

接入与回传一体化小基站的接入控制与资源分配联合优化算法

唐伦 马润琳* 刘云龙 王耀玮 陈前斌

(重庆邮电大学通信与信息工程学院 重庆 400065)

(重庆邮电大学移动通信技术重点实验室 重庆 400065)

摘要: 针对全双工无线接入与回传一体化小基站场景下长期的频谱效率和能效同时最大化问题, 该文提出一种基于近似动态规划理论的接入与回传一体化小基站接入控制与资源分配联合优化算法。该算法首先联合考虑当前基站的资源使用和功率配置情况, 在任一用户需求动态到达以及平均时延、小基站回传速率和传输功率约束下, 使用受限马尔科夫决策过程(CMDP)建立频谱效率最大化和功率消耗最小化的多目标优化模型, 其次运用切比雪夫理论将多目标优化问题转化为单目标问题, 并使用拉格朗日对偶分解法进一步转化为非受限的马尔科夫决策过程(MDP)问题。最后, 为了解决其求解时存在的“维度灾”爆炸问题, 该文提出基于近似动态规划的无线接入与回传一体化小基站资源动态分配算法进行求解, 得到此时的接入与资源分配策略。仿真结果表明, 所提算法能在保证平均时延约束、小基站回传速率约束和传输功率约束的同时最大化长期平均频谱效率和能效。

关键词: 接入与回传一体化小基站; 近似动态规划理论; 受限马尔科夫决策过程; 切比雪夫理论

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2019)06-1389-08

DOI: 10.11999/JEIT180357

Joint Admission Control and Resource Allocation Algorithm for Access and Backhaul Integrated Small Base Station

TANG Lun MA Runlin LIU Yunlong WANG Yaowei CHEN Qianbin

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

(Key Laboratory of Mobile Communication Technology, Chongqing University of Post and Telecommunications, Chongqing 400065, China)

Abstract: To maximize the long-term spectral efficiency and energy efficiency of a full duplex wireless access and backhaul integrated small base station scene, approximate dynamic programming based joint admission control and resource allocation optimization algorithm is proposed. The algorithm firstly considers the resource usage and power configuration of the current base station, the dynamic demand of user, the constraints of average delay as well as backhaul rate and transmission power. The corresponding multi-objective optimization model of maximum spectrum efficiency and minimizes power consumption is established by using the Constrained Markov Decision Process(CMDP). Then, the Chebyshev theory is used to transform the multi-objective into a single-objective optimization, and the Lagrange dual decomposition method is then used to convert the single-objective problem into unrestricted Markov decision process problem. Finally, To solve the “dimension disaster” explosion that generated when solving this unrestricted Markov Decision Process(MDP) problem, a dynamic resource allocation algorithms based on approximate dynamic programming is presented, and the access and resource allocation strategy is obtained during this process. The simulation results show that the algorithm can maximize the long-term average spectrum efficiency and energy efficiency, within the constraints of the average delay, backhaul rate and transmission power, under the scenario of integrated access and backhaul small base station.

收稿日期: 2018-04-18; 改回日期: 2019-02-16; 网络出版: 2019-03-05

*通信作者: 马润琳 357135128@qq.com

基金项目: 国家自然科学基金(61571073), 重庆市教委科学技术研究项目(KJZD-M201800601)

Foundation Items: The National Natural Science Foundation of China (61571073), The Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-M201800601)

Key words: Access and backhaul integrated small base station; Approximate dynamic programming theory; Constrained Markov Decision Process (CMDP); Tchebycheff theory

1 引言

在5G超密集组网中, 频谱资源的严重匮乏和用户要求的高数据传输速率相矛盾, 如何在有限资源的限制下, 提高频谱利用率是一个很重要的问题^[1]。然而, 在宏基站(Macro Base Station, MBS)下大规模部署小基站(Small Base Station, SBS)会增加系统总能耗^[2]。如何对无线接入与回传一体化小基站资源进行高效的分配, 在最大化频谱效率的同时最小化基站的功率消耗具有重要的研究意义。

