

基于联合优化的网络切片资源分配策略

王再见^{1,2}, 谷慧敏^{1,2}

(1. 安徽师范大学物理与电子信息学院, 安徽 芜湖 241002; 2. 安徽省智能机器人信息融合与控制工程实验室, 安徽 芜湖 241002)

摘要: 为解决 5G 网络中各类应用差异性大对通信需求侧重点不同造成的网络资源利用率低的问题, 提出一种基于联合优化的网络切片资源分配策略, 旨在通过综合考虑切片间资源分配和切片内资源调度问题, 最大化网络资源利用率和网络收益。首先, 在切片间资源分配问题中定义一个切片用户平均满意度函数, 基于切片用户数量、切片调度时延以及切片优先级等约束, 提出基于用户服务质量 (QoS) 的比例公平资源分配算法, 以权衡各切片之间的公平性和用户需求。其次, 在切片内资源调度问题中引入服务降级和资源迁移函数, 针对拥塞和非拥塞 2 种情况为内部接入用户和外部接入用户分别建立价格模型。基于所提价格模型建立基站与用户之间的 Stackelberg 博弈, 并采用一种低复杂度的全局搜索算法求解该博弈的最佳响应, 使基站效用和用户效用最优。仿真结果表明, 所提策略能够有效提高资源利用率和网络收益, 并降低网络拥塞, 较好地实现资源分配的公平性。

关键词: 资源分配; 网络切片; 比例公平; Stackelberg 博弈

中图分类号: TN929.5

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2023089

Network slicing resource allocation strategy based on joint optimization

WANG Zaijian^{1,2}, GU Huimin^{1,2}

1. School of Physics and Electronic Information, Anhui Normal University, Wuhu 241002, China

2. Anhui Provincial Engineering Laboratory on Information Fusion and Control of Intelligent Robot, Wuhu 241002, China

Abstract: To improve network resource utilization that was decreased by different applications with different requirements in 5G networks, a network slicing resource allocation strategy based on joint optimization was proposed, which was utilized to maximize both network resource utilization and network revenue by comprehensively considering in tra-slice and inter-slice resource schedule. Firstly, the user's average satisfaction function was defined in the inter-slicing resource allocation problem. Furthermore, in terms of the number of users, slicing schedule delay and priority, a proportional fair resource allocation algorithm based on quality of service (QoS) was proposed, which was employed to achieve the best tradeoff between fairness and the users' requirements among slices. Secondly, after two functions (service degradation and resource migration) were introduced in the inter-slice resource schedule problem, two price models were established for internal access users and external access users respectively, where congestion and non-congestion conditions were analyzed. According to the proposed price models, a Stackelberg game between the base station and users was constructed, and a global search algorithm with low complexity was leveraged to obtain the best response of the game, where the best tradeoff between the base station revenue and user utility was obtained. Simulation results show that the proposed strategy can effectively improve resource utilization and network revenue while reducing network congestion. Therefore, it can better realize fairness in resource allocation.

Keywords: resource allocation, network slicing, proportional fair, Stackelberg game

收稿日期: 2022-12-06; 修回日期: 2023-02-25

基金项目: 安徽省自然科学基金资助项目 (No.2008085MF222)

Foundation Item: The Natural Science Foundation of Anhui Province (No.2008085MF222)

0 引言

5G/B5G 时代催生出大量 4K/8K 超高清视频等对带宽要求很高的应用，以及虚拟现实（VR, virtual reality）等要求时延极小的新型应用，使接入通信网络的用户（UE, user）数量和用户通信数据量呈指数式增长^[1-5]。如何根据业务场景需求提供定制化服务，对网络资源进行合理分配，灵活满足不同业务场景差异化的需求是 5G/B5G 通信系统进一步发展的关键。网络切片可以在同一个物理网络基础设施上同时支持多种具有不同性能要求的业务场景，能够有效解决业务场景差异化和用户需求多样化造成的资源浪费和网络部署成本过高的问题^[6-7]。

近年来，在网络切片资源分配方面已有大量研究工作。文献[8]提出一种基于有序加权算子的通用框架，有效提高了系统的效率和切片间资源分配的公平性，但其考虑的主要因素是用户个人满意度和系统的公平效用。文献[9]联合考虑了切片间资源分配以及切片内资源调度问题，其优化目标在于保证切片的隔离性能，平衡分配效率和服务时延。文献[10]提出一种异构云接入网络下联合拥塞控制和资源分配的网络切片动态资源调度算法，但该算法缺乏对公平性的考虑。文献[11]提出一种同时考虑切片间和切片内优先级的资源分配方案，但其优先级的设定过于简单，仅考虑用户在切片中的数据速率。文献[12]采用共享比例公平（PF, proportional fair）的方法提高资源利用率，但研究场景仅面向车辆网络切片，缺乏对业务多样性的考虑。

文献[8-12]主要针对公平性、服务时延以及吞吐量等方面进行了优化，缺乏对网络收益与用户效用之间存在利益冲突问题的考虑。博弈论是研究理性个体在利益冲突的情况下制定对策，最终达到平衡态的数学工具^[13]。文献[14-18]主要基于博弈模型解决网络资源定价和网络资源分配问题。文献[14]提出一种三级联合优化模型，保证了在跟踪信道、资源效率、效用方面最优。但该方法只针对二层业务，且仅以流量带宽作为度量用户满意度的性能指标，缺乏对用户需求多样性的考虑。文献[15]提出一种基于博弈论的自动化机制，用来协助用户根据需求及时做出决策，但该文献未考虑用户业务的差异性。文献[16]提出一种能收敛到博弈均衡的分布式算法，有效地提高了基础设施提供者（InP, in-

rastructure provider）和 UE 的收益。文献[17]提出了一种基于匹配博弈的资源分配方法，有效地降低了资源分配的响应时延和系统成本。但文献[16-17]均未考虑在网络拥塞的情况下，如何实现资源公平分配的问题。文献[18]量化了纳什均衡之间的偏差，但该文献未考虑用户对服务的满意程度，无法权衡资源分配的公平性和用户服务质量（QoS, quality of service）。

综上所述，现有网络切片资源分配方案普遍以公平性、吞吐量作为优化目标，且未充分考虑用户业务的差异性，同时缺少对资源公平分配与利润最大化共存问题的研究^[19-23]。因此本文考虑资源公平分配和利润最大化共存问题，提出一种基于联合优化的网络切片资源分配策略。本文主要的研究工作如下。

1) 对于切片间资源分配问题，本文提出一种基于 QoS 的 PF 算法。通过联合考虑切片调度时延、切片用户数量以及切片平均满意度函数，构建切片优先级，并按照优先级从高到低的顺序，依次为切片分配资源，直至剩余资源为零。该方法有效地权衡了各切片之间的公平性和用户需求。

2) 对于切片内资源调度问题，本文基于用户和切片已匹配关系，将基站（BS, base station）与 UE 之间的交互过程构建为 Stackelberg 博弈，并考虑到实际网络中可能存在的网络拥塞现象，引入服务降级和资源迁移函数，给出了基于网络拥塞状态的价格模型和资源分配策略，有效地提高了网络收益和用户效用。

3) 仿真结果表明，本文策略能够提高系统资源利用率和网络收益，降低网络拥塞，保证资源分配的公平性。与传统算法的对比验证了本文策略的有效性。

1 系统模型

本文系统模型如图 1 所示。考虑一个 5G 单小区下行链路的切片网络，为满足用户需求，以业务应用类型为单元将网络划分为 3 个切片，分别是超可靠低时延通信（URLLC, ultra-reliable and low-latency communication）、增强型移动宽带（eMBB, enhanced mobile broadband）以及大规模机器型通信（mMTC, massive machine-type communication），并假设 3 个切片的物理基础设施为同一个基站 $k = \{1, \dots, k, \dots, K\}$ 。不失一般性，假设小区中共

