

时分波分无源光网络与云无线接入网联合架构中负载均衡的用户关联与资源分配策略

张 鸿 韩 霄 王汝言* 李职杜 周 敏

(重庆邮电大学通信与信息工程学院 重庆 400065)

(重庆高校市级光通信与网络重点实验室 重庆 400065)

(泛在感知与互联重庆市重点实验室 重庆 400065)

摘 要: 在时分波分无源光网络(TWDM-PON)与云无线接入网(C-RAN)的联合架构中, 由于无线域的负载不均衡问题, 限制了网络整体的传输效率。为了充分利用TWDM-PON与C-RAN联合架构的网络资源, 并保证用户的服务质量(QoS), 该文提出一种负载均衡的用户关联与资源分配算法(LBUARA)。首先根据不同用户的服务质量需求以及分布式无线射频头端(RRH)的负载对用户的影响, 构建用户收益函数。进而, 在保证用户服务质量的前提下, 根据网络状态建立随机博弈模型, 并基于多智能体Q学习提出负载均衡的用户关联和资源分配算法, 从而获得最优的用户关联与资源分配方案。仿真结果表明, 所提的用户关联和资源分配策略能够实现网络的负载均衡, 保证用户的服务质量, 并提高网络吞吐量。

关键词: 云无线接入网; 时分波分无源光网络; 负载均衡; 用户关联; 服务质量

中图分类号: TN915; TP393.04

文献标识码: A

文章编号: 1009-5896(2021)09-2672-08

DOI: 10.11999/JEIT200849

Load Balancing User Association and Resource Allocation Strategy in Time and Wavelength Division Multiplexed Passive Optical Network and Cloud Radio Access Network Joint Architecture

ZHANG Hong HAN Xiao WANG Ruyan LI Zhidu ZHOU Min

(School of Telecommunication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

(Chongqing Key Laboratory of Optical Communication and Network, Chongqing 400065, China)

(Chongqing Key Laboratory of Ubiquitous Sensing and Networking, Chongqing 400065, China)

Abstract: The load imbalance in the wireless domain limits the overall transmission efficiency of the network in the joint architecture of Time and Wavelength Division Multiplexed Passive Optical Network (TWDM-PON) and Cloud Radio Access Network (C-RAN). A Load Balancing User Association and Resource Allocation (LBUARA) algorithm is proposed to ensure the Quality of Service(QoS) of users, and make full use of network resources TWDM-PON jointly with C-RAN architecture. Firstly, the user revenue function is constructed according to the service quality requirements of different users and the impact of Remote Radio Head (RRH) load on users. Furthermore, a random game model is established according to the network state, under the premise of ensuring the quality of user service. A user association and resource allocation algorithm based on multi-agent Q-learning load balancing is proposed to obtain the optimal user association and resource allocation plan. The simulation results show that users association and resource allocation strategies mentioned can

收稿日期: 2020-09-30; 改回日期: 2021-01-29; 网络出版: 2021-03-02

*通信作者: 王汝言 wangry@cqupt.edu.cn

基金项目: 国家自然科学基金(61771082, 61871062), 重庆市高校创新研究群体(CXQT20017), 重庆市教委科学技术研究项目(KJQN201900609, KJQN202000626), 重庆市自然科学基金重点项目(cstc2020jcyj-zdxmX0024)

Foundation Items: The National Natural Science Foundation of China (61771082, 61871062), The University Innovation Research Group of Chongqing (CXQT20017), The Science and Technology Research Program of Chongqing Municipal Education Commission (KJQN201900609, KJQN202000626), The Natural Science Foundation of Chongqing (cstc2020jcyj-zdxmX0024)

achieve load balancing network to ensure quality of service users, and improve network throughput.

Key words: Cloud Radio Access Network (C-RAN); Time and Wavelength Division Multiplexed Passive Optical Network (TWDM-PON); Load balancing; User association; Quality of Service (QoS)

1 引言

越来越多的低延时密集型任务(如实时通信、高清视频播放和游戏等)出现在日常生活中, 日益增长的带宽需求给网络基础设施建设带来了严峻挑战^[1,2]。与传统分布式无线接入网络(Distributed Radio Access Network, D-RAN)不同, 云无线接入网(Cloud Radio Access Network, C-RAN)将传统基站配置分解成集中式的基带处理单元(Base-Band Unite, BBU)和分布式无线射频头端(Remote Radio Head, RRH)两个部分, 且将多个BBU集中到BBU池中, 以实现统计复用增益^[3,4]。专用光纤作为C-RAN的一种前传方案会带来巨大的部署成本^[5]。文献^[6]对C-RAN不同的前传方式进行比较, 作者认为时分波分无源光网络(Time and Wavelength Division Multiplexed Passive Optical Network, TWDM-PON)具有低成本、低延时、低能耗等优点, 是C-RAN最有效的前传解决方案之一。使用TWDM-PON连接BBU池和RRH, 不但可以降低系统升级和维护的资本与运营支出, 还可以提高频谱利用率^[7]。