文献^[3]针对毫米波小基站中满足用户流量需求同时降低网络功耗的问题, 提出了一种启发式组合控制算法, 但小基站的接收和发送是严格限定在两个正交的时间或频率信道内。文献^[4]为在保证系统稳定性的同时最大化能效, 提出了一种eTrans算法, 但只考虑了能效。文献^[5]提出了一种免疫优化算法, 在无线接入与回传速率匹配的约束下实现小基站能效的最大, 但仅考虑了能效。文献^[6]根据信道状态信息为当前的用户需求确定资源分配策略以期节能优化。忽略了当前决策对后续到达的用户需求分配资源的影响。

基于以上研究, 本文首先联合考虑当前基站资源使用和功率配置状态, 在任一用户需求到达及保证平均时延约束、小基站回传速率和传输功率约束的前提下, 使用受限马尔科夫决策过程(Constrained Markov Decision Process, CMDP)建立能效最大化和功耗最小化的多目标优化模型。其次运用切比雪夫理论将其转化为单目标优化问题^[7], 并使用拉格朗日对偶分解法进一步转化为非受限的MDP问题。最后本文提出基于近似动态规划的动态资源分配算法(ADP-based Dynamic Resource Allocation, ADPDRA)进行求解, 得到此时的接入与资源分配策略。

2 系统模型和小基站接入控制与资源分配联合优化问题数学建模

2.1 场景部署

本文系统场景图如图1所示。有一个MBS, 使用大规模MIMO天线, 天线阵列数为 M , 波束赋形组大小是 N , 采用线性迫零波束赋形传输^[8], 有 $I-1$ 个单天线无线SBS。因此基站总数为 I , 基站 $i \in \{1, 2, \dots, I\}$, 每个基站有最大发射功率 p_{\max} , 当 $i=1$ 表示MBS, 当 $i \neq 1$ 表示SBS。为提高频谱利用率, SBS为带内自回传小基站。为使SBSs有效地进

行自回传^[9], 假设每个自回传SBS均配有全双工硬件且工作在全双工模式。MBS和SBSs都有一个可容忍最大队列长度 B 为500个数据包的有限缓存空间。有 J 个单天线用户(Users, UEs) $j \in \{1, 2, \dots, J\}$ 可任意接入任一基站。

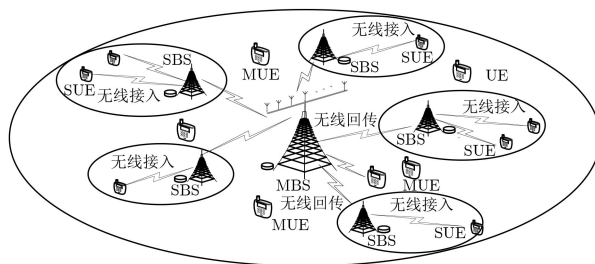


图1 系统场景图

2.2 用户业务模型

时隙 t 时, 宏基站用户(Macro Users, MUEs) j 在MBS处的队列为 $Q_{1,j}(t)$, 则下一时隙MUEs j 在MBS的队列长度为

$$Q_{1,j}(t+1) = \min[Q_{1,j}(t) + a_{1,j}(t) - b_{1,j}(t), B] \quad (1)$$

其中, $a_{1,j}(t)$ 为MUEs j 时隙 t 到达MBS的数据包数量。 $b_{1,j}(t)$ 为时隙 t MBS向MUE j 发送的数据包数量, 且有 $b_{1,j}(t) = r_{ij}/S$, $i=1$, 其中数据包长度 S 值为200 bit, $r_{ij}(t)$ 为用户 j 接入MBS的传输速率。

本文联合考虑MBS和SBS一个整体系统处小基站用户(Small Users, SUEs)的排队情况, 则下一时隙SUEs队列更新为

$$Q_{i,j}(t+1) = \min[Q_{i,j}(t) + a_{i,j}(t) - b_{i,j}(t), 2B], i > 1 \quad (2)$$

其中, $Q_{i,j}(t) = Q_j(t) + D_{i,j}(t)$ 为时隙 t SUEs在联合MBS处的队列长度 $Q_j(t)$ 和SBS处的队列长度 $D_{i,j}(t)$, $a_{i,j}(t)$ 为SUEs j 在时隙 t 到达宏基站的数据包数量。 $b_{i,j}(t)$ 为时隙 t SBS i 发送SUEs j 的数据包数, 且 $b_{i,j}(t) = r_{ij}/S$, $i > 1$, 其中 $r_{ij}(t)$ 为SUEs j 接入小基站的传输速率。 $2B$ 表示联合MBS处SUEs和SBS处SUEs的最大待传数据队列积压值。