有 N 个用户。其中, URLLC 切片中用户数为 N_u , eMBB 切片中用户数为 N_e , mMTC 切片中用户数为 N_m , 且用户均匀分布在小区内。系统工作流程如下: UE 向 BS 申请切片资源, BS 根据 UE 需求和自身服务能力动态调整切片价格, 以最大化自身收益。

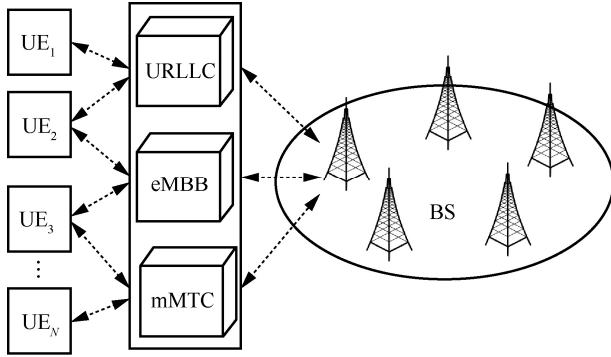


图 1 系统模型

2 问题形成

本文将网络切片资源分配问题分为切片间资源分配问题和切片内资源调度问题。

2.1 切片间资源分配问题

假设基站拥有的无线资源总数为 Q , 初始时根据网络切片类型和负载情况等, 为每个切片分配满足最低要求的资源。其中, eMBB、URLLC 和 mMTC 切片分配到的资源总数分别为 Q_1 、 Q_2 和 Q_3 , 此时, 基站剩余的资源数为 Q' 。

给各切片分配资源之后, 选择合适的调度算法对剩余资源数进行再分配。公平性和吞吐量是相互矛盾的, 往往一方性能的提高会引起另一方性能的下降, 从而使网络资源难以得到最大化利用。PF 算法为经典的资源调度算法, 一定程度上考虑了吞吐量和公平性的权衡。其基本思想是通过引入用户的历史吞吐量来计算优先级。但由于各个切片服务的业务存在较大差异, 切片间的资源分配不能只考虑吞吐量和公平性。因此, PF 算法难以满足 5G 多样化的业务需求。

针对上述问题, 本文结合切片特性, 定义了切片用户平均满意度函数, 提出基于 QoS 的 PF 资源分配算法 (以下简称 PF-Q 算法) 对剩余资源数进行再分配。首先, 根据 PF-Q 算法计算出各切片的优先级; 然后, 按照优先级从高到低的顺序, 依次为各切片分配剩余资源数 Q' , 直至剩余资源数 Q' 为零。

2.1.1 切片用户平均满意度函数

不同切片的服务类型不同, 对速率或时延要求也不同, 即不同切片的满意度函数存在差异。eMBB 切片旨在实现高数据速率, 能容忍一定的时延, 包差错率 (PER, packet error rate) 为 10^{-3} 左右。假设 eMBB 切片所服务用户 $n_e \in N_e$ 从基站 k 处获得的实际传输速率为 $R_{n_e}^k$, 最小速率要求为 $R_{n_e}^{\min}$ 。则用户 n_e 对速率敏感的 eMBB 切片的平均满意度为

$$U_{n_e}^k = \frac{1}{N_e} \sum_{n_e \in N_e} 1 - \alpha e^{-\beta \left(\frac{R_{n_e}^k}{R_{n_e}^{\min}} \right)} \quad (1)$$

其中, $\alpha (0 \leq \alpha \leq 1)$ 是常数系数, β 是 Sigmoid 函数的斜率。

URLLC 切片所需的数据速率通常较低, 旨在通过低时延降低 PER。假设 URLLC 切片所服务用户 $n_u \in N_u$ 发送一个数据分组至基站 k 的时间为 $T_{n_u}^k$, 最大时延需求为 $T_{n_u}^{\max}$ 。则用户 n_u 对时延敏感的 URLLC 切片的平均满意度为

$$U_{n_u}^k = \frac{1}{N_u} \sum_{n_u \in N_u} 1 - \alpha e^{-\beta \left(\frac{T_{n_u}^k}{T_{n_u}^{\max}} \right)} \quad (2)$$

mMTC 切片旨在为海量连接、小数据包、低成本、低功耗的设备提供有效的连接方式, 其目标 PER 远大于 eMBB 和 URLLC。假设 mMTC 切片所服务用户 $n_m \in N_m$ 在单位时间内的 PER 为 $P_{n_m}^k$, 最大 PER 为 $P_{n_m}^{\max}$ 。则用户 n_m 对 mMTC 切片的平均满意度为

$$U_{n_m}^k = \frac{1}{N_m} \sum_{n_m \in N_m} 1 - \alpha e^{-\beta \left(\frac{P_{n_m}^k}{P_{n_m}^{\max}} \right)} \quad (3)$$

为了方便后续讨论, 定义 $x_{n_i}^k$ 为基站 k 中切片 i 的 QoS 因子, $y_{n_i}^k$ 为 QoS 阈值, 则式(1)~式(3)可改写为

$$U_{n_i}^k = \frac{1}{N_i} \sum_{n_i \in N_i} 1 - \alpha e^{-\beta \left(\frac{x_{n_i}^k}{y_{n_i}^k} \right)} \quad (4)$$

2.1.2 基于 QoS 的 PF 资源分配算法

基于上述理论, PF-Q 算法的切片优先级定义为

$$F_{(i, n_i)}^k(t) = U_{n_i}^k \frac{r_{(i, n_i)}^k(t)}{R_{(i, n_i)}^k(t)} \ln(1 + d_i^k) \quad (5)$$

其中，将切片 i 中的所有用户看作一个整体， $r_{(i,n_i)}^k(t) = \sum_{n_i \in N_i} r_{n_i}^k(t)$ 为 t 时刻切片 i 中所有用户的瞬时速率， $R_{(i,n_i)}^k(t) = \sum_{n_i \in N_i} R_{n_i}^k(t)$ 为 t 时刻前 2 个时隙切片 i 中用户 n_i 的平均传输速率， $r_{n_i}^k(t)$ 为 t 时刻用户 n_i 的瞬时速率， $R_{n_i}^k(t)$ 为 t 时刻前 2 个时隙用户 n_i 的平均传输速率。

$R_{(i,n_i)}^k(t)$ 的更新式为

$$R_{(i,n_i)}^k(t) = \begin{cases} \sum_{n_i} \left[\left(1 - \frac{1}{T}\right) R_{n_i}^k(t-1) + \frac{1}{T} r_{n_i}^k(t) \right], & i \text{ 被调度} \\ \sum_{n_i} \left[\left(1 - \frac{1}{T}\right) R_{n_i}^k(t-1) \right], & \text{其他} \end{cases} \quad (6)$$

其中， T 为时间窗。

d_i 为切片 i 在调度单元中的等待时间。当有 2 个或多个切片具有相同优先级时，系统应优先为等待时间长的切片提供服务。通过在算法中引入 $\ln(1+d_i)$ 因子，可以更好地权衡切片之间的公平性和 QoS 需求。根据式(5)计算出每个切片的优先级后，按照优先级顺序对未分配的资源进行再分配。更新 eMBB、URLLC 和 mMTC 切片分配到的资源总数分别为 Q_1 、 Q_2 和 Q_3 。

2.2 切片内资源调度问题

在确定了切片 i 在基站 k 上的资源分布后，基于用户和切片已匹配关系，执行切片内资源调度。

2.2.1 资源模型

假设基站 k 分配给切片 i 的资源为 $\lambda_i^k Q$ ， λ_i^k 为分配比例。本节考虑访问切片 i 的用户包括 $n_i = \{1, 2, \dots, N_i\}$ 个内部接入用户和 $l_i = \{1, 2, \dots, L_i\}$ 个外部接入用户。所有用户都可以根据自身需求和网络状态等因素动态地向基站 k 申请购买所属切片的资源。对于内部接入用户 n_i 而言，其获得的网络接入吞吐量 $C_{n_i}^k$ 与再分配的频谱资源有关，计算式为

$$C_{n_i}^k = x_{n_i}^k \lambda_i^k Q \text{lb} \left(1 + \frac{P_{n_i}^k |f_{n_i}^k|^2}{S_0} \right) \quad (7)$$