虽然TWDM-PON与C-RAN的联合架构具有丰富的带宽资源, 然而无线域的传输容量有限, 整体网络性能受到无线域传输能力的限制^[8]。文献^[9]提出了一种新颖的具有切片特征感知的自然启发式无线资源分配方案。文献^[10]进一步优化了异构云无线接入网中用户关联和功率分配问题。文献^[11]针对资源竞争问题, 提出了基于Stackelberg博弈的虚拟化无线传感网络资源分配策略。文献^[12]提出一种前传感知的用户关联方案, 将RRH所关联用户的信噪比作为决策变量。文献^[13]根据用户供应指标和数据需求指标建立任务价格模型, 转换成用户和移动边缘计算(Mobile Edge Computing, MEC)服务器之间关于任务奖励的博弈。文献^[14]介绍了一种基于顺序拍卖的用户关联算法(Sequential Auction-Based User Association algorithm, SA-BUA), 目的是在负载均衡和用户服务质量(Quality of Service, QoS)之间进行折中。文献^[15]强调了负载均衡在最大化网络效用方面具有显著优势。通过对无线网络业务量的研究表明, 随着RRH负载的增大, 与其相关联的用户服务质量会随之下降^[16]。通过设计用户关联方案可以优化网络结构, 提高无线域频谱资源利用率, 进而提高网络

吞吐量^[17]。传统的用户关联方法根据信号的强弱关联RRH, 极大地影响了无线传输的效率^[18]。因此用户关联方案在TWDM-PON与C-RAN联合架构中具有重要意义。

针对以上问题, 首先根据不同用户的QoS需求和RRH的负载对用户的影响, 构建与RRH负载相关的用户收益函数, 并根据网络状态建立随机博弈模型, 基于多智能体Q学习提出负载均衡的用户关联和资源分配算法(Load Balancing User Association and Resource Allocation algorithm, LBUARA), 得到最优的用户关联策略, 从而在保证用户QoS的同时, 充分利用网络资源, 提高网络的吞吐量。

2 网络架构

TWDM-PON与C-RAN的联合架构如图1所示。BBU池和光线路终端(Optical Line Terminal, OLT)位于中心局(Central Office, CO), 由集中控制器统一调度。在BBU池中, BBU之间通过直通交换机连接, 实现数据迁移和资源共享。OLT内部有多个线卡(Line Card, LC), 可提供多个波长资源, 部分光网络单元(Optical Network Unit, ONU)可以共享同一波长。将RRH的频谱资源划分为资源块(Resource Block, RB), 每个RB的带宽是180 kHz。RRH需要为用户分配足够的RB来保证用户的QoS。由于光域与无线域划分的业务种类不同, 在ONU处设置一个两级缓存队列, 无线域的数据包到达缓存队列, 通过流量整形将无线域业务映射为光域业务, 同时保证了无线域与光域不同优先级业务的QoS性能^[19]。

在物理网络中, 用户首先向多个RRH发送访问请求, RRH将接收到信号转发给与其直连的ONU。之后ONU使用通用公共无线接口(Common Public Radio Interface, CPRI)将用户信息发送到OLT和BBU, 集中控制器对这些信息进行分析计算, 从而获得用户关联策略, 为ONU分配波长和带宽资源, 并将配置信息传送到底层物理网络, 实现用户关联和资源分配。

3 系统模型

在TWDM-PON和C-RAN的联合架构中, 用户可以向多个RRH发送访问请求, 中央控制器通过收集用户准确的信道信息以及QoS需求, 优化用

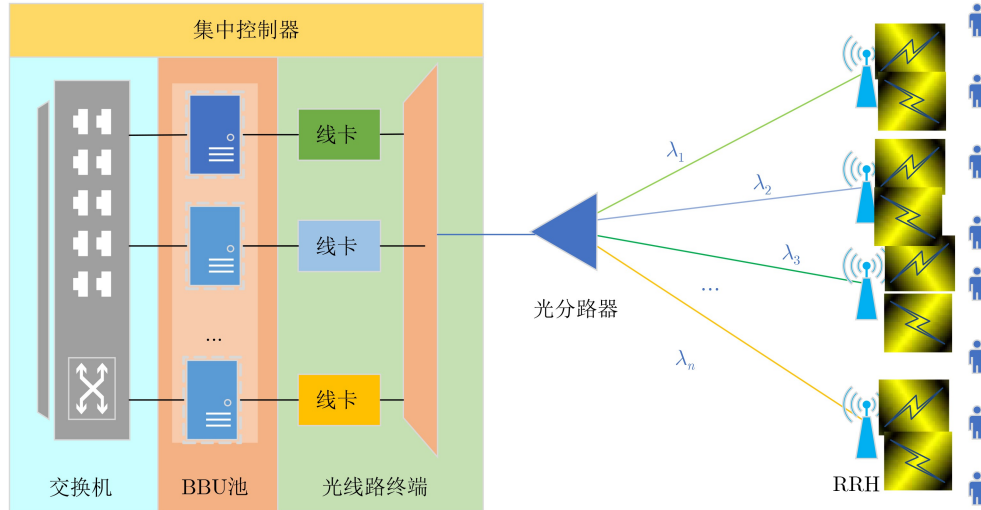


图1 TWDM-PON与C-RAN联合架构

户关联策略，并进行资源分配。令 $\mathcal{R} = \{1, 2, \dots, m\}$ 表示所有RRH的集合， $\mathcal{U} = \{1, 2, \dots, n\}$ 表示所有用户的集合。使用二进制变量 $\alpha_{i,j}(t)$ 表示在时刻 t 用户 i 与RRH j 之间的关联状态。当用户 i 在时刻 t 与RRH j 关联时 $\alpha_{i,j}(t) = 1$ ，否则 $\alpha_{i,j}(t) = 0$ 。每个用户在时刻 t 最多只能关联1个RRH，即

$$\sum_{j \in \mathcal{R}} \alpha_{i,j}(t) \leq 1 \quad (1)$$

若RRH j 拥有 k 个RB，即 $\varphi_j = \{\varphi_{j,1}, \varphi_{j,2}, \dots, \varphi_{j,k}\}$ ，总发射功率是 P_j ，假设将RRH j 的发射功率均匀地分配给RB，则每个RB的功率为 $p_j = P_j/|\varphi_j|$ ，其中 $|\varphi_j|$ 表示RRH j 拥有的RB数量。1个RB在同一时刻只能为1位用户提供服务

$$\sum_{i \in \mathcal{U}} \alpha_{i,j}(t) \varphi_{j,k} \leq 1 \quad (2)$$

且RRH为每个关联用户分配的RB总量不超过其拥有的RB数量

$$\sum_{i \in \mathcal{U}} \sum_{\varphi_{j,k} \in \varphi_j} \alpha_{i,j}(t) \varphi_{j,k} \leq |\varphi_{j,k}| \quad (3)$$