则用户 j 下一时隙队列长度更新为

$$Q_{i,j}(t+1) = \min[Q_{i,j}(t) + a_{i,j}(t) - b_{i,j}(t), 2B], \\ i=1, 2, \dots, I, j=1, 2, \dots, J \quad (3)$$

2.3 能效和能效多目标问题数学建模

定义基站 i 时隙 t 传输功率是 $p_i(t)$, 且 $p_i(t) \leq p_{\max}$ 。UE j 接入MBS时单位带宽接入速率

$$r_{ij}(t) = N \log_2(1 + \text{SINR}_{ij}(p_i(t))), i=1 \quad (4)$$

其中, MUEs信干噪比 $\text{SINR}_{ij}(p_i(t))$ 是

$$\text{SINR}_{ij}(p_i(t)) = \frac{M - N + 1}{N} \cdot \frac{p_i(t)g_{ij}(t)}{\sum_{l \neq i} p_l(t)g_{lj}(t) + \sigma^2}, i=1 \quad (5)$$

其中, $g_{ij}(t)$ 是时隙 t 基站 i 到UE j 的信道增益, $g_{lj}(t)$ 是时隙 t 基站 l 到UE j 的信道增益, σ^2 为噪声功率, M 为天线阵列数, N 为波束赋形组大小。

当SBSs $i, i \neq 1$ 接入MBS时单位带宽回传速率

$$r_i(t) = \log_2 \left(1 + \frac{M - N + 1}{N} \text{SINR}_i(t) \right), i \neq 1 \quad (6)$$

SBSs信干噪比为

$$\text{SINR}_i(t) = \frac{p_1(t)g_i(t)}{\sum_{l \neq i, l \neq 1} p_l(t)g_l(t) + \theta p_i(t) + \sigma^2}, i \neq 1 \quad (7)$$

其中, $p_1(t)$ 是时隙 t 时MBS的传输功率, $g_i(t)$ 是时隙 t 时MBS到第 $i(i \neq 1)$ 个SBSs的信道增益, $\theta p_i(t)$ 为时隙 t 小基站的自干扰功率, $p_i(t)$ 为时隙 t 时SBSs的发送功率, $0 \leq \theta < 1$ 为自干扰消除系数。

UE j 接入SBSs时单位带宽的接入速率

$$r_{ij}(t) = \log_2 (1 + \text{SINR}_{ij}(t)), i > 1 \quad (8)$$

SUEs信干噪比为

$$\text{SINR}_{ij}(t) = \frac{p_i(t)g_{ij}(t)}{\sum_{l \neq i} p_l(t)g_{lj}(t) + \sigma^2}, i > 1 \quad (9)$$

假定每个UE在时隙 t 最多只能连接一个基站, 用二进制变量 $x_{ij}(t)$ 表示用户关联状态, 当UE j 接入基站 i 时, $x_{ij}(t)=1$, 否则 $x_{ij}(t)=0$ 。

定义UE j 和基站 i 通信时用的资源比例为 $y_{ij}(t)$, 且有 $0 \leq y_{ij}(t) \leq 1$ 。则UE j 接入传输速率

$$R_j(t) = \sum_i x_{ij}(t)y_{ij}(t)r_{ij}(t), j=1, 2, \dots, J \quad (10)$$

SBSs $i, i \neq 1$ 接入传输速率

$$R_i(t) = \sum_j x_{ij}(t)y_{ij}(t)r_{ij}(t), i=2, 3, \dots, I \quad (11)$$

SBSs $i, i \neq 1$ 回传传输速率

$$C_i(t) = y_{ij}(t)r_i(t), i \neq 1 \quad (12)$$

为避免UEs数据包在SBS处无限积压而引起数据包的丢失, 在时隙 t 上SBS回传速率需满足

$$R_i(t) \leq C_i(t), i \in I, \forall t \quad (13)$$

定义系统总功率损耗

$$P_{\text{to}}^s(t) = P_{\text{ac}}^s(t) + P_{\text{bh}}^s(t) \quad (14)$$

其中, 回传链路的功耗 $P_{\text{bh}}^s(t)$, 主要是MBS功耗为

$$P_{\text{bh}}^s(t) = P_i^s(t) = \mu_i p_i(t) + M p_t + p_{i,s}, i=1 \quad (15)$$

