫星/CRLB 阈值综合打分）
- 奖励：

$$r_t = w_1 \Delta QoE_t + w_2 J(\mathbf{x}_t) + w_3 \text{Util}_t + w_4 A_t^{\text{pos}} - w_5 \text{Viol}_t - w_6 \text{DelayPen}_t$$

6.5 4.5 PPO 算法实现与训练细节（伪代码）

伪代码：PPO with HypatiaAdmissionEnv（含定位特征）

obs = env.reset()

while not done:

 action = agent.policy(obs)

 next_obs, reward, done, info = env.step(action)

 # obs 包含: 网络 + QoE + (CRLB, GDOP, visible_beams, coop_sats, beam_hint, SINR)

 buffer.add(obs, action, reward, next_obs, done)

 obs = next_obs

agent.update(buffer)

6.6 4.6 DSROQ 执行与接口（合并）

- 接口与流程：
 - 准入动作（接受/降级/部分/延迟）→ 触发路由与带宽联合分配 → 李雅普诺夫风格调度执行。
 - 核心接口：`route, bw = route_and_allocate(flow, topology, link_caps); apply_schedule(current_flows)`。
 - 代价整合与约束（匿名化对齐参考 A/B）：
 - 代价： $C = C_{net} + \lambda_{pos} \Phi(\text{CRLB}, \text{GDOP}, \text{visible_beams}, \text{coop_sats})$
 - 约束：最小可见波束/协作卫星下限，CRLB 阈值过滤；冷却时间限制跨束/跨链路迁移频率以维持定位几何。
 - 重分配触发与稳定性：
 - 触发门限与冷却时间避免频繁全局重分配；优先增量调整与缓存复用以降低震荡。
 - 输出与评估：
 - 产出路由与带宽方案，统计定位几何保持率、CRLB/GDOP 分布变化、 A^{pos} 与 QoE 的联合改进。
-

6.7 附：图表清单（建议）

- 奖励函数分解图与权重敏感性曲线（含定位权重）
 - 定位特征重要性/消融对比图
 - 训练曲线（回报、损失、熵、价值函数误差）
 - 准入-执行联动时序与代价/约束示意图
-

6.8 参考公式对齐

详见 docs/reference/formula_alignment.md。

7 第 5 章 协作定位（匿名化参考 B）

7.1 5.1 概述与目标

- 目标：在 LEO 多波束/多卫星协作环境中提升定位质量与可用性（以 CRLB/GDOP 为核心指标），同时与网络准入-分配联动，形成系统级协同。
- 思路：以协作定位特征与提示（Beam Hint/协作集、可见波束/协作卫星统计）驱动上层策略与下层执行，贯穿状态、奖励、代价与约束。

7.2 5.2 指标与度量（参考 B）

- 定位界限：CRLB（均值/p95）、GDOP（均值/p95）。
- 可用性： $A^{pos} \in [0, 1]$ ，由（可见波束、协作卫星、CRLB 阈值）综合打分获得。
- 协作度：每用户平均可见波束数、协作卫星数；协作集大小分布。

7.3 5.3 协作提示与波束调度（匿名化）

- Beam Hint：基于可见性/相关性与功率/干扰约束生成的“推荐波束/卫星集合”。
- 外层调度建议：优先调度相关性较低的用户/波束组合以抑制干扰，形成更优几何。

7.4 5.4 与策略层（第 4 章）的融合

- 状态注入：在 s_t 中引入 (CRLB、GDOP、visible beams、coop sats、SINR、Beam Hint)。
- 奖励扩展：加入 A_t^{pos} 项，兼顾 QoE 与定位双目标。

7.5 5.5 与执行层（第 4 章-DSROQ 合并节）的融合

- 代价整合： $C = C_{net} + \lambda_{pos} \Phi(\text{CRLB}, \text{GDOP}, \text{visible_beams}, \text{coop_sats})$ 。
- 可行域约束：CRLB 阈值、最小可见波束/协作卫星、迁移冷却时间以维持定位几何。