假设用于下行链路传输的信道是平坦衰落的，RRH j 在时刻 t 为用户 i 分配的RB数量为 $n_{i,j}(t)$ ，那么RRH j 为用户 i 分配的功率 $p_{i,j}(t) = n_{i,j}(t)p_j$ ，则在时刻 t 用户 i 与RRH j 之间的信干噪比(Signal-to-Interference and Noise Ratio, SINR)可以表示为^[14]

$$\text{SINR}_{i,j}(t) = \frac{\alpha_{i,j}(t)p_{i,j}(t)h_{i,j}(t)d_{i,j}^{-\alpha}(t)}{\sum_{k \in \mathcal{R} \setminus \{j\}} \alpha_{i,k}(t)p_{i,k}(t)h_{i,k}(t)d_{i,k}^{-\alpha}(t) + \sigma^2} \quad (4)$$

其中， $h_{i,j}(t)$ 表示用户 i 和RRH j 之间的信道增益系数， $d_{i,j}^{-\alpha}(t)$ 是用户 i 和RRH j 在时刻 t 的欧几里得距

离，路径损耗因子 $\alpha > 2$ ^[14]。 σ^2 表示白噪声的功率谱密度。事实上用户距离RRH越远，信道条件就越差，RRH需要提供更多的RB来保证SINR。与RRH j 关联的用户 i 在时刻 t 的速率可以表示为

$$r_{i,j}(t) = n_{i,j}(t) b \log_2(1 + \text{SINR}_{i,j}(t)) \quad (5)$$

其中， b 表示每个RB的带宽。可见用户的速率与可使用的RB数量有关，每个RRH如果想要尽可能地关联更多的用户，就需要为每个用户分配尽可能少的RB，提高RB的利用率。

每位用户的服务请求不同，对QoS的需求也各不相同，RRH通过为不同用户分配一定数量的RB保证用户的QoS。为了满足不同用户的QoS需求，令 $r_i^{\text{req}}(t)$ 表示用户 i 在时刻 t 的带宽请求

$$n_{i,j}(t) b \log_2(1 + \text{SINR}_{i,j}(t)) \geq r_i^{\text{req}}(t) \quad (6)$$

因此根据式(5)，可以得到RRH j 在时刻 t 为用户 i 提供服务的RB数量为

$$n_{i,j}(t) = \left\lceil \frac{r_i^{\text{req}}(t)}{b \log_2(1 + \text{SINR}_{i,j}(t))} \right\rceil \quad (7)$$

定义用户 i 在时刻 t 的收益为 $v_i(t)$

$$v_i(t) = \ln \left(1 + \frac{|\varphi_j| - \sum_{k \in \mathcal{U}} n_{k,j}(t)}{n_{i,j}(t)} \right) \quad (8)$$

用户在时刻 t 的收益取决于所关联的RRH负载以及为其分配的RB数量，若RRH不能保证用户 i 的QoS，则RRH不为用户 i 提供RB，此时用户 i 的收益 $v_i(t) = 0$ 。根据对数函数的性质，收益函数 $v_i(t)$ 鼓励用户选择距离较远，但负载较轻的RRH。当整个网络的负载过大时，用户会优先选择分配RB资源较少的RRH，从而提高RB的利用率。用户在

任意时刻的收益取决于当前的网络状态和其他用户的关联状态，且用户收益值会随着网络整体负载的上升而下降。为了在实现网络负载均衡的同时，保证每个用户的QoS，因此将目标函数定义为收益最大化问题

$$\max_{i \in \mathcal{U}, j \in \mathcal{R}} \alpha_{i,j}(t)v_i(t) \quad (9)$$

由于用户关联与资源分配联合优化问题具有非凸和组合特性^[20]。因此获得的全局最优策略十分具有挑战性。多智能体强化学习 (Multi-Agent Reinforcement Learning, MARL)通过与环境进行交互获取奖励值来学习改善自己的策略，从而获得该环境下最优策略。考虑到在TWDM-PON与C-RAN的联合架构中，集中控制器可获得每个用户的信息，因此，本文通过建立随机博弈模型，并使用强化学习的方法，基于多智能体Q学习提出负载均衡的用户关联和资源分配算法以获得联合优化问题的最优解。