接入链路的功耗 $P_{\text{ac}}^s(t)$, 主要是SBS功耗为

$$P_{\text{ac}}^s(t) = \sum_i (\mu_i p_i(t) + p_t + p_{i,s}), i > 1 \quad (16)$$

其中, μ_i 为基站 i 功率放大器的损耗系数, p_t 是时隙 t 每个天线的功耗, $p_{i,s}$ 为基站固定功耗。

定义时隙 t 系统频谱效率(Spectral Efficiency, SE)为单位带宽的总吞吐量, 单位是bit/(s·Hz), 则

$$\eta_{\text{SE}}(t) = \sum_j R_j(t) = \sum_j \sum_i x_{ij}(t)y_{ij}(t)r_{ij}(t) \quad (17)$$

定义时隙 t 系统能量效率(Energy Efficiency, EE)为SE与系统总功耗的比, 单位是bit/(J·Hz), 则

$$\eta_{\text{EE}}(t) = \frac{\eta_{\text{SE}}(t)}{P_{\text{to}}(t)} = \frac{\sum_j \sum_i x_{ij}(t)y_{ij}(t)r_{ij}(t)}{P_{\text{ac}}^s(t) + P_{\text{bh}}^s(t)} \quad (18)$$

考虑到用户公平性, 使用用户速率对数和^[8]

$$W(x_{ij}, y_{ij}, p_i) = \sum_j \ln(R_j(t)) \quad (19)$$

因此, 以 S_0 为初始状态, 无限范围系统长期的折扣频效为

$$W(S_0) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\sum_{t=0}^{T-1} \alpha^t W(a(t), S(t)) | S_0 \right] \quad (20)$$

其中, α 为折扣因子, $W(a(t), S(t))$ 为当前系统状态是 $S(t)$, 采取行动 $a(t)$ 时系统的瞬时频效。

系统长期的折扣功率消耗为

$$P_{\text{to}}^s(S_0) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\sum_{t=0}^{T-1} \alpha^t P_{\text{to}}^s(a(t), S(t)) | S_0 \right] \quad (21)$$

第 j 个用户队列积压的长期折扣时延^[9]为

$$Q_{i,j}(S_0) = \limsup_{T \rightarrow \infty} \mathbb{E} \left[\sum_{t=0}^{T-1} \alpha^t Q_{i,j}(a(t), S(t)) | S_0 \right] \quad (22)$$

其中, $Q_{i,j}(a(t), S(t))$ 为当前系统状态 $S(t)$, 采取行动 $a(t)$ 时用户 j 的队列长度。 $Q_{i,j}(S_0)$ 表示用户 j 的平均时延。为保证用户平均时延, 需满足约束条件 $Q_{i,j}(S_0) \leq \varepsilon_{i,j}$, $\varepsilon_{i,j}$ 为平均队列长度的门限值。

则本文中最大化系统SE的同时最小化其系统功率总消耗的多目标函数即可表示为

$$\left. \begin{array}{l} \min -W(S_0) \\ \min P_{\text{to}}^s(S_0) \\ \text{s.t. C1: } Q_{i,j}(S_0) \leq \varepsilon_{i,j}, i \in I, j \in J \\ \text{C2: } p_i(t) \leq p_{\text{max}}, i \in I, \forall t \\ \text{C3: } R_i(t) \leq C_i(t), i \in I, \forall t \end{array} \right\} \quad (23)$$

其中, C1为用户平均时延约束条件, C2为基站传输功率约束条件, C3为SBS回传速率约束条件。

文献[7]中切比雪夫理论, 将式(23)多目标(MOO)转化为单目标(SOO)优化问题。先进行无量纲化, 得

$$\left. \begin{aligned} \min \quad & F_1(S_0) = \frac{W_{\max} - W(S_0)}{W_{\max} - W_{\min}} \\ \min \quad & F_2(S_0) = \frac{P_{\text{to}}^s(S_0)}{P_{\max}^s} \\ \text{s.t.} \quad & \text{C1, C2, C3} \end{aligned} \right\} \quad (24)$$