7.6 5.6 实验设计要点

- 指标：CRLB/GDOP 分布、 A^{pos} 、协作集统计，与 QoE/AR/Util 等网络指标联合呈现。
- 消融：移除 (CRLB/GDOP/协作统计/Beam Hint) 观察对定位与网络双目标的影响。
- 敏感性：对 λ_{pos} 与阈值 (τ, b_{min}, s_{min}) 做曲线/热力图分析。

7.7 5.7 图表建议

- CRLB/GDOP 分布图 (均值/p95、CDF/箱线图)。
- 可用性雷达图与协作柱状图 (visible beams/coop sats)。
- Beam Hint 可视化叠加 (Cesium 层)。

7.8 5.8 匿名化公式片段（通用形态）

- 观测模型 (抽象)： $\mathbf{y} = \mathbf{h}(\mathbf{x}) + \mathbf{n}$ ，其中 $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$ 。
- Fisher 信息矩阵 (线性化近似)： $\mathbf{J}(\mathbf{x}) = \mathbf{H}^\top \mathbf{R}^{-1} \mathbf{H}$ ， $\mathbf{H} = \partial \mathbf{h} / \partial \mathbf{x}$ 。
- CRLB (位置估计下界)： $\text{Cov}(\hat{\mathbf{x}}) \succeq \mathbf{J}^{-1}$ ，可报告 $\text{tr}(\mathbf{J}^{-1})$ 、主对角线或其分位数 (如 p95)。
- GDOP (几何精度因子，归一化)： $\text{GDOP} = \sqrt{\text{tr}((\mathbf{H}^\top \mathbf{H})^{-1})}$ (在等噪声、等功率假设下的常用度量)。
- 可用性 (本稿定义)： $A^{pos} = f(\text{visible beams}, \text{coop sats}, \mathbb{1}[\text{tr}(\mathbf{J}^{-1}) \leq \tau]) \in [0, 1]$ 。

说明：以上为通用形式，仅用于与系统设计对齐；不涉及具体论文题名/作者，详细对齐见 docs/reference/formula_alignment.md。

7.9 参考公式对齐

详见 docs/reference/formula_alignment.md (参考 B)。

8 第 6 章 实验设计与结果分析

8.1 6.1 实验环境搭建

- Hypatia 配置：星座、时间步、ISL、路由选项与仿真时长。
- 计算环境：CPU/GPU、并行评估、随机种子管理、日志与追踪。

8.2 6.2 数据集与参数设置

- 星座与场景：小/中/大规模；静态/慢变/快变；故障类型。
- 流量与 QoS：泊松/突发/周期/混合；EF/AF/BE 配置；ITU-T 参考。
- DRL 超参： γ 、 λ 、学习率、batch、n_steps、clip、epochs 等。

8.3 6.3 基线方法对比

阈值/负载/启发式/ML (SVM/RF/NN) /DRL (DQN/A3C/DDPG/PPO)、无准入与“完美准入”参考。

8.4 6.4 性能指标与统计方法

QoE 均值/分布、公平性指数、准入/拒绝/降级/延迟率、网络利用率、时延/丢包/吞吐、计算复杂度与决策时延；置信区间与显著性检验。

新增定位相关指标：– 定位可用性 $A^{pos} \in [0, 1]$ ：由可见波束/协作卫星/CRLB 阈值综合打分；– 定位质量：CRLB 分布（均值/p95）、GDOP 分布；– 协作度：每用户平均协作卫星数、平均可见波束数。

8.5 6.5 多场景评估结果

- 规模/负载/动态性/故障/QoS 分布的组合场景实验与结果分析。
- 稳定性与鲁棒性：对参数扰动与故障的敏感性评估。

8.6 6.6 消融与敏感性

- 状态/动作/奖励/结构/训练策略消融。
- 超参数敏感性（曲线/热力图）。
- 定位特征消融：去掉（CRLB/GDOP/协作统计/Beam Hint）对子指标与联合指标的影响。