其中， $x_{n_i}^k$ 是切片为内部用户分配资源的分配比例， $P_{n_i}^k$ 是基站 k 为切片 i 中的内部接入用户 n_i 分配的功率， $|f_{n_i}^k|^2$ 是信道增益， S_0 是噪声功率。

2.2.2 价格模型

基站制定的切片价格应该结合市场规律，随着用户需求增加，切片价格应该提高，即切片价格和用户需求有关。结合实际生活，本文分析了网络拥塞和非拥塞 2 种情况下的切片价格。

情况 1 用户对资源的需求量小于或等于切片资源拥有量，即 $\sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i \in L_i} C_{l_i}^k \leq \lambda_i^k Q$ 。

情况 2 用户对资源的需求量大于切片资源拥有量，即 $\sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i \in L_i} C_{l_i}^k > \lambda_i^k Q$ 。

1) 情况 1 中，受市场规律影响下的内部接入用户和外部接入用户切片价格分别为

$$H_k^{n_i} = \frac{\omega_i^k \lambda_i^k Q}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} h_k^{n_i} \quad (8)$$

$$H_k^{l_i} = \frac{\omega_i^k \lambda_i^k Q}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} h_k^{l_i} \quad (9)$$

其中， w_i^k 是价格系数，且为常数； ε 是一个很小的正整数，用来避免切片资源极度紧张以至资源剩余量 $\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k = 0$ ，即分母为零的情况发生； $h_k^{n_i}$ 和 $h_k^{l_i}$ 是基站 k 为切片 i 上内部接入用户 n_i 和外部接入用户 l_i 制定的切片价格。

2) 情况 2 中，本文定义了服务降级和资源迁移函数。特别地，由于基站需要保证内部接入用户 n_i 的基础吞吐量 $C_{n_i}^{\text{min}}$ ，因此，内部接入用户 n_i 不存在服务降级和资源迁移的情况。具体分析如下。

① 服务降级

考虑到网络切片具有弹性，即使在切片资源紧张的情况下，切片仍然可以运作，但切片用户的体验会变差。本文将上述现象称为服务降级，服务降级与切片资源需求量和切片资源拥有量之差有关，差异越大，降级越明显。因此，定义外部接入用户 l_i 的服务降级函数为

$$d_{l_i}^k = 1 - \frac{\lambda_i^k Q}{\sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i' \in L_i, l_i' \neq l_i} C_{l_i'}^k} \quad (10)$$

其中， $d_{l_i}^k$ ($0 \leq d_{l_i}^k \leq 1$) 为常数，且当 $\lambda_i^k Q = 0$ 时， $d_{l_i}^k = 1$ ，服务降级最大；当 $\lambda_i^k Q = \sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i' \in L_i, l_i' \neq l_i} C_{l_i'}^k$

时, $d_{l_i}^k = 0$, 服务无降级。

② 资源迁移

假设在某个时刻, eMBB 切片需求量锐减, 而 URLLC 切片需求量激增。由于 URLLC 切片没有足够的资源支撑其激增的需求量, 此时如果把资源相对过剩的 eMBB 切片已获得的部分资源归还给基站, 再由基站把这部分归还的资源分配给急需增加资源的 URLLC 切片, 那么 URLLC 切片将可以获得足够的资源以支撑其用户需求, 而 eMBB 切片的剩余资源也可以得到充分的利用, 从而提高整个网络的收益。本文将上述现象称为资源迁移, 由上述分析可知, 资源迁移代价与切片资源迁移量有关, 资源迁移量越大, 迁移代价越大。因此, 定义外部接入用户的资源迁移函数为

$$v_{l_i}^k = 1 - \frac{\sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i' \in L_i, l_i' \neq l_i} C_{l_i'}^k - \lambda_i^k Q}{C_{i,j}^k}$$

$$\text{s.t. } C_{i,j}^k > \sum_{n_i \in N_i} C_{n_i}^k + \sum_{l_i' \in L_i, l_i' \neq l_i} C_{l_i'}^k - \lambda_i^k Q \quad (11)$$

其中, $C_{i,j}^k$ 表示当前切片 i 从切片 j 中获得的迁移资源数。结合上述对服务降级和资源迁移现象的分析, 在情况 2 中定义拥塞代价函数为

$$\text{cost}_{l_i}^k = d_{l_i}^k + \eta_{i,j}^k v_{l_i}^k \quad (12)$$

其中, $\eta_{i,j}^k \in \{0,1\}$, $\eta_{i,j}^k = 1$ 表示当前切片 i 从切片 j 中获得迁移资源。综上, 在情况 2 中, 基站 k 为切片 i 上的外部接入用户 l_i 制定的切片价格为

$$H_k^{l_i'} = \left| \frac{\omega_i^k \lambda_i^k Q \text{cost}_{l_i}^k}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} \right| h_k^{l_i} \quad (13)$$

2.2.3 优化目标

1) 用户优化目标

对于内部用户 n_i 而言, 其效用函数包括两部分: 一部分是接入切片 i 获得的收益, 另一部分是购买切片 i 的支出。因此, 内部用户 n_i 的效用函数表示为

$$E_{n_i}^k = \delta_{n_i} \ln \left(1 + \sum_{k \in K} U_{n_i}^k (C_{n_i}^k - C_{n_i}^{\min}) \right) - \sum_{k \in K} C_{n_i}^k H_k^{n_i}$$

$$\text{s.t. } E_{n_i}^k \geq 0$$

$$C_{n_i}^{\min} \leq C_{n_i}^k \leq C_{n_i}^{\max} \quad (14)$$

其中, δ_{n_i} 表示内部接入用户 n_i 的收益系数, $C_{n_i}^{\max}$ 表示内部接入用户 n_i 请求的最大网络接入吞吐量。

基站对外部用户 l_i 收费更高且不保证外部用户 l_i 的基础吞吐量 $C_{l_i}^{\min}$ 要求。因此, 定义外部用户 l_i 的效用函数 $E_{l_i}^k$ 为

$$E_{l_i}^k = \begin{cases} \delta_{l_i} \ln \left(1 + \sum_{k \in K} \phi_{l_i}^k U_{l_i}^k C_{l_i}^k \right) - \sum_{k \in K} C_{l_i}^k (\tau H_k^{l_i}), & \text{情况1} \\ \delta_{l_i} \ln \left(1 + \sum_{k \in K} \phi_{l_i}^k U_{l_i}^k C_{l_i}^k \right) - \sum_{k \in K} C_{l_i}^k (\tau H_k^{l_i'}), & \text{情况2} \end{cases}$$

$$\text{s.t. } E_{l_i}^k \geq 0$$

$$C_{l_i}^{\min} \leq C_{l_i}^k \leq C_{l_i}^{\max} \quad (15)$$

其中, δ_{l_i} 表示外部接入用户 l_i 的收益系数; $\phi_{l_i}^k \in \{0,1\}$, $\phi_{l_i}^k = 1$ 表示切片 i 为用户分配的资源; τ 为调整系数, 控制外部接入用户 l_i 价格函数的变化; $C_{l_i}^{\max}$ 表示外部接入用户 l_i 请求的最大的网络接入吞吐量。

