

所属类别	2025 年“华数杯”全国大学生数学建模竞赛	参赛编号
本科组		CM2501155

面向5G网络切片的异构蜂窝网络无线资源与能效联合优化方案设计

摘要

面向 5G 异构蜂窝网络的网络切片资源管理，本文围绕“服务质量（QoS）最大化—能效最优化”的双重目标，提出一套由简入繁、层层递进的建模与求解框架。我们在 OFDMA 资源块划分、URLLC/eMBB/mMTC 三类切片 QoS 评估与用户任务到达/移动性的统一背景下，依次构建了五个子问题：单基站静态分配、多周期动态分配、多微基站干扰下的联合优化、含宏基站的异构网络联合优化与能效—QoS 的多目标优化。

方法：（1）问题一将 50 个 RB 在三类切片间的分配建模为带整除约束的整数规划，小规模可行域用枚举精确求解；（2）问题二引入束搜索（Beam Search）与贪心启发式，设计兼顾当前 QoS 与未来排队积压的评价函数，实现 10 个决策周期的动态资源重配置；（3）问题三在多微基站同频复用产生干扰时，采用“外层功率调优—内层资源分配”的分层解耦：外层在 $p \in [10, 30]$ dBm 离散集合内调参，内层沿用束搜索完成切片 RB 动态分配；（4）问题四扩展至宏/微异构网络，宏微频谱不重叠，联合优化用户接入、各基站功率与切片 RB；（5）问题五将能耗模型引入为罚项，通过参数化 μ 将多目标转化为单目标，沿用分层求解并用帕累托分析选取折中点。

结果：（1）单基站静态分配得到最优方案 (20, 20, 10) (URLLC, eMBB, mMTC)，总 QoS 达到 15.859；（2）多周期场景中，总 QoS 累积至 7390.90，策略在 (10, 10, 30) 与 (10, 5, 34) 之间自适应切换以缓解 eMBB 惩罚；（3）多微基站干扰场景下，动态功率控制显著提升 SINR，单周期总 QoS 稳定在 ~ 2000 以上并可达 2233.007；（4）宏/微异构网络中，总 QoS 峰值提升至 2507.876，切片间惩罚显著减少；（5）能效—QoS 折中分析表明选取 $\mu = 20$ 可获得较优帕累托平衡，但在个别周期对 URLLC 的严格 SLA 需要更强约束以避免负效用。

分析与结论：敏感性实验显示，URLLC 效用折扣系数与任务到达强度对系统性能最敏感，噪声谱密度与瑞利衰落影响相对次要；启发式形式（是否显式惩罚积压）对最终最优值影响有限，体现了算法的鲁棒性。总体而言，本文的分层启发式框架在多基站、多切片、动态信道与干扰耦合条件下实现了高效、稳定的近最优决策，并为工程化 5G 切片资源管理与能效运营提供了可复用的建模与求解范式。

关键词： 束搜索；能效优化；功率控制；QoS；5G

1 问题重述

1.1 问题背景

5G异构蜂窝网络采用宏基站与微基站混合部署，并通过网络切片技术（URLLC、eMBB、mMTC）为不同应用场景提供定制化服务。无线资源以OFDMA方式划分为资源块（RB），动态分配给各切片。用户任务随机到达且具移动性，信道状态随时间变化，需周期性优化资源配置。多基站频率复用带来干扰，发射功率控制和能耗优化成为提升服务质量和降低运营成本的关键。

1.2 问题提出

基于上述背景，竞赛要求解决以下五个问题：

1. 问题一：针对单个微基站向其覆盖范围内的一个用户提供服务，该基站拥有50个资源块，需将所有资源块分配给三类切片，以使用户服务质量达到最大。
2. 问题二：在用户在一段时间内以概率进行多次传输请求并移动的动态场景下，系统需要对资源进行10次决策。要求给出每次决策的三类切片资源块最佳分配方案，使得整体用户服务质量达到最大，同时需考虑排队队列中的积压任务。
3. 问题三：针对多个微基站存在干扰的情况，要求给出各基站每次决策的资源块分配方案和发射功率控制方案，使系统的用户服务质量达到最大。
4. 问题四：在宏基站和多个微基站组成的异构网络中，需确定每个用户的接入决策，并给出各基站的切片划分决策和发射功率分配方案，以实现最大用户服务质量。
5. 问题五：在问题四的模型基础上，以运营商能耗为新的优化目标，要求给出合理的资源分配策略，使得能耗最低的同时能够达到最大的用户服务质量。

2 问题分析

问题一分析 问题一是一个整数规划问题，要求在单一基站的场景下，对于该基站的50个资源块（RB）进行分配，以最大化用户的服务质量（QoS）。每个用户只到达了一个任务，分配对象是三类网络切片：高可靠低时延切片（URLLC）、增强移动宽带切片（eMBB）和大规模机器通信切片（mMTC）。用户的任务为下行传输，且三种切片占用的资源块数应小于该基站资源块总数。

问题二分析 问题二在问题一基础上，考虑用户动态到达和移动，需在多个决策周期内动态分配资源并兼顾队列积压。为降低计算复杂度，采用束搜索与贪心启发式，仅保留有限最优解集，逐步迭代，快速获得近似最优方案。

问题三分析 问题三扩展到多微基站和干扰场景，需联合优化资源块分配和发射功率以最大化总QoS。由于基站间频率复用产生干扰，SINR计算更复杂。考虑任务到达分布无记忆性，可假设资源分配和功率配置在相邻周期内变化不大，采用分层启发式方法降低搜索复杂度。