8.7 6.7 本章小结

总结在多维度指标与场景下的优势、局限与可推广性。

8.8 附：指标定义（公式）

- 平均 QoE: $\bar{Q} = \frac{\sum_k w_k Q_k}{\sum_k w_k}$
- 准入率: $AR = \frac{N_{accept}}{N_{requests}}$, 拒绝率: $RR = 1 - AR$
- 降级/延迟率: 同理定义为相应动作发生次数占比
- QoS 违规率: $Viol = \frac{\sum_f \mathbb{1}[QoS_f \text{ 违反}]}{N_{active}}$
- Jain 公平性: $J(\mathbf{x}) = \frac{(\sum_i x_i)^2}{n \sum_i x_i^2}$
- 吞吐量: $Thr = \frac{\sum \text{bits delivered}}{\Delta t}$
- 平均时延: $\bar{T} = \frac{1}{N} \sum_i T_i$; 丢包率: $PLR = \frac{N_{lost}}{N_{sent}}$
- 网络利用率: $Util = \frac{\text{used capacity}}{\text{total capacity}}$
- 定位可用性: $A^{pos} = f(\text{visible beams, coop sats}, \mathbb{1}[CRLB \leq \tau])$

8.9 附：统计方法

- 95% 置信区间: $\bar{x} \pm t_{\alpha/2, \nu} \cdot \frac{s}{\sqrt{n}}$
- 两独立样本 t 检验: $t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{s_1^2/n_1 + s_2^2/n_2}}$
- 多组比较: ANOVA/非参检验; 必要时 Holm-Bonferroni 校正

8.10 附：图表清单（建议）

- 综合对比: QoE/AR/Jain/Util/Delay/PLR/Thr
 - 定位相关: CRLB/GDOP 分布、可用性雷达图、协作统计柱状图
 - CDF/箱线图: QoE 分布、时延分布、CRLB 分布
 - 热力图: 状态特征重要性、敏感性曲面
 - 雷达图: 多指标综合表现
-

8.11 6.8 端到端演示流程（含可视化联动）

- 1) 启动后端 API（默认端口 5000）:

```
python src/api/main.py
```

- 2) 运行仿真/训练或回放（示例）:

```
# 训练（示意）
```

```
python src/admission/train.py --config experiments/configs/drl_params.yaml
```

```
# 或回放/评测脚本
```

```
python experiments/run_baselines.py
```

- 3) 前端/可视化联动（示例调用）:

```
# 获取定位指标
```

```
curl "http://127.0.0.1:5000/api/positioning/metrics?time=0&users=[]"
```

```
# 获取波束调度提示
```

```
curl -X POST "http://127.0.0.1:5000/api/positioning/beam_hint" \
  -H "Content-Type: application/json" \
  -d '{"time":0, "users":[], "budget":{}}'
```

- 4) 生成图表（示例数据见 experiments/results/）:

```
# QoE 对比
```

```
python scripts/plots/qoe_metrics_plot.py \
  --input experiments/results/positioning_ablation_example.json \
  --output docs/assets/ablation_qoe_positioning.png
```

```
# 准入/拒绝/降级 速率对比
```

```
python scripts/plots/admission_rates_plot.py \
  --input experiments/results/positioning_ablation_example.json \
  --output docs/assets/ablation_rates_positioning.png
```

```
# 敏感性热力图
```

```
python scripts/plots/fairness_heatmap.py \
  --input experiments/results/sensitivity_matrix.json \
  --output docs/assets/ablation_sensitivity_positioning.png
```

- 5) 生成论文 PDF（Pandoc + XeLaTeX）:

```
bash docs/latex/build.sh
# 输出: docs/latex/thesis.pdf
```