其中, P_{\max}^s 为每个基站功率是 p_{\max} 时系统长期折扣功耗的最大值; W_{\max} 为不考虑各个干扰时系统长期折扣频效的最大值; 系统长期折扣频效的最小值 $W_{\min} = J \ln(\delta)$, δ 是预先定义的一个足够小的值。

则式(23)可转化为SOO如式(25)

$$\left. \begin{aligned} \min \quad & \max\{\omega F_1(S_0), (1 - \omega)F_2(S_0)\} \\ \text{s.t.} \quad & \text{C1, C2, C3} \end{aligned} \right\} \quad (25)$$

其中, ω 是加权系数, 取值为 $0 \leq \omega \leq 1$ 。

定理 1 对于任意给定的 $\omega \in [0, 1]$, 若单目标问题式(25)存在唯一的最优解 $(x_{ij}^*, y_{ij}^*, p_i^*)$, 则此最优解也是初始MOO问题(23)的帕累托Pareto最优解^[7]。证明略。

使用拉格朗日法将CMDP问题(式(25))转化为非受限的MDP问题^[10], 表示为

$$J(\lambda) = \limsup_{T \rightarrow \infty} E \left[\sum_{t=0}^{T-1} \alpha^t \{g(\lambda, S(t), a(t))\} \right] \quad (26)$$

其中, λ 为约束条件C1引入的拉格朗日乘子矢量。

其中,

$$\begin{aligned} & g(\lambda, S(t), a(t)) \\ &= \max_{a(t) \in A(S(t))} \left\{ \omega \frac{W_{\max} - W(a(t), S(t))}{W_{\max} - W_{\min}}, \right. \\ & \quad \left. (1 - \omega) \frac{P_{\text{tot}}^s(a(t), S(t))}{P_{\max}^s} \right\} \\ & \quad + \sum_{i \in I, j \in J} \lambda_{ij} (Q_{i,j}(a(t), S(t)) - \varepsilon_{i,j}) \end{aligned} \quad (27)$$

为 t 时隙状态为 $S(t)$ 时, 采取行动 $a(t)$ 的单步回报函数。且有其行动空间 $A(s(t))$ 满足约束条件C2: $p_i(t) \leq p_{\max}, i \in I, \forall t$ 和C3: $R_i(t) \leq C_i(t), i \in I, \forall t$ 。为了最大化长期平均SE的同时最小化其系统总功耗, 下面进行MDP建模。

2.4 基于MDP的接入控制与资源分配联合优化问题建模

(1) 状态空间 \mathcal{S} : 时隙 t 全局状态空间 $S(t)$ 由时

隙 t 用户 j 在基站处队列积压长度 $Q_{i,j}(t)$ 和时隙 t 新用户请求到达前基站 i 的频谱和功率资源使用状态 $U_i(t)$ 组成, 即 $S(t) = \{Q_{i,j}(t), U_i(t)\}$ 。每个基站资源容量为 $C_{R,i}$ 。

(2) 行动空间 A : MDP的一个策略 π 是一个从状态空间 \mathcal{S} 到行动空间 A 的映射 $\pi: \mathcal{S} \rightarrow A$ 。定义为

$$\pi(S(t)) = a(t) \in A(S(t)) \quad (28)$$

$A(S(t))$ 表示状态为 $S(t)$ 时, 对当前状态可采取的行动集合。其中, 每个决策 $a(t)$ 为

$$\begin{aligned} a(t) &= (x_{ij}(t), \beta_{ij}(t), p_i(t)) \\ &= \begin{cases} (1, y_{ij}(t)C_{R,i}, p_i(t)), & \text{基站 } i \text{ 接收 UE } j \text{ 的接入} \\ (0, 0, 0), & \text{基站 } i \text{ 拒绝 UE } j \text{ 的接入} \end{cases} \end{aligned} \quad (29)$$

其中, $\beta_{ij}(t) = y_{i,j}(t)C_{R,i}$ 为时隙 t 基站 i 分给新到达用户 j 的无线资源量。行动空间满足约束条件 $A = \{(x_{ij}(t), \beta_{ij}(t), p_i(t)) : p_i(t) \leq p_{\max}; R_i(t) \leq C_i(t), i \in I, \forall t$