用户的优化目标是在考虑体验质量的前提下, 最大化自身收益。因此, 对于用户而言, 其资源优化问题可映射为收益最大化问题。由于内部接入用户与外部接入用户的效用函数类似, 故后续分析中只考虑外部接入用户。将式(4)、式(9)、式(13)代入式(15), 则外部接入用户 l_i 效用函数 $E_{l_i}^k$ 可转化为

$$E_{l_i}^k = \delta_{l_i} \ln \left[1 + \sum_{k \in K} \phi_{l_i}^k C_{l_i}^k \left(\frac{1}{N_i} \sum_{l_i \in L_i} 1 - \alpha e^{-\beta \left(\frac{x_{n_i}^k}{y_{n_i}^k} \right)} \right) \right] - \sum_{k \in K} C_{l_i}^k \left| \frac{\omega_i^k \lambda_i^k Q (\text{cost}_{l_i}^k)^\theta h_k^{l_i} \tau}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} \right| \quad (16)$$

其中, $\theta \in \{0,1\}$, $\theta = 0$ 表示情况 1, $\theta = 1$ 表示情况 2。

因此, 外部接入用户对应的收益最大化问题为

$$\max_{C_{l_i}^1, \dots, C_{l_i}^k, \dots, C_{l_i}^K} E_{l_i} (C_{l_i}^1, \dots, C_{l_i}^k, \dots, C_{l_i}^K, \mathbf{h}^{l_i}) \quad (17)$$

其中, $\mathbf{h}^{l_i} = \{h_1^{l_i}, \dots, h_k^{l_i}, \dots, h_K^{l_i}\}$ 。

2) 基站优化目标

基站的效用函数也包括两部分: 一部分是制定切片价格带来的收益, 另一部分是构建切片时所支出的成本。因此, 基站的效用函数表示为

$$E_k^i = \begin{cases} \sum_{n_i \in N_i} C_{n_i}^k H_k^{n_i} - \sum_{n_i \in N_i} C_{n_i}^k F_k^{n_i} + \sum_{l_i \in L_i} C_{l_i}^k (\tau H_k^{l_i}) - \sum_{l_i \in L_i} C_{l_i}^k F_k^{l_i}, & \text{情况1} \\ \sum_{n_i \in N_i} C_{n_i}^k H_k^{n_i} - \sum_{n_i \in N_i} C_{n_i}^k F_k^{n_i} + \sum_{l_i \in L_i} C_{l_i}^k (\tau H_k^{l_i}) - \sum_{l_i \in L_i} C_{l_i}^k (F_k^{l_i} + \gamma_{l_i}^k \cos t_{l_i}^k), & \text{情况2} \end{cases} \quad (18)$$

其中, $F_k^{n_i}$ 和 $F_k^{l_i}$ 分别表示基站 k 为内部接入用户 n_i 和外部接入用户 l_i 提供切片服务所支出的成本; $\gamma_{l_i}^k \cos t_{l_i}^k$ 表示在情况 2 中, 外部接入用户 l_i 因服务降级或服务迁移造成的代价成本, $\gamma_{l_i}^k$ 为代价成本系数。同上, 基站的效用函数可简化为

$$E_k^i = \sum_{n_i \in N_i} C_{n_i}^k \frac{\omega_i^k \lambda_i^k Q h_k^{n_i}}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} - \sum_{n_i \in N_i} C_{n_i}^k F_k^{n_i} + \sum_{l_i \in L_i} C_{l_i}^k \left(\frac{\omega_i^k \lambda_i^k Q (\cos t_{l_i}^k)^\theta h_k^{l_i} \tau}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} \right) - \sum_{l_i \in L_i} C_{l_i}^k (F_k^{l_i} + \theta \gamma_{l_i}^k \cos t_{l_i}^k) \quad (19)$$

由于基站的优化目标是在考虑服务能力的前提下, 最大化自身收益^[24]。因此, 其资源优化问题也可以映射为收益最大化问题, 即

$$\max_{h_k^i} E_k^i (h_k^i, C_{l_i}^k) \quad (20)$$

2.2.4 博弈构成与求解

本文根据对系统模型的分析以及 Stackelberg 博弈的相关要素, 分别定义博弈方、策略和效用如下。

1) 领导者: 基站。领导者决定切片 i 的定价并影响跟随者的决策。

2) 跟随者: 所有访问切片 i 的用户。用户考虑基站的切片定价和自身吞吐量需求, 调整切片购买策略, 最大化自身收益。

3) 策略: 外部接入用户的策略是向基站申请购买的切片资源数 $C_{l_i} = \{C_{l_i}^1, \dots, C_{l_i}^k, \dots, C_{l_i}^K\}$, 基站的策略是制定切片 i 的价格 $h^i = \{h_1^i, \dots, h_k^i, \dots, h_K^i\}$ 。

4) 效用: 外部接入用户的效用函数用 E_i^k 表示, 基站的效用函数用 E_k^i 表示。

为了便于理解和分析, 本文在博弈过程中不考虑业务时变的场景。本文采用全局搜索算法来确定该博弈的纳什均衡点, 该算法的时间复杂度为 $O(M)$, 其包括确定最优值的最大迭代步骤 M 。该算法能够根据用户需求和单位带宽价格

建立一个搜索表并进行更新, 通过直接从表中搜索最佳响应, 能够快速完成新的算法过程, 达到平衡。

均衡点可以通过个体最优求得, 令 $C_{l_i}^{k*}$ 和 $h_k^{l_i*}$ 分别为式(16)和式(19)的最佳响应, 定义纳什均衡点为^[25]

$$E_{l_i}^k (C_{l_i}^*, h^{l_i*}) \geq E_{l_i}^k (C_{l_i}, h^{l_i*}) \quad (21)$$

$$E_k^i (h_k^{l_i*}, C_{l_i}^{k*}) \geq E_k^i (h_k^{l_i}, C_{l_i}^{k*}) \quad (22)$$

其中, $C_{l_i} = \{C_{l_i}^1, \dots, C_{l_i}^k, \dots, C_{l_i}^K\}$ 。

本文采用逆向归纳法来求解博弈。首先, 求解 Stackelberg 博弈的第二阶段, 得到用户的最优吞吐量需求 $C_{l_i}^{k*}$ 。然后, 将 $C_{l_i}^{k*}$ 用于求解博弈的第一阶段, 得到基站对于切片 i 的最优定价 $h_k^{l_i*}$ 。本文以用户和基站的一次交互为例, 讨论纳什均衡求解过程。

① 用户最佳响应

假设基站已做出决策, 即给定切片 i 的价格为 h_k^i , 对于给定的价格 h_k^i 和 $C_{l_i}^{\min} \leq C_{l_i}^k \leq C_{l_i}^{\max}$, 用户的效用函数是凹函数。这可以通过计算用户效用函数的二阶导数得到

$$\frac{\partial E_{l_i}^k}{\partial C_{l_i}^k} = \frac{\delta_{l_i}^k \phi_{l_i}^k U_{l_i}^k}{1 + \sum_{k \in K} \phi_{l_i}^k C_{l_i}^k U_{l_i}^k} - \left(\frac{\omega_i^k \lambda_i^k Q (\cos t_{l_i}^k)^\theta h_k^{l_i} \tau}{\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon} \right) - \frac{C_{l_i}^k \omega_i^k \lambda_i^k Q (\cos t_{l_i}^k)^\theta h_k^{l_i} \tau}{\left(\lambda_i^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon \right)^2} \quad (23)$$