4 用户关联与资源分配的联合优化

4.1 随机博弈模型

在任意时刻 t ，由于用户的收益估值受当前网络状态和用户动作影响，引入有限状态的马尔可夫决策过程(Markov Decision Process, MDP)，将联合优化问题转化为随机博弈模型MDP($S, A_i, P_{ss'}, V_i$)^[21]。 S 是可能的状态集合， A_i 表示第 i 个用户的动作集合， $P_{ss'}$ 表示用户从状态 s 转换为状态 s' 的概率， V_i 是第 i 个用户的收益。定义每个用户在时刻 t 的状态为 $s(t)$

$$s(t) = \{s_1(t), s_2(t), \dots, s_n(t)\} \quad (10)$$

其中， $s_i(t) \in \{0, 1\}$ ，当 $s_i(t) = 1$ 时意味着用户 i 的QoS需求得到满足，当 $s_i(t) = 0$ 时说明没有满足用户 i 的QoS。所有的用户在时刻 t 都需要选择关联RRH，则定义每个用户的动作集合 A_i 为

$$A_i = \{a_{i,1}(t), a_{i,2}(t), \dots, a_{i,m}(t)\} \quad (11)$$

每个用户可能做出的动作数量与当前可关联的RRH数量相同，并且用户当前动作会影响下一时刻的网络状态。在任意时刻 t ，定义其他 $n-1$ 个用户的动作为 $a_{-i}(t) = \{a_1(t), \dots, a_{i-1}(t), a_{i+1}(t), \dots, a_n(t)\}$ 。当用户 i 采取动作 $a_i(t)$ 时，用户 i 在时刻 t 的收益为 $v_i(t)$ ，若其他用户采取动作 $a_{-i}(t)$ ，则用户 i 在 t 时刻的收益可以表示为 $v_i(t) = v_i(s, a_i^*, a_{-i}^*)$ 。若 (a_i, a_{-i}) 为该博弈可能的解决方案，则当每个用户的状态满足不等式(12)时，博弈将达到纳什均衡状态，此时每个用户的动作对于其他用户动作来说都是最佳响应

$$v_i(s, a_i^*, a_{-i}^*) \geq v_i(s, a_i, a_{-i}^*), \forall a_i \in A_i \quad (12)$$

在未知的随机环境中，通过使用强化学习的方法，可以获得最优的策略 $\pi_i^*(s_i, a_i)$ ，即用户 i 在状态 s 时选择动作 a_i 的收益最大。在联合架构中，通过集中控制器可以获得全局信息，所有用户基于状态空间学习最优策略 $a_i^* = \pi_i^*(s_i, a_i)$ 。用户 i 通过在每个状态下获得最优的策略 $\pi_i^*(s_i, a_i)$ ，最大化其在 t 时刻的状态价值函数

$$v_i(s, \pi_i, \pi_{-i}) = \mathbb{E} \left[\sum_{t=0}^{T-1} \gamma^t v_i(s, \pi_i(t), \pi_{-i}(t)) | s(0) = s \right] \quad (13)$$

状态价值函数 $v_i(s, \pi_i, \pi_{-i})$ 是一个期望值，其中 $\gamma \in [0, 1]$ 是收益衰减因子， $\pi_{-i} = (\pi_1, \dots, \pi_{i-1}, \pi_{i+1}, \dots, \pi_n)$ 表示其他用户的策略向量。根据马尔可夫的性质，令 $u_i(s, \pi_i, \pi_{-i}) = \mathbb{E}[v_i(s, \pi_i, \pi_{-i})]$ ，将状态价值函数表述为

$$v_i(s, \pi_i, \pi_{-i}) = u_i(s, \pi_i, \pi_{-i}) + \gamma \sum_{s' \in S} p_{ss'}(\pi_i, \pi_{-i}) v_i(s', \pi_i, \pi_{-i}) \quad (14)$$

其中， $p_{ss'}(\pi_i, \pi_{-i})$ 是用户从状态 $s \rightarrow s'$ 的状态转移概率。每个用户的预期积累的收益由 $v_i(s, \pi_i, \pi_{-i})$ 确定，并且每个状态可以从每个可能的策略获得即时奖励。若用户策略 (π_i^*, π_{-i}^*) 中任意一个 π_i 均满足如式(15)的条件，则达到纳什均衡状态，其中 $\pi_{-i}^* = (\pi_1^*, \dots, \pi_{i-1}^*, \pi_{i+1}^*, \dots, \pi_n^*)$

$$v_i(s, \pi_i^*, \pi_{-i}^*) \geq v_i(s, \pi_i, \pi_{-i}^*), \forall s_i \in S \quad (15)$$

由于任何有限博弈都有混合策略均衡，因此总存在满足贝尔曼最优方程(Bellman Optimality Equation, BOE)的纳什均衡策略 $v_i^*(s, \pi_i, \pi_{-i}) = v_i(s, \pi_i^*, \pi_{-i}^*)$ ^[22]

$$v_i(s, \pi_i^*, \pi_{-i}^*) = \max_{a_i \in A_i} \left(u_i(s, \pi_i, \pi_{-i}^*) + \gamma \sum_{s' \in S} p_{ss'}(a_i, \pi_{-i}^*) v_i(s', \pi_i^*, \pi_{-i}^*) \right) \quad (16)$$

4.2 基于多智能体Q学习的用户关联和资源分配

多智能体Q学习是解决MDP问题的RL方法之一，通过将一个随机博弈看作一个多智能体RL的过程，与环境不断交互学习每个状态的收益函数，进而得到最优的纳什策略。根据式(14)可以得到最佳Q值函数

$$Q_i^*(s, a_i) = u_i(s, \pi_i, \pi_{-i}^*) + \gamma \sum_{s' \in S} p_{ss'}(a_i, \pi_{-i}^*) v_i(s', \pi_i^*, \pi_{-i}^*) \quad (17)$$

虽然对于用户来说可能存在多个纳什均衡策略 $\pi_i^*(s_i, a_i)$ ，但是根据式(16)可得，最优的策略 $\pi_i^*(s_i, a_i)$ 可以通过 $Q_i^*(s, a_i)$ 得到

$$v_i(s, \pi_i^*, \pi_{-i}^*) = \max_{a_i \in A_i} Q_i^*(s, a_i) \quad (18)$$

将式(17)与式(18)整合，可以得到

$$Q_i^*(s, a_i) = u_i(s, \pi_i, \pi_{-i}^*) + \gamma \sum_{s' \in S} p_{ss'}(a_i, \pi_{-i}^*) \max_{a'_i \in A_i} Q_i^*(s', a'_i) \quad (19)$$