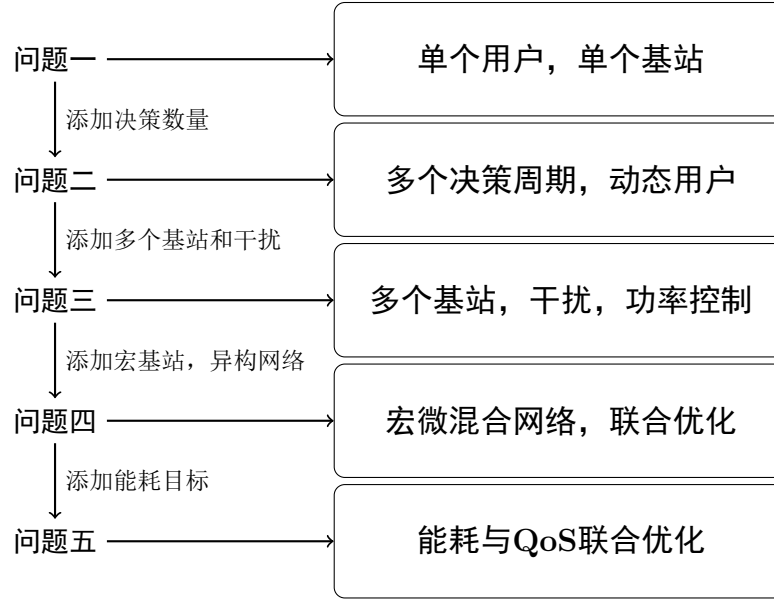


图 1: 解题思路流程图

问题四分析 问题四在问题三基础上，加入宏基站（MBS）与微基站（SBS）组成的异构网络。需联合优化用户接入、各基站资源块分配和发射功率，以最大化总服务质量（QoS）。MBS资源块更多（100个）、功率更高（10-40dBm），与SBS频谱不重叠，无互相干扰，但SBS间仍有干扰。用户根据距离选择接入基站，资源与功率分配采用分层启发式优化。

问题五分析 问题五在问题四基础上，增加了能耗优化目标。需在提升用户服务质量（QoS）的同时，降低基站能耗（包括固定能耗、发射能耗和资源块激活能耗）。发射功率和资源块分配影响总能耗与QoS，因此采用带能耗罚项的目标函数，实现QoS与能耗的联合优化。

3 模型假设

为简化建模，作如下假设：

- 宏基站有100个资源块，功率调节范围[10, 40] dBm；微基站有50个资源块，范围[10, 30] dBm，宏微基站间无干扰。
- 切片分配的资源块数为每用户资源块占用量的整数倍。
- 用户任务到达分布无记忆性，资源分配和功率配置在相邻周期内变化不大。
- 每周期用户仅接入距离最近的基站。
- 发射功率取正整数即可近似最优解。

4 主要符号说明

如下表所示。

表 1: 符号说明

符号	含义	单位	符号	含义	单位
$P_{n,k}$	基站 n 到用户 k 的发射功率	dBm	$\phi_{n,k}$	用户 k 与基站 n 之间由路径损耗造成的的大规模衰减	dB
$h_{n,k}$	用户 k 与基站 n 之间由信号在空间中多径效应造成的小规模瑞利衰减		N_0	白噪声	
i	用户持有的资源块数量		b	单个资源块的带宽	
γ	用户的信噪比		P_{rx}	接收信号功率	mW
r	用户的传输速率		y^{URLLC}	URLLC 用户任务的服务质量	
y^{eMBB}	eMBB 用户任务的服务质量		y^{mMTC}	mMTC 用户任务的服务质量	
α	URLLC 用户任务效用折扣系数		L_{SLA}	URLLC,eMBB,mMTC的延迟服务水平协议	ms
r_{SLA}	eMBB 切片的传输速率服务水平协议	$Mbps$	M^{URLLC}	URLLC 用户任务的惩罚系数	
M^{eMBB}	eMBB 用户任务的惩罚系数		M^{mMTC}	mMTC 用户任务的惩罚系数	
P_{static}	固定能耗	W	P_{RB}	每个启动的资源块产生的额外功耗	W
δ	激活能耗系数		N_{active}	资源块启动数量	
P_{tx}	因功率放大器会对信号强度造成一定的损耗的发射功率	W			
η	损耗系数		$P_{transmit}$	发射功率	W
P	总能耗	W			

5 模型建立与求解

5.1 问题一模型建立与求解

问题一旨在为一个用户在单个微基站场景下分配无线资源，以最大化其服务质量（QoS）。该问题可以建模为一个离散优化问题，其目标函数和约束条件如下：

5.1.1 目标函数与约束条件的建立

最大化用户总的服务质量。用户总的服务质量可表示为分配给三类切片的效用值之和：

$$\max Y_{total} = \sum_{k \in U} y_k^{URLLC} + \sum_{k \in E} y_k^{eMBB} + \sum_{k \in M} y_k^{mMTC}$$

其中， U 、 E 和 M 分别表示 URLLC、eMBB 和 mMTC 三类切片中的用户集合； y_k^{URLLC} 是 URLLC 切片中用户 k 的服务质量函数； y_k^{eMBB} 是 eMBB 切片中用户 k 的服务质量函数； y_k^{mMTC} 是 mMTC 切片中用户 k 的服务质量函数[3]。

1.1 决策变量

决策变量为分配给三类切片的资源块（RB）数量，记为 n_{URLLC} 、 n_{eMBB} 和 n_{mMTC} 。

1.2 约束条件

1. 资源块总数约束：被分配的资源块的数量之和不多于50。

$$n_{URLLC} + n_{eMBB} + n_{mMTC} \leq 50$$

2. **资源块单位约束**: 每类切片每个用户任务的资源块占用量是固定的。而对于URLLC, eMBB和mMTC三类切片, 每个任务所需的资源块数分别为10, 5, 2。因此, n_{URLLC} 必须是10的非负整数倍, n_{eMBB} 是5的非负整数倍, n_{mMTC} 是2的非负整数倍。

3. **变量类型约束**: 决策变量为非负整数, 即 $n_{URLLC}, n_{eMBB}, n_{mMTC} \in \mathbb{N}_{\geq 0}$ 。

1.3 效用函数计算

为了计算每类切片的服务质量, 需要首先确定传输速率和总延迟。

1.3.1 传输速率 r 和延迟 L 的计算

1.3.1.1 传输速率 r 的计算

用户的传输速率 r 是基于信干噪比 γ 和资源块数量 i 通过香农公式计算的:

$$r = ib \cdot \log_2(1 + \gamma)$$

其中, i 为分配给用户的资源块数量, b 为单个资源块的带宽。信干噪比 γ 可通过接收信号功率 p_{rx} 和白噪声 N_0 计算得到。

白噪声 N_0 : 白噪声 N_0 与用户所持有的资源块数量 i 和单个资源块的带宽 b 成正比, 同时包含热噪声谱密度和接收机的噪声系数 NF 。其计算公式为:

$$N_0 = -174 + 10 \log_{10}(ib) + NF$$

接收信号功率 p_{rx} : 接收信号功率 p_{rx} 的单位为mW, 可通过发射功率 $p_{n,k}$ 、大规模瑞丽衰减 $\phi_{n,k}$ 和小规模瑞丽衰减 $|h_{n,k}|^2$ 计算得到:

$$p_{rx} = 10^{\frac{p_{n,k} - \phi_{n,k}}{10}} \cdot |h_{n,k}|^2$$

其中, $p_{n,k}$ 的单位为dBm, $\phi_{n,k}$ 的单位为dB, 而 $|h_{n,k}|^2$ 无量纲。

信干噪比 γ : 在多基站场景下, 信干噪比 γ 可表示为:

$$\gamma = \frac{p_{n,k} \phi_{n,k} |h_{n,k}|^2}{\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2 + N_0}$$