9 第 7 章 系统实现与部署

9.1 7.1 后端核心实现

- Hypatia 封装与统一接口：状态提取、拓扑/链路查询、仿真控制。
- DRL Agent：PPO 实现、训练/推理接口与并行评估。
- DSROQ 适配：MCTS 路由、李雅普诺夫调度的触发与数据结构。
- API：Flask REST，仿真控制、状态查询与决策下发。

9.2 7.2 前端可视化

- Cesium 3D：星座与链路、拓扑变化动画、决策可视化。
- 指标看板：QoE、准入统计、网络性能与训练曲线（ECharts）。

9.3 7.3 部署与运维

- Docker/K8s：镜像、编排与资源限制。
- 监控与日志：Prometheus/Grafana、结构化日志、追踪与告警。

9.4 7.4 测试与性能

- 单元/集成/性能测试用例与覆盖。
- 压测：并发请求、场景回放与最坏情况导致的退化分析。

9.5 7.5 本章小结

展示系统工程落地与可维护性，保证实验与演示的稳定可靠。

9.6 附：实现伪代码

```
# Flask API 示例
@app.post('/api/admission/decision')
def decide():
    req = parse_request(request.json)
    s = env.extract_state(req)
    a = agent.predict(s)
    result = dsroq_interface.execute_admission(a, req, env.time())
    return jsonify(result.to_dict())

# 训练入口
def train(total_steps):
    for _ in range(total_steps):
        s = env.reset_episode()
        done = False
        while not done:
            a = agent.sample_action(s)
```

```

s2, r, done, info = env.step(a)
agent.buffer.add(s, a, r, s2, done)
s = s2
agent.update()

```

9.7 附：图表清单（建议）

- 系统组件与调用关系图
 - API 接口列表与时序
 - 训练/推理性能曲线与资源占用图
-

9.8 附：定位 API 示例（Flask）

```

@app.get('/api/positioning/metrics')
def get_positioning_metrics():
    t = int(request.args.get('time', env.time()))
    users = request.args.get('users', '[]') # JSON list of user ids/locations
    users = json.loads(users)
    metrics = pos_module.metrics(t, users)
    # metrics: { 'crlb': {'mean':..., 'p95':...}, 'gdop': {...},
    #           'visible_beams': avg, 'coop_sats': avg, 'beam_hint': {...} }
    return jsonify(metrics)

@app.post('/api/positioning/beam_hint')
def get_beam_hint():
    payload = request.json # { 'time': t, 'users': [...], 'budget': {...} }
    hint = pos_module.beam_schedule_hint(payload)
    # hint: per-user recommended beams/sat set to support positioning
    return jsonify(hint)

```

9.9 附：前端联动示例（Vue 风格伪代码）

```

<template>
  <div class="dashboard">
    <div ref="cesiumContainer" class="cesium" />
    <charts-panel :qoe="qoe" :pos="pos" :rates="rates" />
  </div>
</template>
<script>
export default {
  data(){ return { qoe:{}, pos:{}, rates:{} } },
  mounted(){ this.tick() },
  methods:{
    async tick(){
      setInterval(async()=>{
        const t = Date.now()/1000|0;
        const pos = await fetch(`/api/positioning/metrics?time=${t}`).then(r=>r.json())
        const rates = await fetch('/api/stats/admission').then(r=>r.json())
        const qoe = await fetch('/api/stats/qoe').then(r=>r.json())
        this.pos = pos; this.rates = rates; this.qoe = qoe;
      })
    }
  }
}

```

```

        // 可选：依据 beam_hint 更新 Cesium 图层
        const hint = await fetch('/api/positioning/beam_hint',{method:'POST',body:JSON.stringify({time:t, users:this.users}),
        this.updateCesiumLayers(hint)
        }, 1000)
    },
    updateCesiumLayers(hint){ /* 渲染推荐波束/协作卫星 */ }
}
}
</script>