通过Q学习的方法，基于可用信息 $(s, a_i, a'_i, s', u(s, \pi_i, \pi_{-i}))$ 以递归的方式找到最佳策略，Q值函数的更新方程为

$$Q_i(s, a_i) = Q_i(s, a_i) + \delta \left(u_i(s, a_i, \pi_{-i}) + \gamma \max_{a'_i \in A_i} Q_i^*(s', a'_i) - Q_i(s, a_i) \right) \quad (20)$$

其中， δ 表示 $Q_i(s, a_i)$ 的更新学习速率， δ 增加学习的收敛速率会更快，然而较高的学习速率可能会导致局部最优而不是全局最优。考虑到实际的算法执行， δ 通常设置为0.1^[23,24]。Q学习方法更新 $Q_i(s, a_i)$ 时趋于收敛。在学习过程中，平衡探索(exploration)与利用(exploitation)的关系是用户关联过程中的重要问题，探寻的目的是为了不断尝试关联新的RRH，而利用的目的是为了使用已经探寻过的RRH。通常使用概率选择作为执行动作的策略，这是一种平衡探索与利用的有效方法。但是在进行探索时， ε -贪婪策略会在所有可用动作中进行平均选择，这意味着可能选择最差的动作。因此使用玻尔曼兹分布，将Q值的分级函数 $\phi_i(s_i, a_i)$ 作为动作的概率^[25]

$$\phi_i(s_i, a_i) = \frac{e^{Q_i(s_i, a_i)/\tau}}{\sum_{a_k \in A_i} e^{Q_i(s_i, a_k)/\tau}} \quad (21)$$

若各个动作的平均收益差异较小，则选择各动作的概率差异也较小，若某些动作的平均收益明显高于其他动作，则这些动作被选择的概率也会更高。即最高的收益被选择的概率也最大，其他的动作则根据Q值进行排名和加权。其中 τ 是一个正参数， τ 越小，则选择平均收益较高的动作的概率越高， τ 越大，则各个动作被选择的概率趋于相同。当 τ 趋于0时，算法将趋于利用，反之则趋于探索。

针对TWDM-PON与C-RAN联合架构中的优化问题，基于多智能体Q学习提出负载均衡的用户关联和资源分配算法。用户从一个状态开始，根据某个策略执行一系列动作直到结束，环境给出终止

状态的即时收益为止为一个完整的训练周期(episode)。当用户过多时，当前网络的可用资源总量可能不能满足所有用户的QoS需求，因此设置 t 为每个episode的最大执行次数。负载均衡的用户关联和资源分配算法如表1所示。

表1 负载均衡的用户关联和资源分配算法

-
- (1) 初始化episode，每个用户的Q值 $Q_i(s, a_i)$ 以及 $\phi_i(s_i, a_i)$
 - (2) **for** each step of an episode to t steps **do**
 - (3) **for** each UE i **do**
 - (4) 在状态 s_i 时通过式(21)选择动作 a_i
 - (5) 通过式(7)计算为每个用户分配的RB数量
 - (6) 通过式(13)计算 V_i
 - (7) 每个用户获取关联状态 s' ，设置 $s' \rightarrow s$
 - (8) 通过式(20)更新 $Q_i(s, a_i)$
 - (9) 更新 $\phi_i(s_i, a_i)$
 - (10) **end for**
 - (11) **if** 当前状态集合 $S = \{1, 1, \dots, 1\}$
 - (12) **break**
 - (13) **end if**
 - (14) 最终所有的用户得到关联策略 (s_i, a_i)
-

在每个episode开始时初始化网络状态，之后每个用户根据式(21)选择执行动作 a_i ，计算用户执行动作后的收益以及下一个状态 s' ，并更新 $Q_i(s, a_i)$ 和 $\phi_i(s_i, a_i)$ 。当满足所有用户的QoS需求或循环次数达到 t 时，开始下一个episode。

5 性能测试与分析

本文采用Pycharm 2019平台，并使用Python 3.6进行仿真。考虑在覆盖面积为10000 m²小规模的网络场景中评估算法的性能，将本文提出的LBUARA与文献[14]中的SA-BUA算法和文献[18]中提及的最大信干噪比算法(Max-Signal to Interference plus Noise Ratio algorithm, Max-SINR)进行比较。设置7个RRH均匀分布在覆盖区域，用户在区域中的位置随机生成，且每个用户所需的传输速率服从在[1 Mbit/s, 5 Mbit/s]之间的随机分布。对于无线资源来说，每个RRH的发射功率为10 W，拥有100个RB，每个RB的带宽为180 kHz。设置学习速率 $\delta=0.1$ ， $\gamma=0.9$ ，路径损耗因子 $\alpha=4$ ，信道功率增益服从在[5,14]之间的随机分布^[10]。高斯白噪声的功率谱密度为 10^{-10} W/Hz。

图2所示为在网络中有50个用户时，不同算法下的RRH与用户的关联情况。如图2(a)和图2(c)所示，使用本文LBUARA算法和SABUA算法时，RRH关联的用户数量更为均匀。从图2(b)可以看出，使