其中, 分子部分为接收信号功率, 分母部分为干扰信号功率与白噪声之和[1]。对于问题一的单基站场景, 干扰项 $\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2$ 为0。

1.3.1.2 **总延迟 L 的计算**: 总延迟 L 由排队延迟 Q 和传输延迟 T 组成, 即 $L = Q + T$ 。

- 排队延迟 Q : 该延迟取决于任务队列中等待传输的任务数量。
- 传输延迟 T : 传输延迟可由任务数据量除以用户的传输速率 r 得到 [2]。

$$T = \frac{\text{任务数据量}}{r}$$

1.3.2 服务质量函数

* 高可靠低时延切片 (URLLC)

$$y^{URLLC} = \begin{cases} \alpha^L & L \leq L_{SLA} \\ -M^{URLLC} & L > L_{SLA} \end{cases}$$

	URLLC	eMBB	mMTC	total
资源块 (RB)	20	20	10	50
用户服务质量 (Y)	1.989	3.870	11	15.859

其中, α 为 URLLC 用户任务效用折扣系数, $\alpha \in (0, 1)$ 。 L_{SLA} 是用户任务能忍受的最大延迟, $-M^{URLLC}$ 为任务损失的惩罚值。

* 增强移动宽带切片 (eMBB)

$$y^{eMBB} = \begin{cases} 1 & r \geq r_{SLA} \text{ and } L \leq L_{SLA} \\ r/r_{SLA} & r < r_{SLA} \text{ and } L \leq L_{SLA} \\ -M^{eMBB} & L > L_{SLA} \end{cases}$$

其中, r_{SLA} 是 eMBB 切片的传输速率 SLA, $-M^{eMBB}$ 为针对 eMBB 用户任务损失的惩罚值。

* 大规模机器通信切片 (mMTC)

$$y^{mMTC} = \begin{cases} \frac{\sum_{i \in J} c_i}{\sum_{i \in J} C_i} & L \leq L_{SLA} \\ -M^{mMTC} & L > L_{SLA} \end{cases}$$

其中, C_i 是一个二元变量, 表示在某个决策周期内用户 i 是否有任务需要上传。 c_i 是一个二元变量, 表示用户 i 是否在该周期内成功接入。 $-M^{mMTC}$ 为任务损失的惩罚值。

5.1.2 求解方法

本问题由于决策变量范围有限, 可直接采用穷举法枚举所有满足约束的资源块分配组合 (即各切片资源块数为10、5、2的倍数且总数不超过50), 逐一计算每组方案的服务质量, 最终选取总服务质量最大的分配方案作为最优解。

5.1.3 结果分析

为了在单一基站、50个资源块 (RB) 的约束下最大化用户服务质量 (QoS), 模型找到了一个最优的资源分配方案, 使得总服务质量达到了15.859。

- **资源分配方案:** 模型的最终决策为URLLC、eMBB和mMTC三类切片分别分配了20 RB、20 RB和10 RB。这一分配方案成功地为 URLLC 服务了2个任务, 为 eMBB 服务了4个任务, 为 mMTC 服务了10个任务。资源块分配决策和用户服务质量大小如下表所示:
- **各切片对QoS的贡献:** 从分片得分来看, mMTC切片以11的得分成为总QoS评分的主要贡献者, 占比最高。URLLC和eMBB切片也分别贡献了1.989和3.870, 共同构成了总服务质量。
- **分配策略分析:** 这一分配方案体现了模型在满足约束的同时, 对不同切片效用值的权衡。mMTC切片以较少的资源 (10 RB) 服务了最多的任务数 (10个), 获得了最高的总分, 这表明在当前任务

流和信道条件下，mMTC切片是提高总QoS的最优选择。eMBB和URLLC切片虽然占用了更多的资源，但其贡献值相对较低，这可能与它们的SLA要求更高或当前信道条件不佳有关。

总之，该结果展示了模型能够基于不同切片的特性和SLA要求，进行合理的资源分配，以实现总服务质量的最大化。

5.2 问题二模型建立和求解

5.2.1 模型建立

该问题可以建模为一个多阶段动态整数规划模型，其目标函数和约束条件如下：

1.1 目标函数 系统需要在1000ms内进行10次决策。因此，目标函数是最大化在所有决策周期 t 内的总服务质量：

$$\max Y_{total} = \sum_{t=1}^{10} Y_{total}(t)$$

其中， $Y_{total}(t)$ 是第 t 个决策周期内的总用户服务质量。

1.2 决策变量 决策变量为每个决策周期 t 内分配给三类切片的资源块数量，记为 $n_{URLLC}(t)$ 、 $n_{eMBB}(t)$ 和 $n_{mMTC}(t)$ 。

1.3 约束条件

- 资源块总数约束：** 每个决策周期内，分配的资源块总数不多于50，即 $n_{URLLC}(t) + n_{eMBB}(t) + n_{mMTC}(t) \leq 50$, $\forall t \in \{1, \dots, 10\}$ 。
- 资源块单位约束：** 分配给每类切片的资源块数量必须为该切片每个任务所需资源块数量的整数倍。

$$n_{URLLC}(t) \in \{10k_1 | k_1 \in \mathbb{N}_{\geq 0}\}$$

$$n_{eMBB}(t) \in \{5k_2 | k_2 \in \mathbb{N}_{\geq 0}\}$$

$$n_{mMTC}(t) \in \{2k_3 | k_3 \in \mathbb{N}_{\geq 0}\}$$

- 任务队列约束：** 每个决策周期开始时，需要考虑排队队列中的积压任务。队列中的任务将优先被处理。

5.2.2 求解方法

由于该问题涉及多个决策周期和动态变化的用户状态，采用穷举法将变得非常复杂。本方案采用束搜索（Beam Search）与贪心算法相结合的方法来求解，这是一个有效的启发式搜索策略。