(3) 状态转移概率 $P(S(t+1)|S(t), a(t))$: 假设当前有用户需求到达时, 在当前系统状态 $S(t)$ 下, 根据其所对应的决策 π 采取行动 $a(t)$ 后, 转移到下一状态 $S(t+1)$ 的概率为

$$\begin{aligned} & P(S(t+1)|S(t), a(t)) \\ &= P(Q_{i,j}(t+1)|Q_{i,j}(t), a(t)) \\ & \quad \cdot P(U_i(t+1)|U_i(t), a(t)) \end{aligned} \quad (30)$$

(4) 回报函数 $g(\lambda, S(t), a(t))$: 时隙 t , 状态 $S(t)$ 采取行动 $a(t)$ 后到下一状态 $S(t+1)$ 的最大单步收益函数 $g(\lambda, S(t), a(t))$ 。

3 优化性能与算法分析

3.1 优化性能分析

定义最优值函数为

$$V^*(S) = \min_{a \in A(S)} \left\{ g(\lambda, S, a) + \alpha \sum_{S' \in \mathcal{S}} P(S'|S, a) V^*(S') \right\} \quad (31)$$

系统中不能确定状态转移概率, 故而引入决策后状态 S^a 。定义下一状态 S' 、当前决策后状态 S^a 对应的最优值函数是 $V^*(S')$, $V^a(S^a)$, 则有 $V^a(S^a) = E[V^*(S')]$ 。因此可将式(30)改为

$$V^*(S) = \min_{a \in A(S)} \{g(\lambda, S, a) + \alpha V^a(S^a)\} \quad (32)$$

3.2 算法分析

常规求解式(31)值函数的方法在用户数增加时会存在“维度灾”。本文用近似动态规划(Approximate Dynamic Programming, ADP)方法求解。

定义影响决策的关键特征集为 F , 定义基函数如表1, 基函数是关于决策后状态 $S^a(t)$ 中某一个特征 $f, f \in F$ 的函数, 记为 $\phi_f(S^a(t))$ 。定义基函数向量 $\Phi = (\phi_1(S^a(t)), \phi_2(S^a(t)), \dots, \phi_{|F|}(S^a(t)))^T$, 可将决策后状态 $S^a(t)$ 最优值表示为

$$V^a(S^a(t)) \approx \mu^T \Phi = \sum_{f \in F} \mu_f \phi_f(S^a(t)) \quad (33)$$

其中, μ_f 是待估的回归参数, $\mu = (\mu_1, \mu_2, \dots, \mu_{|F|})^T$ 是待估的回归参数向量。则式(32)可改写为

$$V^*(S) = \min_{A(S(t))} \left\{ g(\lambda, S, a) + \alpha \sum_{f \in F} \mu_f \phi_f(S^a(t)) \right\} \quad (34)$$

表1 基函数的定义

基函数	含义
1	常数
$Q_{i,j}(t)$	用户 j 在基站 i 处队列积压的长度
$U_i(t)$	新用户请求到达前基站 i 的资源使用状态
$[Q_{i,j}(t)]^2$	用户 j 在基站 i 处队列积压长度的平方
$Q_{i,j}(t)U_i(t)$	用户 j 在基站 i 处队列积压长度和新用户请求到达前基站 i 资源使用状态的乘积
$[U_i(t)]^2$	新用户请求到达前基站 i 资源使用状态的平方

参考文献[11], 样本值 $\hat{\mu}_f$ 表示为

$$\hat{\mu}_f = \widehat{V^\pi(S)} - \widehat{V^\pi(S')} | S=S(t), S'=S(t+1) \quad (35)$$

其中, $\widehat{V^\pi(S)}$ 为时隙 t 用TD(0)方法取的值函数样本, $\widehat{V^\pi(S')}$ 为 $t+1$ 时隙 $S(t+1)$ 的估计样本值函数。

当观察到 $\widehat{V^\pi(S)}$ 后, 无线资源调度器根据梯度法沿着梯度方向更新 μ , 关于待估参数向量 μ 梯度

$$\begin{aligned} & \nabla_{\mu} V^a(S^a(t-1) | \mu) \\ &= (\phi_1(S^a(t-1)), \phi_2(S^a(t-1)), \dots, \phi_{|F|}(S^a(t-1))) \\ &= \Phi(S^a(t-1)) \end{aligned} \quad (36)$$