为体现多轮交互对用户需求的影响, 本文将用户 l_i 在上一轮交互中向基站 $k' \in K, k' \neq k$ 请求的资源, 视为用户 l_i 在本轮交互中向基站 k' 请求的资源。同时, 将用户 $l_i' \in L_i, l_i' \neq l_i$ 在上一轮交互中向基站 $k \in K$ 请求的资源, 视为用户 l_i' 在本轮交互中向基站 k 请求的资源。因此, 得到二阶导数为

$$\frac{\partial E_i^k}{(\partial C_{l_i}^k)^2} = -\frac{\delta_{l_i} (\phi_{l_i}^k U_{l_i}^k)^2}{\left(1 + \sum_{k \in K} \phi_{l_i}^k C_{l_i}^k U_{l_i}^k\right)^2} - \frac{\omega_{i,k} \lambda_{i,k} Q (\cos t_{l_i}^k)^\theta h_k^l \tau}{\left(\lambda_{l_i}^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon\right)^2} - \frac{\omega_{i,k} \lambda_{i,k} Q (\cos t_{l_i}^k)^\theta h_k^l \tau}{\left(\lambda_{l_i}^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon\right)^3} \cdot \frac{\lambda_{i,k} Q - \sum_{n_i \in N_i} C_{n_i}^k + \varepsilon - \sum_{l_i \in L_i, l_i' \neq l_i} C_{l_i'}^k + C_{l_i}^k}{\left(\lambda_{l_i}^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i} C_{l_i}^k + \varepsilon\right)^3} < 0 \quad (24)$$

式(24)表明二阶导数为负，证明了用户效用函数是凹函数。令一阶导数为零，得到用户最佳响应 $C_{l_i}^{k*}$ 为

$$C_{l_i}^{k*} = q_{l_i}^k + \frac{\zeta_{l_i}^k q_{l_i}^k h_k^l}{2\delta_{l_i}} + \sqrt{\left(\frac{\zeta_{l_i}^k q_{l_i}^k G_{l_i}^{k'} h_k^l}{\delta_{l_i} \phi_{l_i}^k U_{l_i}^k} - (q_{l_i}^k)^2\right) + \frac{(2\delta_{l_i} q_{l_i}^k + \zeta_{l_i}^k q_{l_i}^k h_k^l)^2}{4(\delta_{l_i})^2}} \quad (25)$$

其中， $\zeta_{l_i}^k$ 、 $q_{l_i}^k$ 和 $G_{l_i}^{k'}$ 分别为

$$\begin{aligned} \zeta_{l_i}^k &= \omega_{i,k} \lambda_{i,k} Q (\cos t_{l_i}^k)^\theta \tau \\ q_{l_i}^k &= \lambda_{l_i}^k Q - \sum_{n_i \in N_i} C_{n_i}^k - \sum_{l_i \in L_i, l_i' \neq l_i} C_{l_i'}^k + \varepsilon \\ G_{l_i}^{k'} &= 1 + \sum_{k' \in K, k' \neq k} \phi_{l_i}^{k'} C_{l_i}^{k'} U_{l_i}^{k'} \end{aligned}$$

② 基站最佳响应

得到用户的最佳响应 $C_{l_i}^{k*}$ 后，本文提出一种低复杂度的全局搜索算法，用于求解基站对于切片 i 的最优定价 $h_k^{l_i*}$ ，具体步骤如算法 1 所示。

算法 1 基于全局搜索的最优定价更新算法

- 1) 输入其他初始化参数
- 2) 初始化 $E_k^l = E_{l_i}^k = C_{l_i}^{k*} = h_k^{l_i*} = 0$
- 3) 计算 $(h_k^l)_{\min}$ 和 $(h_k^l)_{\max}$ ，详见附录 1
- 4) 求解式(18)关于 h_k^l 的偏导数， $\frac{\partial E_k^l}{\partial h_k^l} < 0$ 表明

E_k^l 是可微的实函数，因此 E_k^l 在闭区间 $[(h_k^l)_{\min}, (h_k^l)_{\max}]$ 内存在最大值，详见附录 2

- 5) 迭代
- 6) let $H = (h_k^l)_{\min} : j : (h_k^l)_{\max}$
- 7) 循环
- 8) for $i = 1 : \text{length}(H)$
- 9) set $h = H(i)$;
- 10) 根据式(18)计算 $E_k^l(h)$
- 11) if $E_k^l(h) > E_k^l$
- 12) 更新 $E_k^l = E_k^l(h)$
- 13) set $h_k^{l_i*} = h$
- 14) 根据式(24)计算 $C_{l_i}^{k*}$
- 15) 根据式(15)求解 $E_{l_i}^k$
- 16) until $E_k^l(h) \leq E_k^l$ ，输出 E_k^l 、 $E_{l_i}^k$ 、 $C_{l_i}^{k*}$ 、 $h_k^{l_i*}$
- 17) end if
- 18) end for

3 仿真分析

为验证本文所提策略的有效性，本节首先对本文提出的 PF-Q 算法与 PF 算法、轮询 (RR, round robin) 算法和最大载干比 (Max C/I) 算法分别在资源分配比例和资源分配的公平性两方面进行对比。仿真参数如表 1 所示。

表 1 仿真参数

仿真参数	参数值
基站数量/个	2
切片数/个	3
用户数/个	100
用户分布	随机撒点
带宽/MHz	5

随机 T 时隙中不同资源分配算法下不同切片被分配的资源总数如图 2 所示。由图 2 可知，每个切片在不同时隙中被分配的资源总数是不同的，但 4 种算法都将较多的资源分配给用户传输数据量比较大的 eMBB 切片，而对 mMTC 切片和 URLLC 切片的分配存在一定差异。这是因为 mMTC 切片所服务的用户数虽多但数据量小，URLLC 切片所服务的用户虽然对实时性要求严

格，但数据量远没有 eMBB 切片大。由此可见，4 种算法都能够有效地根据用户数据量对资源进行分配。

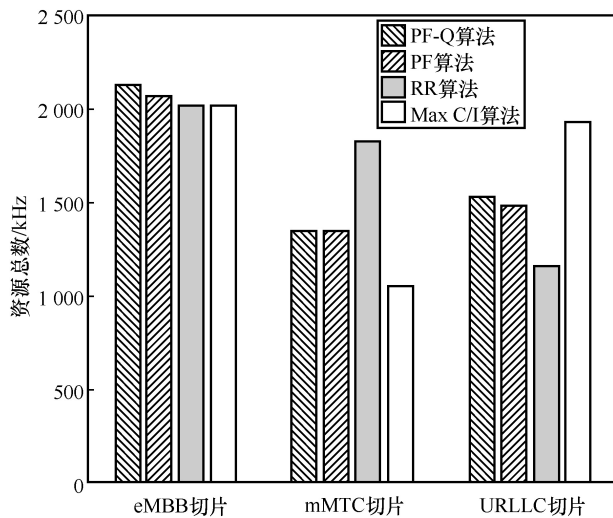


图2 不同资源分配算法下不同切片被分配的资源总数

从图2中还可以看出，RR算法和Max C/I算法分配给各切片的资源差异较大。在RR算法中分配给URLLC切片的资源少于其他3种算法，这是因为RR算法按照用户申请资源的先后顺序，在相同的时间长度下依次循环占用系统资源，以保证每个用户都获得相同的调度机会。而URLLC切片中用户数量少且传输数据量小，因此RR算法分配给URLLC切片的资源最少。Max C/I算法与RR算法类似，其分配给mMTC切片的资源同样少于其他3种算法。这是因为Max C/I算法受信道条件影响，如果切片中的用户多数靠近基站，信道条件好则优先级高，分配的资源数就多；如果切片中用户多数分布在小区边缘，信道条件差则优先级低，分配的资源就少。PF算法和PF-Q算法的分配结果类似，虽然这2种算法注重公平性，但每个切片上的资源分配数也有不同，因为mMTC切片中的业务大部分时间不需要分配资源，导致该切片分配的资源比其他切片少。经过观察多次运行的仿真结果，由于PF-Q算法引入了切片用户平均满意度和调度时延因子，因此PF-Q算法在公平性方面要略优于PF算法。

不同算法吞吐量公平性指数的对比如图3所示。从图3可以看出，PF-Q算法和PF算法的吞吐量公平性指数都随着调度时隙的增加而增大，但PF-Q算法的公平性指数整体大于PF算法。而RR