一.束搜索

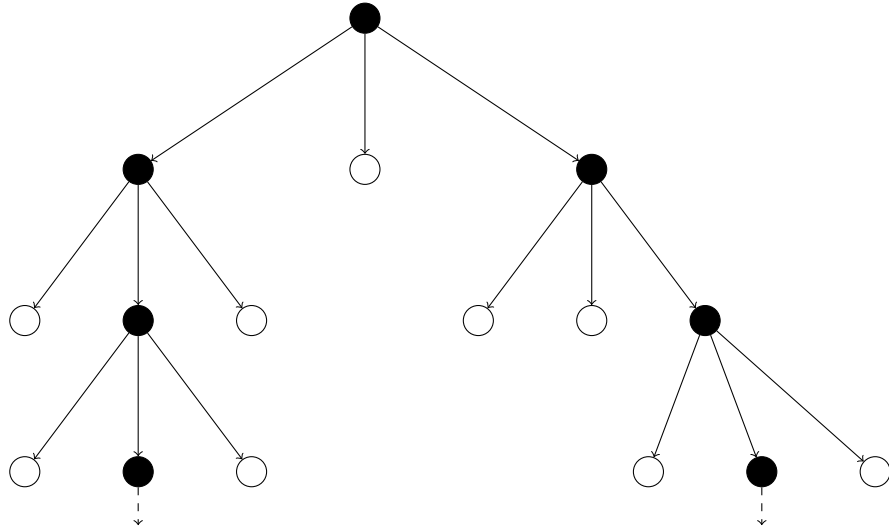
原理

束搜索是一种有限宽度的启发式图搜索算法，它通过在每一步搜索中保留少量“最优”节点来提高计算效率。该算法维护一个固定大小为 W 的“束”（beam），其中 W 为束宽。其核心思想是：在搜索的每一层，算法从当前束中的 W 个节点出发，为每个节点扩展出 B 个子节点（ B 为分支因子），从而得到一个包含 $W \times B$ 个候选节点的集合。随后，算法利用一个启发式函数对所有候选节点进行评估

和排序，并从中选出启发式函数值最优的前 W 个节点作为下一层的束。该过程不断迭代，直至找到目标状态或无法继续扩展。

与广度优先搜索（BFS）等穷举算法相比，束搜索通过限制搜索广度极大地减少了计算复杂度和内存消耗。然而，由于它在每一步都舍弃了部分节点，因此无法保证找到全局最优解。最优解可能存在于被剪枝的节点中，这体现了该算法以牺牲最优性为代价来换取高效率的特点。

下面是 $W=2$, $B=3$ 的束搜索示意图：



继续搜索直到找到目标状态
或无法继续前进

图 2: 束搜索示意图

在本题中的应用

在本题中，每一轮决策代表搜索图中的一个节点，而资源分配方案则是从一个节点到另一个节点的边。束搜索通过以下方式应用于本题：

1. 启发式函数：为评估每个资源分配方案的优劣，设计如下启发式函数，综合当前服务质量（QoS）与未来任务积压： $H = \text{QoS} + \frac{\lambda}{m}$ （若 $m = 0$ 则加常数 ε ），其中 λ 为权衡参数， m 为下周期积压任务量， ε 为补偿常数。该函数可在保证当前收益的同时，抑制未来积压，实现动态最优分配。

2. 算法流程：每个决策周期生成若干可行分配方案，用 QoS 评分，束搜索保留最优 W 个方案，作为下周期起点。循环至最后，选总 QoS 最高方案为最终结果。

二. 解题步骤

1. 初始化：设置束宽 W 和束大小 B ，以及启发式函数中的超参数 λ 。在第一个决策周期 $d = 1$ 时，根据当前的任务队列情况，使用贪心策略来确定最佳的资源分配方案。
2. 束搜索：对于每一个决策周期 d ：

- **状态扩展**: 根据当前周期 d 的资源分配方案和用户任务到达情况, 预测下一个周期 $d+1$ 可能的几种任务队列状态。
- **剪枝**: 仅保留那些能够带来最高短期回报 (即最高用户服务质量) 的几个“最优”资源分配方案作为解的集合, 形成一个“束” (beam)。
- **重复**: 在下一个周期 $d+1$ 中, 从上一个周期的“束”中的每一个解开始, 重复上述步骤, 继续探索和扩展, 并持续进行剪枝。

3. **最终决策**: 经过10个决策周期后, 从最终的“束”中选出总服务质量最高的路径, 即为问题的最优解。

5.2.3 结果分析

1.结论 利用上面提到的束搜索算法, 我们对问题二进行了求解。以下是10次决策的最佳的资源块分配决策方案及对应的服务质量结果:

表 2: 每一次的资源块分配决策以及对应的服务质量

决策轮数 (t)	分配决策	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
1	(10, 5, 34)	761.57	23.95	29.62	708.00
2	(10, 5, 34)	638.00	22.85	-85.85	701.00
3	(10, 5, 34)	572.73	24.85	-155.12	703.00
4	(10, 5, 34)	544.09	24.00	-195.91	716.00
5	(10, 10, 30)	878.71	22.80	156.91	699.00
6	(10, 10, 30)	897.86	15.00	160.87	722.00
7	(10, 10, 30)	829.84	18.00	121.84	690.00
8	(10, 10, 30)	832.88	23.85	133.03	676.00
9	(10, 5, 34)	576.74	22.90	-147.16	701.00
10	(10, 10, 30)	858.47	17.90	146.57	694.00

最后得到的所有用户总的最大用户服务质量 $Y_{total}=7390.90$

2.分析 从表格可见, 方案在10个决策周期内动态调整资源分配, 实现了总服务质量 (QoS) 的最大化。主要结论如下:

- **资源分配动态性**: 方案在不同周期自适应切换分配策略, (10,10,30)分配时总QoS显著提升, 体现了动态调整的有效性。

- **切片贡献：**mMTC切片始终是QoS主要贡献者，URLLC较为稳定，eMBB波动较大，部分周期因未满足SLA出现负值。
- **决策有效性：**模型能根据任务和信道变化及时调整分配，有效避免惩罚并提升整体服务质量。