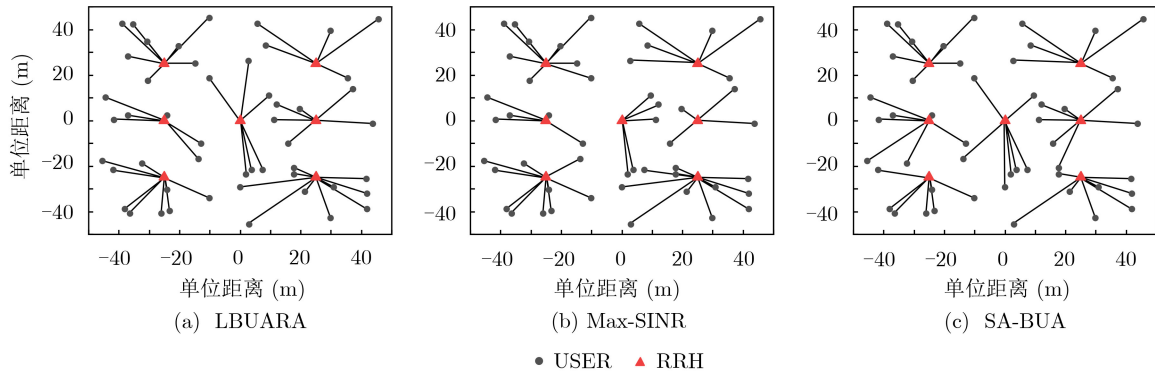


图 2 不同算法的用户关联状态

用Max-SINR算法时，不同RRH关联的用户数量有较大差距。这是由于LBUARA算法和SA-BUA算法会根据RRH的负载状态，使用户关联距离较远但负载较低的RRH，且Max-SINR算法会导致用户优先关联附近SINR较高的RRH。

本文使用Jain公平性指数 η 衡量整个网络的负载均衡状态^[26]。 ω_j 表示RRH j 所消耗的RB资源， n 为当前网络中RRH的数量

$$\eta = \frac{\left(\sum_{j=1}^n \omega_j / |\varphi_j|\right)^2}{n \cdot \sum_{j=1}^n (\omega_j / |\varphi_j|)^2} \quad (22)$$

如图3所示，当网络中用户数量较少时，使用LBUARA算法与SA-BUA算法的网络负载均衡状态远优于Max-SINR算法。由于SA-BUA算法为每个用户分配的RB数量固定，并根据RRH关联的用户数量衡量负载状态，因此SA-BUA算法更容易实现RRH的负载均衡。随着用户数量的增多，网络整体的负载增大，区域内所有RRH负载会随之增加。当用户数量逐渐到当前网络的承载极限时，每个RRH的负载状态趋于一致。

不同用户数量的网络在不同时刻的吞吐量如图4所示。当用户数量较少时，网络有充足的资源满足所有用户的需求，且SA-BUA算法为用户分配的RB数量可能高于用户需求。不同时刻用户的位置和需求会发生变化，若当前每个RRH附近的用户数量分布较为均匀，则缓解了Max-SINR算法导致的负载不均衡问题，从而提高了网络中RB资源的利用率，满足了更多用户的需求。同样，由于SA-BUA算法忽略了不同用户QoS需求，可能造成RB资源的过度配置，而均匀的用户分布使SA-BUA算法下的RB资源利用率更高。与其他两种算法相比，LBUARA算法受用户分布的影响较小，且在资源

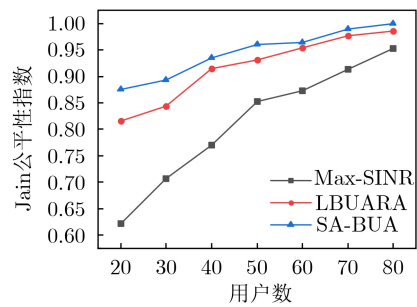


图 3 Jain公平性指数与用户数量的关系

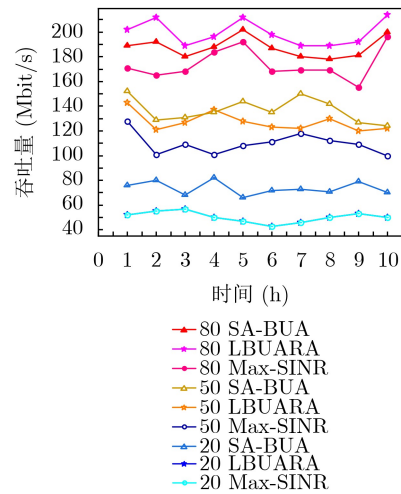


图 4 不同时刻的网络吞吐量

有限的情况下，随着用户数量的增加，LBUARA算法能够保障更多用户的QoS。

满足QoS约束的用户数量如图5所示。Max-SINR算法使附近用户密度较大的RRH关联过多的用户，网络中的RB资源不能充分利用，导致部分用户的QoS得不到保证。SA-BUA算法为每个用户提供相同数量的RB，导致部分用户的实际速率大于其请求速率，而部分用户的需求得不到满足。LBUARA算法根据用户需求分配RB资源，并根据RRH负载进行用户关联，使网络资源得到更充分的利用。随着用户数量的增加，各个RRH附近的

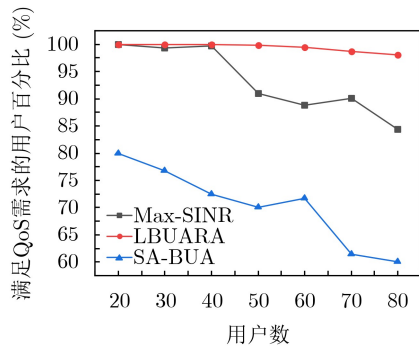


图5 满足QoS需求的用户数量

用户分布相对之前更为均匀,因此在Max-SINR算法和SA-BUA算法中满足QoS的用户数量会在持续下降后出现上升。随着用户的继续增加,网络内的RB资源不能满足所有用户的需求,满足QoS的用户数量整体呈现下降趋势。