```

9.10 参考公式对齐

详见 docs/reference/formula_alignment.md。

10 第 8 章 消融实验与参数分析

10.1 8.1 状态空间消融

移除时间/历史/预测要素对性能的影响，分析关键特征的重要性。

10.2 8.2 动作空间消融

从二元（接收/拒绝）到细粒度动作的增量效果评估。

10.3 8.3 奖励函数消融

仅 ΔQoE 、加入公平性、加入效率、加入违规惩罚的组合对比与权重敏感性。

10.4 8.4 超参数敏感性

学习率、折扣、clip、n_steps、batch、epochs 等的敏感性曲线与热力图。

10.5 8.5 本章小结

总结各组件与超参对整体性能与稳定性的贡献与影响。

10.6 附：消融实验流程（伪代码）

```

def run_ablation(configs):
    results = {}
    for name, cfg in configs.items():
        env = make_env(cfg.env)
        agent = make_agent(cfg.agent)
        res = evaluate(agent, env, seeds=cfg.seeds)
        results[name] = aggregate(res)
    return compare(results)

```

10.7 附：图表清单（建议）

- 组件消融柱状/折线对比图
 - 权重/超参数敏感性热力图与曲线
 - 稳定性（方差/置信区间）对比图
 - 定位特征消融：
 - CRLB/GDOP 移除/加入对 QoE/AR/定位可用性的影响
 - 可见波束/协作卫星移除/加入的影响
 - Beam Hint 移除/加入对定位与网络双目标的影响
-

10.8 附：图表生成命令示例（定位消融）

```
# QoE 对比（含 95%CI）
python scripts/plots/qoe_metrics_plot.py \
  --input experiments/results/positioning_ablation_example.json \
  --output docs/assets/ablation_qoe_positioning.png

# 定位可用性/准入率/降级率对比
python scripts/plots/admission_rates_plot.py \
  --input experiments/results/positioning_ablation_example.json \
  --output docs/assets/ablation_rates_positioning.png

# 敏感性热力图（示例：自备矩阵 JSON）
python scripts/plots/fairness_heatmap.py \
  --input experiments/results/sensitivity_matrix.json \
  --output docs/assets/ablation_sensitivity_positioning.png
```

11 第 9 章 结论与展望

11.1 9.1 研究工作总结

提出面向 LEO 卫星网络的 DRL 准入控制，与 DSROQ 形成分层协同，实现 QoE 驱动的全局优化并在 Hypatia 上完成高保真验证。

11.2 9.2 主要贡献

- 首个面向 LEO 的实时 DRL 准入控制与时间感知状态空间。
- 准入-分配分层架构与细粒度动作空间。
- 多维度实验、消融与敏感性评估，展现综合优势。

11.3 9.3 不足与局限

- 泛化与跨星座迁移能力仍需强化；
- 训练成本与收敛速度受限于仿真复杂度；
- 现实系统中的在线约束与安全性尚需进一步验证。

11.4 9.4 未来工作方向

- 联邦/分布式强化学习；

- 在线/自适应与安全约束的策略优化；
- 多域协同（空天地海）与跨域 QoE 管理；
- 真实轨道与测量数据驱动的校准与部署。

11.5 附：图表清单（建议）

- 论文贡献总览图
- 核心结果回顾（关键指标雷达图/对比表）

11.6 附：未来路线图（占位）

- 短期：实验规模扩展、更多场景与数据公开、代码复现完善
- 中期：在线学习/安全约束、与定位/时频一体化协同
- 长期：真实系统试验、跨域融合、多任务多目标联合优化