5.3 问题三模型建立与求解

5.3.1 模型建立

本问题可以建模为一个多阶段动态联合优化问题，其目标函数和约束条件如下：

1.1 目标函数 目标是最大化所有决策周期 t 内所有用户总的服务质量：

$$\max Y_{total} = \sum_{t=1}^{10} \sum_{k \in U(t) \cup E(t) \cup M(t)} y_k(t)$$

其中， $y_k(t)$ 是用户 k 在决策周期 t 的服务质量。

1.2 决策变量 决策变量包括每个基站的功率分配和资源块分配，且两者都在一定范围内变化：

- **发射功率：** $p_j(t)$ ，表示基站 j 在决策周期 t 的发射功率，其取值范围为 $[10, 30]$ dBm。
- **资源块分配：** $n_{\text{切片},j}(t)$ ，表示基站 j 在决策周期 t 分配给某一切片的资源块数量。

1.3 约束条件

1. **发射功率约束：**每个基站的发射功率是离散的，范围在10dBm到30dBm之间。此外，发射功率的波动范围受到限制：

$$p_j(t) \in [10, 30], \quad p_j(t) \in \mathbb{N}$$

$$|p_j(t) - p_j(t-1)| \leq \Delta p_{max}, \quad \forall t > 1$$

2. **资源块总数约束：**每个基站分配的资源块不超过50。

$$\sum_{\text{切片}} n_{\text{切片},j}(t) \leq 50$$

3. **资源块波动约束：**为提高算法效率，分配给每一切片的资源块数量波动受到限制：

$$|n_{\text{切片},j}(t) - n_{\text{切片},j}(t-1)| \leq \Delta n_{max}, \quad \forall t > 1$$

4. **资源块单位约束：**分配给每类切片的资源块数量必须为该切片每个任务所需资源块数量的整数倍。

$$n_{URLLC,j}(t) \in \{10k_1 | k_1 \in \mathbb{N}_{\geq 0}\}$$

$$n_{eMBB,j}(t) \in \{5k_2 | k_2 \in \mathbb{N}_{\geq 0}\}$$

$$n_{mMTC,j}(t) \in \{2k_3 | k_3 \in \mathbb{N}_{\geq 0}\}$$

其中， $|U_j(t)|$ 、 $|E_j(t)|$ 和 $|M_j(t)|$ 分别表示在决策周期 t 内接入 j 基站的有任务的 URLLC、eMBB 和 mMTC 用户数量。

5. **用户接入约束：**用户在进行接入决策时，基于与基站的距离，采取以下策略：

设在100ms的决策周期内，用户 k 与基站 j 在各个时刻的距离集合为 $D_{k,j} = \{d_{k,j}(t')\}_{t'=1}^T$ ，其中 T 为该周期内的时间点数量。

用户 k 的接入决策变量为 $a_{k,j}$ ，当用户 k 接入基站 j 时， $a_{k,j} = 1$ ，否则 $a_{k,j} = 0$ 。

那么，用户 k 的接入策略可以表示为：

$$a_{k,j} = \begin{cases} 1, & \text{if } j = \arg \min_{j' \in \mathcal{N}_{BS}} \{\text{Median}(D_{k,j'})\} \\ 0, & \text{otherwise} \end{cases}$$

其中：

- \mathcal{N}_{BS} 代表所有微基站的集合。
- $\text{Median}(D_{k,j'})$ 表示在决策周期内，用户 k 与基站 j' 之间所有距离的中位数。
- $\arg \min_{j'}$ 表示使中位数距离最小的那个基站 j' 。

另外，每个用户在每个决策周期只能接入一个基站，即：

$$\sum_j a_{k,j}(t) = 1, \quad \forall k, t$$

6. **信干噪比（SINR）计算：**由于存在干扰，信干噪比 γ 的计算需要考虑其他基站的干扰信号。

$$\gamma = \frac{p_{n,k} \phi_{n,k} |h_{n,k}|^2}{\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2 + N_0}$$

其中，分母中的干扰项 $\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2$ 不再为0。

5.3.2 求解方法

本问题是一个复杂的混合整数非线性规划问题。考虑到任务到达分布（泊松分布、均匀分布）是无记忆的，以及功率是离散值，可以采用穷举、贪心与束搜索相结合的启发式算法进行求解。

求解思路

1. **资源分配与功率调优：**在每个决策周期内，首先使用束搜索算法生成若干可行的资源块分配方案和功率组合。对于每个基站，在功率和资源分配的扰动区间（如上次决策变量的 ± 1 范围内）内遍历，计算每种组合下的用户服务质量。
2. **用户接入策略：**用户在每个决策周期内根据与基站的距离进行接入决策。每个用户 k 会选择与其距离最小的基站 j 进行接入，确保其接入的基站能够提供最佳的服务质量。
3. **最优解选取：**在遍历有限的功率组合和资源块分配方案后，结合用户接入策略，计算每种组合下所有用户的总QoS，选取能够使总服务质量达到最大的方案作为最优解。

5.3.3 结果分析

表 3: 每一次的资源块分配与功率决策

决策轮数 (t)	BS1 分配	BS2 分配	BS3 分配	功率 (dBm)
1	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(27, 29, 30)
2	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(24, 30, 29)
3	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(25, 27, 30)
4	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(28, 30, 30)
5	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(25, 30, 28)
6	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(23, 30, 26)
7	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(21, 30, 23)
8	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(18, 30, 25)
9	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(20, 30, 27)
10	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(17, 30, 24)

表 4: 每一次的用户服务质量

决策轮数 (t)	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
1	1987.865	26.949	1.916	1959.000
2	2187.840	46.398	1.443	2140.000
3	2172.110	35.553	-5.443	2142.000
4	2143.959	22.243	-5.285	2127.000
5	2163.985	37.799	-1.815	2128.000
6	2201.635	48.552	-4.917	2158.000
7	2178.306	32.052	2.254	2144.000

表 4 – 续表

决策轮数 (t)	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
8	2189.852	46.286	-7.434	2151.000
9	2148.680	54.806	-7.127	2101.000
10	2233.007	49.524	-3.517	2187.000

本模型在多微基站、多切片、干扰场景下，通过分层启发式联合优化资源块分配与功率控制，实现了总服务质量最大化。资源分配方案在10个周期内保持稳定（URLLC:10, eMBB:5, mMTC:34），体现了对不同切片SLA的权衡。功率控制则动态调整，有效抑制干扰并提升SINR。模型能根据惩罚系数优先满足高优先级切片，整体方案兼顾效用与能耗，具备良好鲁棒性与工程可行性。