算法和Max C/I算法的吞吐量公平性指数随着时间的增加都有所减小，但RR算法的公平性指数始终是最大的，这是因为在RR算法中所有的用户具有相同机会来获得系统资源；而Max C/I算法在每一时刻调度都会选择信道条件最好的用户，以牺牲公平性实现吞吐量最大化，因此Max C/I算法的公平性最低。虽然RR算法公平性最高，但RR算法没有考虑信道状况条件，数据传输的可靠性并不高，而本文所提的PF-Q算法能够较好地权衡公平性和吞吐量之间的动态平衡，因此所提PF-Q算法优于其他3种算法。

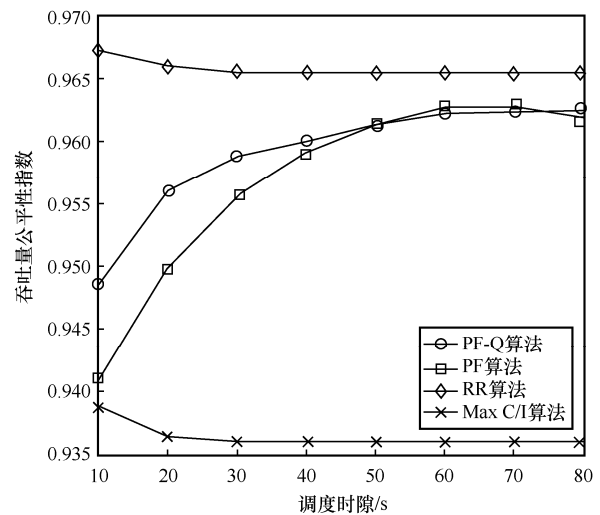


图3 不同算法吞吐量公平性指数的对比

本文在确定切片 i 在基站上的资源分配比例后，基于混合条件执行片内资源调度，分别从用户效用、基站效用和纳什均衡三方面验证本文所提的基于联合优化的Stackelberg博弈的可行性。仿真网络环境包括2个基站、20个内部接入用户以及40个外部接入用户，网络中切片 i 的最大带宽频谱资源为 $\lambda_i^*Q=2130$ kHz，每个内部接入用户可租用的最大频谱带宽为 $C_i^{\max}=100$ kHz，每个外部接入用户可租用的最大频谱带宽为 $C_i^{\max}=180$ kHz。

3种情况下用户效用与购买带宽之间的关系如图4所示。从图4可以看出，用户效用随购买带宽的增加而逐渐减小，这是由于随着带宽的增加，单位带宽价格会上升，从而导致用户效用减小。从图4中还可以看出，当 δ_i 相等时，用户效用会随着 τ 的降低而增大；而当 τ 相等时，用户效用随着 δ_i 的增加而增大。

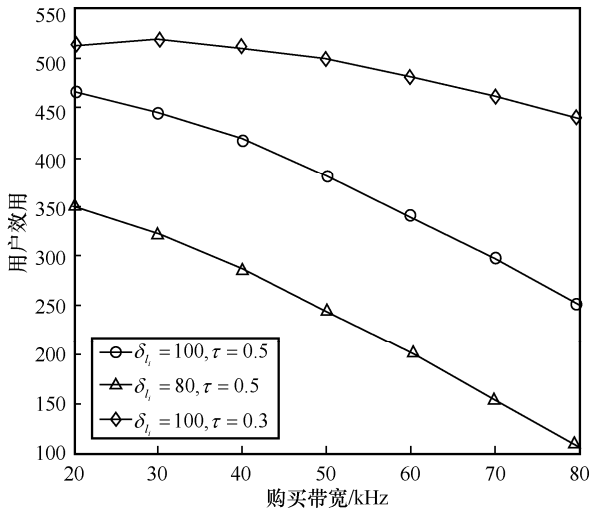


图 4 用户效用与购买带宽之间的关系

在网络拥塞和非拥塞 2 种情况下单位带宽价格和用户数之间的关系如图 5 所示。从图 5 可以看出，随着用户数的增加，2 种情况下的单位带宽价格都呈上升趋势，这也反映了本文中基站制定的切片价格是符合市场规律的，即随着用户需求增加，切片价格应该提高。此外，在拥塞情况下的单位带宽价格比非拥塞情况下的单位带宽价格更高，这不仅表明了基站可以通过提高单位带宽价格有效降低网络拥塞，还可以提高基站的收益，从而最大化网络收益。

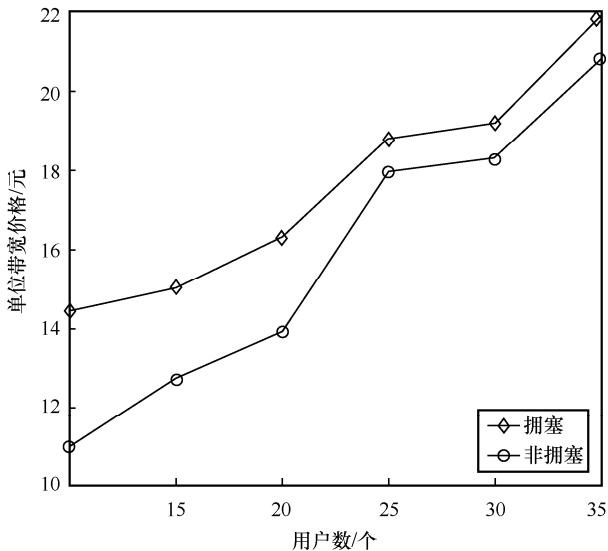


图 5 单位带宽价格和用户数之间的关系

图 6 证明了纳什均衡的存在。从图 6 中可以看出，随着单位带宽价格的增加，用户效用和基站效用都增加后减少。这是由于对于用户而言，一方面

价格的提升会提高用户的满意度从而增加用户效用，另一方面价格的提升会导致用户支出的成本增加从而减少用户效用，因此用户效用呈先增加后减少趋势；对于基站而言，提升价格会增加其收益，但价格提升过高，会导致用户流失数增多，从而引起基站效用降低，因此基站效用也呈先增加后减少的趋势。

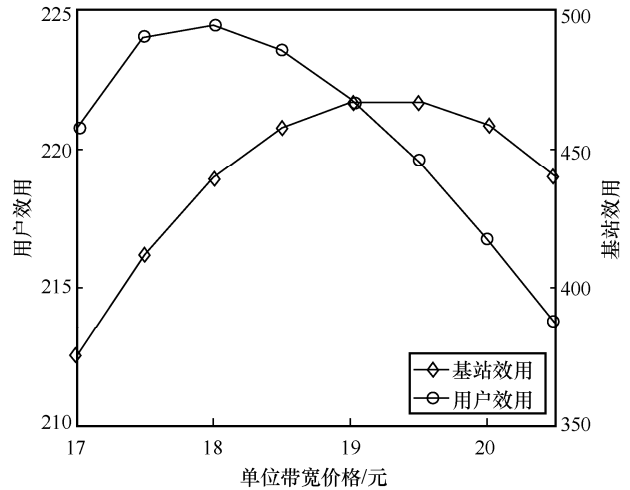


图 6 纳什均衡的证明

图 7 和图 8 分别比较了所提基于联合优化的 Stackelberg 博弈、文献[3]中基于随机定价的 Stackelberg 博弈、文献[11]中基于信道定价的 Stackelberg 博弈与传统的非博弈方法的效用。从图 7 和图 8 中可以看出，随着带宽的增加，基站效用和用户效用都呈先增加后减少的趋势，且本文所提基于联合优化的 Stackelberg 博弈的效用明显大于另外 3 种方法。这是因为基于随机定价的 Stackelberg 博弈和传统的非博弈方法均采用随机定价，其结果不具备确定性和稳定性。在基于随机定价的 Stackelberg 博弈中，虽然价格随机，但为了保证效用为正，单位带宽价格必须大于某一值，其价格有一定范围，而传统的非博弈方法中价格完全随机。基于信道定价的 Stackelberg 博弈其价格受信道质量影响，该模型具有提高信道跟踪能力和保持切片隔离性的优点。但从图 7 和图 8 可以看出，所提基于联合优化的 Stackelberg 博弈的基站效用和用户效用均大于基于信道定价的 Stackelberg 博弈。这是由于基于信道定价的 Stackelberg 博弈中没有考虑价格随用户需求变化而变化，导致价格制定不合理，使效用减少。