12 术语与缩写表（Glossary）

说明：本文统一采用中文为主、括号给出英文与缩写（如有）。

- LEO (Low Earth Orbit, 近地轨道)：高度约 160—2000 km 的轨道层，具备低时延与广覆盖特性。
- QoE (Quality of Experience, 用户体验质量)：从用户感知出发的综合体验指标，受时延、丢包、吞吐、波动等影响。
- QoS (Quality of Service, 服务质量)：网络层面可测度的性能保障指标，如带宽、时延、抖动、丢包。
- EF/AF/BE (Expedited/Assured/Best Effort, 业务分类)：分别表示加速转发、可保障转发、尽力而为三类业务优先级。
- 准入控制 (Admission Control)：决定新业务/新流是否接入网络的策略与机制。
- DRL (Deep Reinforcement Learning, 深度强化学习)：结合深度学习与强化学习的序贯决策方法。
- PPO (Proximal Policy Optimization, 近端策略优化)：常用的策略梯度类 DRL 算法，收敛稳定性较好。
- DSROQ (Routing/Bandwidth/Scheduling Joint Optimization, 联合路由/带宽/调度优化)：本文指代的下层执行优化与控制框架。
- Hypatia：LEO 卫星网络仿真/建模平台（本项目作为仿真后端与数据源接口）。
- MDP (Markov Decision Process, 马尔可夫决策过程)：由状态、动作、转移、奖励、折扣构成的决策建模框架。
- 状态 (State)：用于决策的特征集合，包括网络、业务、QoE/QoS、定位等信息。
- 动作 (Action)：策略可采取的决策，如接受、拒绝、降级、延迟、部分接受等。
- 奖励 (Reward)：用于评价动作优劣的信号，如 ΔQoE 、公平性、效率、违规惩罚、定位可用性等加权组合。
- Jain 公平性 (Jain's Fairness Index)：度量分配公平性的指标，范围 (0,1]，越接近 1 越公平。
- 网络利用率 (Utilization, Util)：链路或全网资源的使用程度占比。
- 违规率 (Violations, Viol)：QoS/约束被违反的比例。
- 定位可用性 (A_{pos})：本文定义的定位可用性评分，范围 [0,1]。
- CRLB (Cramér–Rao Lower Bound, 克拉美–罗下界)：估计方差的理论下界，用于度量定位精度下限。
- GDOP (Geometric Dilution of Precision, 几何精度因子)：由几何构型导致的定位精度劣化因子，越小越好。
- SINR (Signal-to-Interference-plus-Noise Ratio, 信干噪比)：接收信号与干扰/噪声的比值，越大越好。
- Beam Hint (波束调度提示)：为提升定位/通信目标而推荐的波束/卫星组合集合。
- 可见波束 (Visible Beams)：在给定时刻对用户可见/可服务的波束数量或集合。
- 协作卫星 (Cooperative Satellites)：参与为同一用户提供几何/测量支撑的卫星集合。

- MCTS (Monte Carlo Tree Search, 蒙特卡洛树搜索): 常用于路由/调度搜索与策略改进的启发式方法。
- 李雅普诺夫优化 (Lyapunov Optimization): 在队列/稳定性约束下的在线控制与调度方法。
- 置信区间 (Confidence Interval, CI): 统计上对均值等指标区间置信度的度量, 常用 95%CI。
- CDF (Cumulative Distribution Function, 累积分布函数): 描述随机变量取值不超过某值的概率分布函数。
- 消融实验 (Ablation Study): 移除或替换系统组件以评估其对性能的影响。
- 敏感性分析 (Sensitivity Analysis): 研究超参数或权重变化对结果的影响。
- 可视化 (Visualization): 将系统状态、指标与结果以图形化方式呈现 (如 ECharts、Cesium)。
- API (Application Programming Interface, 应用编程接口): 模块间通信的接口定义。
- KPI (Key Performance Indicator, 关键绩效指标): 用于衡量系统目标达成情况的核心指标集合。
- 资源重分配 (Reallocation): 在网络状态变化或约束触发下对路由/带宽/调度方案进行调整。
- 冷却时间 (Cooldown): 限制频繁迁移/切换以稳定拓扑/几何的时间窗口设置。

若发现术语缺漏或需统一翻译, 请在本文件中追加或修订, 并在首次出现处保持与此表一致的写法。