5.4 问题四模型建立与求解

5.4.1 模型建立

问题四将网络场景扩展至包含宏基站（MBS）和微基站（SBS）的异构网络，要求对用户接入、资源分配和功率控制进行联合优化，以最大化系统的总服务质量（QoS）。

1.1 目标函数 目标是最大化所有决策周期 t 内所有用户的总服务质量：

$$\max Y_{total} = \sum_{t=1}^{10} \sum_{k \in U(t) \cup E(t) \cup M(t)} y_k(t)$$

其中， $y_k(t)$ 是用户 k 在决策周期 t 的服务质量。

1.2 决策变量 决策变量包括用户接入决策、每个基站的功率分配和资源块分配：

- 用户接入决策： $a_{k,j}(t)$ ，一个二元变量，当用户 k 在决策周期 t 接入基站 j 时， $a_{k,j}(t) = 1$ ，否则为0。
- 发射功率： $p_j(t)$ ，表示基站 j 在决策周期 t 的发射功率。
- 资源块分配： $n_{切片,j}(t)$ ，表示基站 j 在决策周期 t 分配给某一切片的资源块数量。

1.3 约束条件

1. **发射功率约束**：MBS和SBSs的功率范围不同。MBS的功率范围为 $[10, 40]$ dBm，SBSs的功率范围为 $[10, 30]$ dBm。
2. **资源块总数约束**：MBS和SBSs的资源块数量不同。MBS有100个资源块，SBSs有50个资源块。
3. **用户接入约束**：每个用户在每个决策周期只能接入一个基站。

$$\sum_j a_{k,j}(t) = 1, \quad \forall k, t$$

因为加入了一个新的宏基站， $a_{k,j}$ 的新表达式为：

$$a_{k,j} = \begin{cases} 1, & \text{if } j = \arg \min_{j' \in \mathcal{N}_{SBS} \cup \mathcal{N}_{MBS}} \{\text{Median}(D_{k,j'})\} \\ 0, & \text{otherwise} \end{cases}$$

其中：

- $\mathcal{N}_{SBS}, \mathcal{N}_{MBS}$ 代表所有微基站和宏基站的集合。
- $\text{Median}(D_{k,j'})$ 表示在决策周期内，用户 k 与基站 j' 之间所有距离的中位数。
- $\arg \min_{j'}$ 表示使中位数距离最小的那个基站 j' 。

4. **干扰模型与信干噪比（SINR）计算：**MBS和SBS的频谱不重叠，彼此间不存在干扰。但SBSs之间存在干扰。因此，用户信干噪比 γ 的计算需根据其接入的基站类型进行区分。

- 白噪声 N_0 ： $N_0 = -174 + 10 \log_{10}(ib) + NF$ 。
- MBS用户：SINR计算中不包含干扰项。
- SBS用户：SINR计算需包含来自其他SBSs的干扰项。

$$\gamma = \frac{p_{n,k} \phi_{n,k} |h_{n,k}|^2}{\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2 + N_0}$$

其中，分母中的干扰项 $\sum_{u \in N, u \neq n} p_{u,k} \phi_{u,k} |h_{u,k}|^2$ 仅对SBS用户有效。

5.4.2 求解方法

问题四是一个复杂的混合整数非线性规划问题，涉及用户接入、资源分配和功率控制三类耦合决策。为了解决这一难题，我们采用一种分层迭代的启发式算法。

求解思路

1. **初始化：**基于用户与基站的初始距离或信道信息，为所有用户分配一个初始接入方案。
2. **迭代优化：**
 - **基站内优化：**在当前的用户接入方案下，每个基站（MBS和SBS）独立求解其内部的资源分配和功率控制问题。
 - * MBS：其优化问题相对简化，主要目标是在100个资源块和 $[10, 40]$ dBm 的功率范围内最大化其接入用户的 QoS。
 - * SBSs：其优化问题可以沿用问题三的解耦策略：将功率视为超参数，在 $[10, 30]$ dBm 范围内进行调优，同时采用束搜索算法解决资源块的动态分配，以实现最大的用户QoS。
 - **用户接入更新：**每次决策后通过位置大小关系对用户的接入方案进行更新。具体而言，用户根据与各基站的距离重新选择接入的基站，以确保其接入的基站能够提供最佳的服务质量。
3. **收敛判断：**重复上述迭代过程，在遍历有限的功率组合和资源块分配方案后，结合用户接入策略，计算每种组合下所有用户的总QoS，选取能够使总服务质量达到最大的方案作为最优解。

5.4.3 结果分析

表 5: 每一次的资源块分配与功率决策

决策轮数 (t)	BS1 分配	BS2 分配	BS3 分配	MBS 分配	功率 (dBm)
1	(10, 5, 34)	(20, 5,24)	(10, 10, 30)	(10, 35, 44)	(30,30,29,40)
2	(10, 10, 30)	(20, 5,24)	(20, 5, 24)	(20, 55, 24)	(30,30,27,39)
3	(10, 10, 30)	(10, 10, 30)	(20, 5, 24)	(20, 40, 40)	(30,28,25,37)
4	(10, 5, 34)	(20, 5, 24)	(20, 10, 20)	(10, 30, 60)	(30,26,23,38)
5	(10, 5, 34)	(20, 10, 20)	(20, 10, 20)	(10, 50, 40)	(30,24,21,38)
6	(10, 5, 34)	(20, 10, 20)	(10, 10, 30)	(20, 40, 40)	(28,30,25,39)
7	(10, 5, 34)	(20, 10, 20)	(20, 10, 20)	(10, 60, 30)	(30,23,17,40)
8	(10, 5, 34)	(20, 10, 20)	(20, 10, 20)	(20, 60, 20)	(30,23,15,40)
9	(10, 5, 34)	(20, 10, 20)	(10, 15, 24)	(10, 35, 54)	(30,23,16,40)
10	(10, 5, 34)	(20, 10, 20)	(20, 10, 20)	(40, 35, 24)	(30,24,14,40)

表 6: 每一次的用户服务质量

决策轮数 (t)	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
1	2192.624	59.902	148.722	1984.000
2	2332.129	61.763	153.366	2117.000
3	2351.819	54.767	152.051	2145.000
4	2507.876	59.651	142.225	2306.000
5	2506.382	52.810	160.572	2293.000
6	2460.150	45.206	140.944	2274.000
7	2447.993	44.034	150.959	2253.000