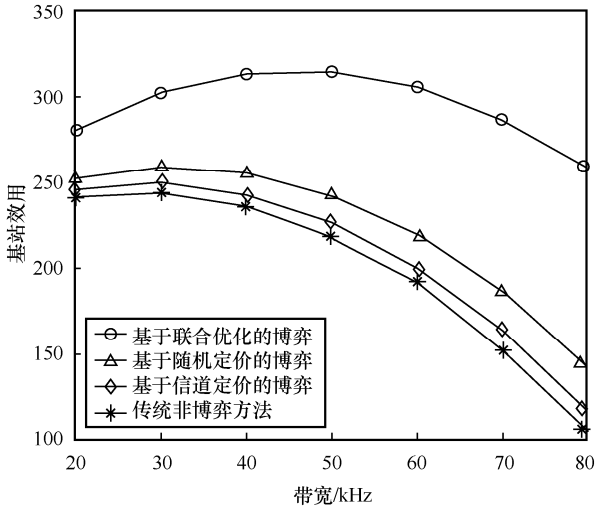


图 7 不同策略下基站效用的对比

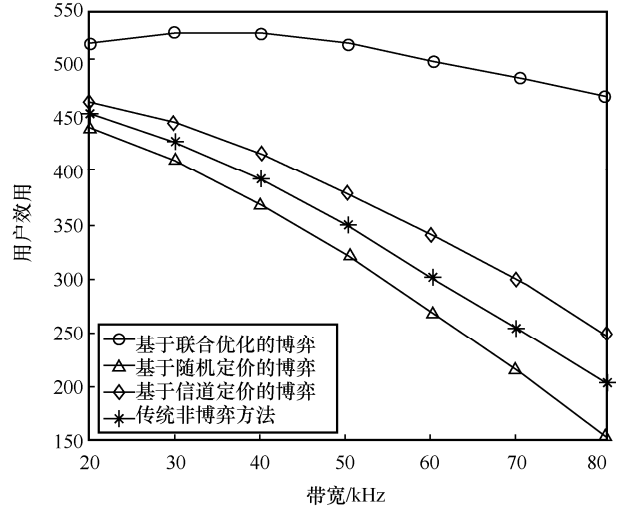


图 8 不同策略下用户效用的对比

4 结束语

本文对基于联合优化的网络切片资源分配策略进行研究。通过综合考虑切片间资源分配和切片内资源调度问题,建立一个以最大化网络资源利用率和网络收益为目标的资源分配模型。在切片间资源分配问题中,定义切片用户平均满意度函数和切片优先级,

提出基于 QoS 的比例公平资源分配方法,权衡了各切片间的公平性和用户需求。在切片内资源分配问题中,基于用户和切片已匹配关系,构建基站与用户之间的博弈,给出基于拥塞控制的定价和分配算法,使基站效用和用户效用达到最优。仿真结果表明,本文策略能够有效提高资源利用率和网络收益,降低网络拥塞,保证资源分配的公平性。

附录 1 价格区间证明

将 $C_{l_i}^{\min}$ 和 $C_{l_i}^{\max}$ 代入式(25), 可得

$$(h_k^l)_{\min} = \frac{4(C_{l_i}^{\min} - q_{l_i}^k)(\delta_{l_i})^3 + q_{l_i}^k}{\zeta_{l_i}^k q_{l_i}^k (4(\delta_{l_i})^2 - 1)} + \frac{4G_{l_i}^{k'}(\delta_{l_i})^2}{\varphi_{l_i}^k U_{l_i}^k \zeta_{l_i}^k q_{l_i}^k \sigma_{l_i}} + \sqrt{\left[\frac{4(C_{l_i}^{\min} - q_{l_i}^k)(\delta_{l_i})^3 + q_{l_i}^k}{\zeta_{l_i}^k q_{l_i}^k (4(\delta_{l_i})^2 - 1)} + \frac{4G_{l_i}^{k'}(\delta_{l_i})^2}{\varphi_{l_i}^k U_{l_i}^k \zeta_{l_i}^k q_{l_i}^k \sigma_{l_i}} \right]^2 + \left[-\frac{16(\delta_{l_i})^4 (C_{l_i}^{\min} - q_{l_i}^k)^2}{(\zeta_{l_i}^k q_{l_i}^k)^2 \sigma_{l_i}} - \frac{1}{(\delta_{l_i})^2} \right]} \quad (26)$$

$$(h_k^l)_{\max} = \frac{4(C_{l_i}^{\max} - q_{l_i}^k)(\delta_{l_i})^3 + q_{l_i}^k}{\zeta_{l_i}^k q_{l_i}^k (4(\delta_{l_i})^2 - 1)} + \frac{4G_{l_i}^{k'}(\delta_{l_i})^2}{\varphi_{l_i}^k U_{l_i}^k \zeta_{l_i}^k q_{l_i}^k \sigma_{l_i}} + \sqrt{\left[\frac{4(C_{l_i}^{\max} - q_{l_i}^k)(\delta_{l_i})^3 + q_{l_i}^k}{\zeta_{l_i}^k q_{l_i}^k (4(\delta_{l_i})^2 - 1)} + \frac{4G_{l_i}^{k'}(\delta_{l_i})^2}{\varphi_{l_i}^k U_{l_i}^k \zeta_{l_i}^k q_{l_i}^k \sigma_{l_i}} \right]^2 + \left[-\frac{16(\delta_{l_i})^4 (C_{l_i}^{\max} - q_{l_i}^k)^2}{(\zeta_{l_i}^k q_{l_i}^k)^2 \sigma_{l_i}} - \frac{1}{(\delta_{l_i})^2} \right]} \quad (27)$$

附录 2 基站效用函数是可微的实函数证明

定理 1 可微的实函数必须是连续函数。

定理 2 闭区间上的连续实函数必须包含最大值和最小值。

式(19)关于 h_k^l 的偏导数为

$$\frac{\partial E_k^l}{\partial h_k^l} = -\frac{\zeta_{l_i}^k q_{l_i}^k A_1}{2\delta_{l_i}} - \frac{\zeta_{l_i}^k q_{l_i}^k 8G_{l_i}^{k'}(\delta_{l_i})^3 A_1}{16\varphi_{l_i}^k U_{l_i}^k (\delta_{l_i})^4 A_2} - \frac{\zeta_{l_i}^k q_{l_i}^k (2q_{l_i}^k \delta_{l_i} \varphi_{l_i}^k U_{l_i}^k + \zeta_{l_i}^k q_{l_i}^k \varphi_{l_i}^k U_{l_i}^k h_k^l) A_1}{16\varphi_{l_i}^k U_{l_i}^k (\delta_{l_i})^4 A_2}$$

$$\frac{\delta_i \zeta_i^k q_i^k h_k^l + q_i^k (\zeta_i^k h_k^l)^2 + 2\delta_i \zeta_i^k h_k^l}{(4\delta_i q_i^k + \zeta_i^k q_i^k h_k^l + 2\delta_i A_2)^2} A_2 + \frac{2\delta_i (2\delta_i \zeta_i^k q_i^k + 2q_i^k (\zeta_i^k)^2)}{(4\delta_i q_i^k + \zeta_i^k q_i^k h_k^l + 2\delta_i A_2)^2} - \frac{2\zeta_i^k q_i^k A_2 + \delta_i A_3 + 2\delta_i (2\delta_i \zeta_i^k (A_2)^2 + \zeta_i^k h_k^l A_3)}{(4\delta_i q_i^k + \zeta_i^k q_i^k h_k^l + 2\delta_i A_2)^2} A_2 < 0 \quad (28)$$