6 结论

为了解决在TWDM-PON与C-RAN的联合架构中无线域的负载不均衡问题,本文提出一种负载均衡的用户关联与资源分配算法。首先,给出TWDM-PON与C-RAN的联合架构的优化框架,然后根据网络中RRH的负载情况,构建用户收益函数,并建立随机博弈模型,最后基于多智能体Q学习提出负载均衡的用户关联与资源分配算法解决用户关联和资源分配问题。通过仿真验证,所提算法能够有效解决网络的负载均衡问题,保证用户QoS,提高网络吞吐量。

参考文献

- [1] REN Hong, LIU Nan, PAN Cunhua, *et al.* Low-latency C-RAN: An next-generation wireless approach[J]. *IEEE Vehicular Technology Magazine*, 2018, 13(2): 48–56. doi: [10.1109/MVT.2018.2811244](https://doi.org/10.1109/MVT.2018.2811244).
- [2] ZHAO Wentao and WANG Shaowei. Traffic density-based RRH selection for power saving in C-RAN[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(12): 3157–3167. doi: [10.1109/JSAC.2016.2600414](https://doi.org/10.1109/JSAC.2016.2600414).
- [3] KALIL M, Al-DWEIK A, SHARKH M A, *et al.* A framework for joint wireless network virtualization and cloud radio access networks for next generation wireless networks[J]. *IEEE Access*, 2017, 5: 20814–20827. doi: [10.1109/ACCESS.2017.2746666](https://doi.org/10.1109/ACCESS.2017.2746666).
- [4] SIMEONE O, MAEDER A, PENG Mugen, *et al.* Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems[J]. *Journal of Communications and Networks*, 2016, 18(2): 135–149. doi: [10.1109/JCN.2016.000023](https://doi.org/10.1109/JCN.2016.000023).
- [5] ZHANG Zhiguo, CHEN Yanxu, CAI Shanyong, *et al.* Colorless-light and tunable-light-source schemes for TWDM and WDM PONs[J]. *IEEE Communications Magazine*, 2018, 56(8): 120–128. doi: [10.1109/MCOM.2018.1700751](https://doi.org/10.1109/MCOM.2018.1700751).
- [6] WANG Xinbo, CAVDAR C, WANG Lin, *et al.* Virtualized cloud radio access network for 5G transport[J]. *IEEE Communications Magazine*, 2017, 55(9): 202–209. doi: [10.1109/MCOM.2017.1600866](https://doi.org/10.1109/MCOM.2017.1600866).
- [7] 王汝言, 徐宁宁, 吴大鹏. 能耗和时延感知的虚拟化云无线接入网络资源分配机制[J]. *电子与信息学报*, 2019, 41(1): 83–90. doi: [10.11999/JEIT180063](https://doi.org/10.11999/JEIT180063).
WANG Ruyan, XU Ningning, and WU Dapeng. Energy consumption and delay-aware resource allocation mechanism for virtualization cloud radio access network[J]. *Journal of Electronics & Information Technology*, 2019, 41(1): 83–90. doi: [10.11999/JEIT180063](https://doi.org/10.11999/JEIT180063).
- [8] ALNOMAN A, CARVALHO G H S, ANPALAGAN A, *et al.* Energy efficiency on fully cloudified mobile networks: Survey, challenges, and open issues[J]. *IEEE Communications Surveys & Tutorials*, 2018, 20(2): 1271–1291. doi: [10.1109/COMST.2017.2780238](https://doi.org/10.1109/COMST.2017.2780238).
- [9] WU Dapeng, ZHANG Zhihao, WU Shaoen, *et al.* Biologically inspired resource allocation for network slices in 5G-enabled internet of things[J]. *IEEE Internet of Things Journal*, 2019, 6(6): 9266–9279. doi: [10.1109/JIOT.2018.2888543](https://doi.org/10.1109/JIOT.2018.2888543).
- [10] ALI M, RABBANI Q, NAEEM M, *et al.* Joint user association, power allocation, and throughput maximization in 5G H-CRAN networks[J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(10): 9254–9262. doi: [10.1109/TVT.2017.2715229](https://doi.org/10.1109/TVT.2017.2715229).
- [11] 王汝言, 李宏娟, 吴大鹏. 基于Stackelberg博弈的虚拟化无线传感网络资源分配策略[J]. *电子与信息学报*, 2019, 41(2): 377–384. doi: [10.11999/JEIT180277](https://doi.org/10.11999/JEIT180277).
WANG Ruyan, LI Hongjuan, and WU Dapeng. Stackelberg game-based resource allocation strategy in virtualized wireless sensor network[J]. *Journal of Electronics & Information Technology*, 2019, 41(2): 377–384. doi: [10.11999/JEIT180277](https://doi.org/10.11999/JEIT180277).
- [12] ELHATTAB M K, ELMESALAWY M M, and ISMAIL T. Fronthaul-aware user association in 5G heterogeneous cloud radio access networks: A matching game perspective[C]. 2018 International Symposium on Networks, Computers and Communications (ISNCC), Rome, Italy, 2018. doi: [10.1109/ISNCC.2018.8531056](https://doi.org/10.1109/ISNCC.2018.8531056).
- [13] WU Dapeng, LIU Jia, and YANG Zhigang. Bilateral satisfaction aware participant selection with MEC for mobile crowd sensing[J]. *IEEE Access*, 2020, 8: 48110–48122. doi: [10.1109/ACCESS.2020.2978774](https://doi.org/10.1109/ACCESS.2020.2978774).
- [14] SU Gongchao, CHEN Bin, LIN Xiaohui, *et al.* A sequential