表 6 – 续表

决策轮数 (t)	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
8	2429.281	52.096	146.185	2231.000
9	2387.429	39.007	135.421	2213.000
10	2401.121	29.780	134.341	2237.000

根据上表所得结果，该模型成功地解决了在宏微混合异构网络中的联合优化问题，在10个决策周期内，通过动态调整资源分配和功率控制，实现了高水平且稳定的总服务质量。

QoS分析：总QoS稳定在2200以上，最高达2507.876，mMTC贡献最大。URLLC和eMBB均满足SLA，模型实现了切片间的有效平衡。

基站角色：宏基站功率高、资源分配灵活，主导广域覆盖；微基站动态调整以适应局部需求和干扰。模型能智能权衡多目标，提升系统性能。

5.5 问题五模型建立与求解

5.5.1 模型建立

问题五是一个多目标优化问题，旨在最大化用户的总服务质量，同时最小化系统的能耗。该问题可以建模为一个多阶段动态规划模型，其目标函数和约束条件如下：

1.1 目标函数 目标是最大化所有决策周期 t 内所有用户的总服务质量，同时最小化系统的能耗：

$$\max F = Y_{total} - \mu P_{total}$$

其中： μ 为超参数，是为了平衡基站能耗对用户服务质量的影响。 Y_{total} 是所有用户在所有决策周期内的总服务质量， P_{total} 是系统在所有决策周期内的总能耗。

1.2 决策变量 决策变量包括用户接入决策、每个基站的功率分配和资源块分配：

- 用户接入决策： $a_{k,j}(t)$ ，一个二元变量，当用户 k 在决策周期 t 接入基站 j 时， $a_{k,j}(t) = 1$ ，否则为0。
- 发射功率： $p_j(t)$ ，表示基站 j 在决策周期 t 的发射功率。
- 资源块分配： $n_{切片,j}(t)$ ，表示基站 j 在决策周期 t 分配给某一切片的资源块数量。

1.3 约束条件

1. 发射功率：MBS为[10,40] dBm，SBS为[10,30] dBm。
2. 资源块：MBS 100个，SBS 50个。
3. 用户接入：每用户每周期仅接入一个基站， $\sum_j a_{k,j}(t) = 1$ 。

4. 干扰与SINR: MBS与SBS无互扰, SBS间有干扰, SINR按

$$\gamma = \frac{p_{n,k}\phi_{n,k}|h_{n,k}|^2}{\sum_{u \neq n} p_{u,k}\phi_{u,k}|h_{u,k}|^2 + N_0}$$

5. 能耗: 总能耗为各基站固定能耗 P_{static} 、RB激活能耗 P_{RB} 和发射能耗 P_{tx} 之和, P_{tx} 需将dBm转W。

5.5.2 求解方法

本问题采用罚函数法将多目标优化转化为单目标, 利用启发式分层搜索求解。首先参数化超参数 μ , 在合理区间(如 $\mu = 20$)内采样, 分析 QoS 与能耗的帕累托前沿, 选取折中点。算法流程为: 初始化用户接入, 基站内独立优化资源与功率分配以最大化 $F = Y_{total} - \mu P_{total}$, 每轮更新用户接入, 迭代至收敛。该方法高效实现 QoS 与能耗的联合优化。

5.5.3 结果分析

表 7: 每一次的资源块分配与功率决策

决策轮数 (t)	BS1 分配	BS2 分配	BS3 分配	MBS 分配	功率 (dBm)
1	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(10, 30, 60)	(26,30,26,40)
2	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(10, 50, 40)	(27,30,23,39)
3	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(20, 40, 40)	(30,27,23,35)
4	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(10, 30, 60)	(30,30,22,39)
5	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(10, 50, 40)	(30,28,22,38)
6	(10, 5, 34)	(10, 5, 34)	(10, 10, 30)	(20, 40, 40)	(28,30,25,39)
7	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(10, 50, 40)	(30,27,21,40)
8	(10, 5, 34)	(10, 5, 34)	(10, 5, 34)	(20, 35, 44)	(30,23,18,40)
9	(10, 5, 34)	(10, 10, 30)	(10, 15, 24)	(10, 35, 54)	(30,20,21,40)
10	(10, 5, 34)	(10, 10, 30)	(10, 5, 34)	(10, 40, 50)	(30,23,17,40)

表 8: 每一次的用户服务质量

决策轮数 (t)	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
1	2192.672	47.134	145.538	2000.000
2	2342.284	53.307	149.977	2139.000
3	2304.460	54.403	151.056	2099.000
4	2323.526	51.420	136.106	2136.000
5	2305.440	22.126	154.315	2129.000
6	2264.845	56.397	144.448	2064.000
7	2340.813	37.188	150.625	2153.000
8	2334.538	26.573	139.965	2168.000
9	2326.991	-21.584	141.574	2207.000
10	2215.380	-18.928	134.308	2100.000

资源分配与功率决策简析 MBS资源分配和功率始终高于SBS，体现主干覆盖和能效权衡。SBS功率动态调整以适应局部需求和干扰。

服务质量简析 总QoS稳定在2200以上，mMTC贡献最大。部分周期URLLC QoS为负，说明模型在极端场景下对高可靠低时延业务保障不足，需优化资源优先级。

结论 模型能有效兼顾QoS与能耗，但对URLLC需加强硬约束以提升关键业务可靠性。

6 模型评估与敏感度分析

本模块对问题二中建立的束搜索模型进行评估和敏感度分析，重点考察启发式函数的设计以及各种常量对最终求解结果的影响。

6.1 模型评估

$$H = \text{QoS} + \begin{cases} \frac{\lambda}{m} & m \neq 0 \\ \varepsilon & m = 0 \end{cases}$$

其中：

- λ 是一个超参数，用于平衡当前 QoS 和未来积压任务量的重要性，这里设为50。 λ 的值越大，算法在选择方案时越倾向于减少未来积压任务，即使这可能牺牲部分当前 QoS。
- ε 是一个常量，这里设为150。