其中, A_1 、 A_2 和 A_3 分别为

$$A_1 = F_k^l + \theta \gamma_i^k \cos t_i^k$$

$$A_2 = \sqrt{\left(\frac{\zeta_i^k q_i^k G_i^k h_k^l}{\delta_i \varphi_i^k U_i^k} - (q_i^k)^2 \right) + \frac{(2\delta_i q_i^k + \zeta_i^k q_i^k h_k^l)^2}{4(\delta_i)^2}}$$

$$A_3 = \frac{\zeta_i^k q_i^k G_i^k}{\delta_i \varphi_i^k U_i^k} + \frac{\zeta_i^k (q_i^k)^2}{4(\delta_i)^3} + \frac{(\zeta_i^k q_i^k)^2 h_k^l}{8(\delta_i)^4}$$

结合定理 1、定理 2、式(26)和式(27)可以证明 E_k^l 是可微的实函数。因此, E_k^l 在闭区间 $[(h_k^l)_{\min}, (h_k^l)_{\max}]$ 内存在最大值。

参考文献:

- [1] CHEN X M, NG D W K, YU W, et al. Massive access for 5G and beyond[J]. *IEEE Journal on Selected Areas in Communications*, 2021, 39(3): 615-637.
- [2] QI Q, CHEN X M, ZHONG C J, et al. Integrated sensing, computation and communication in B5G cellular Internet of things[J]. *IEEE Transactions on Wireless Communications*, 2021, 20(1): 332-344.
- [3] CAO T F, XU C Q, DU J P, et al. Reliable and efficient multimedia service optimization for edge computing-based 5G networks: game theoretic approaches[J]. *IEEE Transactions on Network and Service Management*, 2020, 17(3): 1610-1625.
- [4] 伏玉笋, 杨根科. 无线超可靠低时延通信: 关键设计分析与挑战[J]. *通信学报*, 2020, 41(8): 187-203.
FU Y S, YANG G K. Wireless ultra-reliable and low-latency communication: key design analysis and challenge[J]. *Journal on Communications*, 2020, 41(8): 187-203.
- [5] IMT-2020 (5G) Promotion Group. White paper on 5G concept[S]. 2015.
- [6] 孙国林, 欧睿杰, 刘贵松. 基于深度强化学习的应急物联网切片资源预留算法[J]. *通信学报*, 2020, 41(9): 8-20.
SUN G L, OU R J, LIU G S. Deep reinforcement learning-based resource reservation algorithm for emergency Internet-of-things slice[J]. *Journal on Communications*, 2020, 41(9): 8-20.
- [7] CHIEN H T, LIN Y D, LAI C L, et al. End-to-end slicing with optimized communication and computing resource allocation in multi-tenant 5G systems[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(2): 2079-2091.
- [8] FOSSATI F, MORETTI S, PERNY P, et al. Multi-resource allocation for network slicing[J]. *IEEE/ACM Transactions on Networking*, 2020, 28(3): 1311-1324.
- [9] HAN Y, TAO X F, ZHANG X F, et al. Hierarchical resource allocation in multi-service wireless networks with wireless network virtualization[J]. *IEEE Transactions on Vehicular Technology*, 2020, 69(10): 11811-11827.
- [10] 唐伦, 魏延南, 谭硕, 等. H-CRAN 网络下联合拥塞控制和资源分配的网络切片动态资源调度策略[J]. *电子与信息学报*, 2020, 42(5): 1244-1252.
TANG L, WEI Y N, TAN Q, et al. Joint congestion control and resource allocation dynamic scheduling strategy for network slices in heterogeneous cloud radio access network[J]. *Journal of Electronics & Information Technology*, 2020, 42(5): 1244-1252.
- [11] AJIBARE A T, FALOWO O E. Resource allocation and admission control strategy for 5G networks using slices and users priorities[C]//*Proceedings of 2019 IEEE AFRICON*. Piscataway: IEEE Press, 2020: 1-6.
- [12] 吴大鹏, 郑豪, 崔亚平. 面向服务的车辆网络切片协调智能体设计[J]. *电子与信息学报*, 2020, 42(8): 1910-1917.
WU D P, ZHENG H, CUI Y P. Service-oriented coordination agent design for network slicing in vehicular networks[J]. *Journal of Electronics & Information Technology*, 2020, 42(8): 1910-1917.
- [13] FOSSATI F, HOTEIT S, MORETTI S, et al. Fair resource allocation in systems with complete information sharing[J]. *IEEE/ACM Transactions on Networking*, 2018, 26(6): 2801-2814.
- [14] 曾菊玲, 张春雷, 蒋砺思, 等. 基于信道定价的无线虚拟网络资源分配策略: 匹配/Stackelberg 分层博弈[J]. *电子与信息学报*, 2021, 43(1): 108-115.
ZENG J L, ZHANG C L, JIANG L S, et al. Channel's price-based resource allocation for wireless virtual network: a hierarchical matching/Stackelberg game approach[J]. *Journal of Electronics & Information Technology*, 2021, 43(1): 108-115.
- [15] LIETO A, MALANCHINI I, MANDELLI S, et al. Strategic network slicing management in radio access networks[J]. *IEEE Transactions on Mobile Computing*, 2022, 21(4): 1434-1448.
- [16] TRAN T D, LE L B. Resource allocation for multi-tenant network slicing: a multi-leader multi-follower Stackelberg game approach[J].

- IEEE Transactions on Vehicular Technology, 2020, 69(8): 8886-8899.
- [17] RAVEENDRAN N, GU Y N, JIANG C X, et al. Cyclic three-sided matching game inspired wireless network virtualization[J]. IEEE Transactions on Mobile Computing, 2021, 20(2): 416-428.
- [18] GUIJARRO L, VIDAL J R, PLA V. Competition between service providers with strategic resource allocation: application to network slicing[J]. IEEE Access, 2021, 9: 76503-76517.
- [19] NHU C N, PARK M. Dynamic network slice scaling assisted by attention-based prediction in 5G core network[J]. IEEE Access, 2022, 10: 72955-72972.
- [20] SONG F, LI J, MA C, et al. Dynamic virtual resource allocation for 5G and beyond network slicing[J]. IEEE Open Journal of Vehicular Technology, 2020, 1: 215-226.
- [21] SANTOS E, SOUZA R D, REBELATTO J L, et al. Network slicing for URLLC and eMBB with max-matching diversity channel allocation[J]. IEEE Communications Letters, 2020, 24(3): 658-661.
- [22] XIANG H Y, YAN S, PENG M G. A realization of fog-RAN slicing via deep reinforcement learning[J]. IEEE Transactions on Wireless Communications, 2020, 19(4): 2515-2527.
- [23] LI Y J, ZHAO Y L, LI J, et al. Side channel attack-aware resource allocation for URLLC and eMBB slices in 5G RAN[J]. IEEE Access, 2019, 8: 2090-2099.
- [24] TANG L, ZHANG Y, LIANG R, et al. Virtual resource allocation algorithm for network utility maximization based on network slicing[J]. Journal of Electronics & Information Technology, 2017, 39(8): 1812-1818.
- [25] CABALLERO P, BANCHS A, VECIANA G D, et al. Network slicing games: enabling customization in multi-tenant networks[J]. IEEE/ACM Transactions on Networking, 2019, 27(2): 662-675.

[作者简介]



王再见（1980—），男，安徽定远人，博士，安徽师范大学教授、博士生导师，主要研究方向为面向 5G 的无线多媒体通信、多媒体大数据技术、深度学习、人工智能等。



谷慧敏（1998—），女，安徽马鞍山人，安徽师范大学硕士生，主要研究方向为无线多媒体通信。