- auction game for QoS-aware user association in heterogeneous cellular networks[C]. Proceedings of 2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN), Prague, Czech Republic, 2018: 464–468. doi: [10.1109/ICUFN.2018.8436793](https://doi.org/10.1109/ICUFN.2018.8436793).
- [15] GE Xin, LI Xiuhua, JIN Hu, *et al.* Joint user association and user scheduling for load balancing in heterogeneous networks[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(5): 3211–3225. doi: [10.1109/TWC.2018.2808488](https://doi.org/10.1109/TWC.2018.2808488).
- [16] PROSKOCHYLO A, ZRIAKHOV M, and AKULYNICHEV A. The effects of queueing algorithms on QoS for real-time traffic in process of load balancing[C]. 2018 International Scientific-Practical Conference Problems of Infocommunications. Science and Technology (PIC S&T), Kharkiv, Ukraine, 2018: 575–580. doi: [10.1109/INFOCOMMST.2018.8632161](https://doi.org/10.1109/INFOCOMMST.2018.8632161).
- [17] WANG Feng, CHEN Wen, TANG Hongying, *et al.* Joint optimization of user association, subchannel allocation, and power allocation in multi-cell multi-association OFDMA heterogeneous networks[J]. *IEEE Transactions on Communications*, 2017, 65(6): 2672–2684. doi: [10.1109/TCOMM.2017.2678986](https://doi.org/10.1109/TCOMM.2017.2678986).
- [18] AWAIS M, AHMED A, NAEEM M, *et al.* Efficient joint user association and resource allocation for cloud radio access networks[J]. *IEEE Access*, 2017, 5: 1439–1448. doi: [10.1109/ACCESS.2017.2663758](https://doi.org/10.1109/ACCESS.2017.2663758).
- [19] 王汝言, 池文祥, 张鸿. WOBAN中基于令牌桶算法的自适应流量整形策略[J]. 电子与信息学报, 2017, 39(6): 1401–1408. doi: [10.11999/JEIT160974](https://doi.org/10.11999/JEIT160974).
- WANG Ruyan, CHI Wenxiang, and ZHANG Hong. Adaptive traffic shaping policy based on token bucket algorithm of wireless-optical broadband access network[J]. *Journal of Electronics & Information Technology*, 2017, 39(6): 1401–1408. doi: [10.11999/JEIT160974](https://doi.org/10.11999/JEIT160974).
- [20] WANG Kaidi, LIU Yuanwei, DING Zhiguo, *et al.* User association and power allocation for multi-cell non-orthogonal multiple access networks[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(11): 5284–5298. doi: [10.1109/TWC.2019.2935433](https://doi.org/10.1109/TWC.2019.2935433).
- [21] SCHERER W T, ADAMS S, and BELING P A. On the practical art of state definitions for Markov decision process construction[J]. *IEEE Access*, 2018, 6: 21115–21128. doi: [10.1109/ACCESS.2018.2819940](https://doi.org/10.1109/ACCESS.2018.2819940).
- [22] JANG B, KIM M, HARERIMANA G, *et al.* Q-learning algorithms: A comprehensive classification and applications[J]. *IEEE Access*, 2019, 7: 133653–133667. doi: [10.1109/ACCESS.2019.2941229](https://doi.org/10.1109/ACCESS.2019.2941229).
- [23] LI Dong, ZHANG Haijun, LONG Keping, *et al.* User association and power allocation based on q-learning in ultra dense heterogeneous networks[C]. 2019 IEEE Global Communications Conference (GLOBECOM), Waikoloa, USA, 2019: 1–5. doi: [10.1109/GLOBECOM38437.2019.9013455](https://doi.org/10.1109/GLOBECOM38437.2019.9013455).
- [24] ZHAO Nan, LIANG Yingchang, NIYATO D, *et al.* Deep reinforcement learning for user association and resource allocation in heterogeneous cellular networks[J]. *IEEE Transactions on Wireless Communications*, 2019, 18(11): 5141–5152. doi: [10.1109/TWC.2019.2933417](https://doi.org/10.1109/TWC.2019.2933417).
- [25] CHEN Xianfu, ZHAO Zhifeng, and ZHANG Honggang. Stochastic power adaptation with multiagent reinforcement learning for cognitive wireless mesh networks[J]. *IEEE Transactions on Mobile Computing*, 2013, 12(11): 2155–2166. doi: [10.1109/TMC.2012.178](https://doi.org/10.1109/TMC.2012.178).
- [26] SEDIQ A B, GOHARY R H, SCHOENEN R, *et al.* Optimal tradeoff between sum-rate efficiency and Jain's fairness index in resource allocation[J]. *IEEE Transactions on Wireless Communications*, 2013, 12(7): 3496–3509. doi: [10.1109/TWC.2013.061413.121703](https://doi.org/10.1109/TWC.2013.061413.121703).
- 张 鸿: 男, 1987年生, 博士, 研究方向为物联网、光无线融合网络等。
- 韩 霄: 男, 1996年生, 硕士生, 研究方向为光无线融合网络。
- 王汝言: 男, 1969年生, 教授, 博士, 研究方向为泛在网络、多媒体信息处理等。
- 李职杜: 男, 1990年生, 博士, 研究方向为边缘计算、网络演算等。
- 周 敏: 女, 1996年生, 硕士生, 研究方向为光无线融合网络。

责任编辑: 余 蓉