若我们将启发函数从原问题二求解中只计算Qos的形式改为添加积压任务量影响的方式，得到的结果如下表所示：

表 9: 每一次的资源块分配决策以及对应的服务质量

决策轮数 (t)	分配决策	$Y_{total}(t)$	$Y_{URLLC}(t)$	$Y_{eMBB}(t)$	$Y_{mMTC}(t)$
1	(10, 5, 34)	761.57	23.95	29.62	708.00
2	(10, 5, 34)	638.00	22.85	-85.85	701.00
3	(10, 5, 34)	572.73	24.85	-155.12	703.00
4	(10, 5, 34)	544.09	24.00	-195.91	716.00
5	(10, 10, 30)	878.71	22.80	156.91	699.00
6	(10, 10, 30)	897.86	15.00	160.87	722.00
7	(10, 10, 30)	829.84	18.00	121.84	690.00
8	(10, 10, 30)	832.88	23.85	133.03	676.00
9	(10, 5, 34)	576.74	22.90	-147.16	701.00
10	(10, 10, 30)	858.47	17.90	146.57	694.00

对于上表的结果，我们发现与仅考虑QoS的启发式函数相比，加入了积压任务量影响的启发式函数并不影响最后的结果，这表明问题二的建模方法和结果具有极强的稳定性和鲁棒性。

6.2 敏感度分析

为了评估模型对关键参数的敏感性，我们对以下几个常量进行了系统的敏感度分析：

- α : URLLC 用户任务效用折扣系数
- 任务流:用户请求的任务量
- 热噪声谱密度: $N_0 = -174 + 10 \log_{10}(ib) + NF$ 中的-174
- 噪声系数NF: $N_0 = -174 + 10 \log_{10}(ib) + NF$ 中的NF,原值为7
- 瑞丽衰减 $h_{n,k}$: 由信号在空间中多径效应造成的小规模衰减

通过不断改变这几个常量，我们可以得出它们对于最后用户总的服务质量（Qos）的影响程度。结果如下图所示：

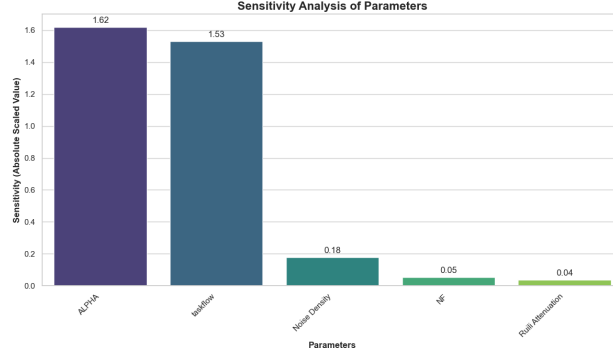


图 3: 敏感度分析结果

通过该直方图，我们可以看出URLLC 用户任务效用折扣系数 α 和用户请求的任务量对最终的服务质量影响最大，而热噪声谱密度、噪声系数NF和瑞利衰减 $h_{n,k}$ 对结果的影响则较小。这表明在实际应用中，优化用户请求的任务量和合理设置折扣系数 α 是提升系统服务质量的关键因素。

7 结论与讨论

7.1 研究结论

本研究通过对五个核心问题的分析与建模，获得了明确的量化结果，并得出以下主要结论：

- **决策收敛性：** 在问题二的求解中，研究发现当用户行为模式稳定并遵循特定概率分布时，在100ms的决策周期内，系统决策会收敛于一个特定的最优策略。
- **关键影响因素：** 模型分析表明，用户请求的数据流量以及URLLC用户的服务质量（QoS）评价函数中的折扣因子 α 是影响决策目标函数的关键变量。相比之下，其他因素（如系统噪声）的影响则不显著。
- **算法鲁棒性：** 通过对启发式函数进行变体测试，我们发现其形式对最终求解结果影响甚微。这证明了本研究采用的基于束搜索的近似贪心策略具有良好的鲁棒性。

基于以上分析，我们评估了该模型在实际应用中的潜力和局限性：

- **适用场景：** 该模型特别适用于用户行为可预测且稳定的环境，例如工业自动化、智能楼宇等。在这些场景下，模型能在保障服务质量的同时，有效提升网络资源的利用效率。
- **个性化服务：** 模型充分考虑了不同用户类型在服务质量上的差异化需求，能够实现资源的按需分配，从而提高用户满意度。
- **求解效率：** 模型对同类型用户的排队情况进行了精细建模，并采用束搜索算法求解，能够在有限的计算资源下，高效地获得近似最优的决策方案。

- **局限性：** 该模型在用户行为模式发生剧烈变化的场景（如应急通信、自然灾害响应）中适用性较差，可能导致决策失效，无法满足关键时刻的用户需求。
- **核心假设：** 模型的有效性依赖于一个核心假设，即用户请求的数据流量在100ms决策周期内服从特定分布。在实际应用中，流量模式可能受时间、地点等多种外部因素影响，分布的动态变化可能削弱决策的有效性。

7.2 未来工作展望

为进一步提升模型的实用性与先进性，未来的研究可从以下方向展开：

- **模型扩展：** 引入更多样化的用户类型和服务质量评价函数，以更精确地映射真实世界网络应用的多样化需求。
- **动态适应性：** 开发更具灵活性和自适应能力的决策算法，以应对用户行为的动态变化，特别是在突发事件和高动态场景下。
- **智能决策：** 融合机器学习与深度学习技术，特别是强化学习方法，以增强模型的环境感知、流量预测与自主决策能力。
- **实证验证：** 结合具体应用场景，开展仿真实验与物理测试，对模型的有效性、可靠性和性能进行全面的评估与验证。
- **算法优化：** 持续优化算法的计算复杂度与执行效率，以增强模型在大规模、高密度网络环境下的可扩展性与应用潜力。

8 文献

- [1] Ji M , Chen J , Liu Z , et al. Multi-level quantization and blind equalization based direct transmission method of digital baseband signal[J]. Physical Communication, 2017:S1874490717300861.
- [2] Wikipedia contributors. "Transmission time." Wikipedia, The Free Encyclopedia. Wikipedia, The Free Encyclopedia, 26 Apr. 2024. Web. 10 Aug. 2025.
- [3] Chinchilla-Romero, L., Prados-Garzon, J., Ameigeiras, P., Muñoz, P., Lopez-Soler, J. M. (2022). 5G Infrastructure Network Slicing: E2E Mean Delay Model and Effectiveness Assessment to Reduce Downtimes in Industry 4.0. Sensors, 22(1), 229. <https://doi.org/10.3390/s22010229>