具体的随机梯度法更新待估参数算法如表2所示, 且待估回归参数 μ_f 通过采样获得, 具体方法如表3所示。

本文使用的ADP易导致算法的局部收敛。因此, 本文算法用 ε -贪心策略对行为空间进行探索

$$\pi(S) \leftarrow \begin{cases} 1 - \varepsilon + \frac{\varepsilon}{|A(S(t))|}, & a(t) = \pi^*(S) \\ \frac{\varepsilon}{|A(S(t))|}, & a(t) \neq \pi^*(S) \end{cases} \quad (37)$$

其中, ε 为“探索因子”, 探索率值为 $\varepsilon=0.1$ 。

综上所述, 基于近似动态规划的接入与回传一体化小基站资源动态分配算法可总结如表4。具体执行过程描述如表4所示。

表2 基于随机梯度法更新待估参数的样本值函数逼近

近似值函数算法

- (1) 输入目标函数:

$$\mu^* = \arg \min_{\mu} E \left\{ \frac{1}{2} [V^a(S^a(t-1) | \mu) - \widehat{V^\pi(S)}]^2 \right\}$$
- (2) 初始化步长 γ_t , 均方差误差门限 δ 等;
- (3) 使用TD(0)方法取样本值函数的初始值;
- (4) 用数值微分法对待估参数 μ_f 取样本 $\hat{\mu}_f$;
- (5) 计算当前状态 $S(t)$ 的值函数样本 $\widehat{V^\pi(S)}$;
- (6) 根据式(36), 按随机梯度法沿着梯度方向更新待估参数向量 μ , γ_t 值为 $0.1 \sim 0.7$;
- (7) 根据式(33), 求最优近似值函数;
- (8) 判定(1)中目标函数是否小于 $\delta=10^{-5}$, 若不满足, 则执行(6), 否则依次执行;
- (9) 获得最接近样本值函数的近似值函数 V^a 。

表3 基于蒙特卡罗方法的外部用户需求随机变量采样算法

- (1) 初始化采样路径 $k, k=1$, 总采样次数 $K, K=200$, 折扣因子 α 等参数;
- (2) for $k=1$ to K do
- (3) 选择一条采样路径 ω^k ;
- (4) 对于每一个采样时刻 $t=1, 2, \dots, 20$, 令 $\omega_t^k = W_t(\omega^k)$ 表示采样的实现;
- (5) 得到数据包数 $a_{i,j}(t)$ 与无线资源量 $\beta_{ij}(t)$;
- (6) $k=k+1$, 重复1。

表4 基于近似动态规划的动态资源分配算法

- (1) 对每个状态 $S(t)$, 初始化其初始值函数 $V^*(S)$, 探索率 $\varepsilon=0.1$ 等参数;
- (2) for $t=0$ to T do
- (3) 根据式(37), 使用 ε -贪心策略对行为空间进行探索;
- (4) 根据式(33)求解, 并令此时的 $\pi^*(S)$ 为以上优化问题的解;
- (5) 根据表2求近似值函数 V^a , 得到此时的接入与频谱和功率资源分配策略;
- (6) 根据表3进行1次采样;
- (7) 根据梯度法更新拉格朗日乘子 λ_{ij} ;
- (8) 更新 $t=t+1$, 并更新 $t+1$ 时隙的系统状态, 重复3。

4 性能仿真与结果分析

为验证本文为无线接入与回传一体化小基站系统在相对长期的资源分配中获得更高的频效和能效而提出基于近似动态规划的动态资源分配算法ADPDRA和多个用户的仿真场景并进行仿真, 先证所提算法的收敛性, 再将文献[12]中贪心算法(Greedy Algorithm, GA)和文献[13]中值迭代算法(Value Iteration algorithm, VI)、轮询调度算法(Round-Robin scheduling algorithms, RR)作为基准和本文ADPDRA算法对比分析。参数设置如表5。

图2为后决策状态近似值函数线性拟合图, 随

表5 仿真参数

参数	数值
信道模型	瑞利衰落
噪声功率谱密度 N_0	-174 dBm/Hz
训练周期 T	30
宏基站发送功率 P_0	43 dBm
小基站发送功率 P	33 dBm
基站个数 I	4
时隙时长	1 ms
总用户数	30, 50, 70
折扣因子 α	0.3, 0.5, 0.9
天线阵列数为 M	100
波束成形组大小 N	10
平均队列长度门限值 $\varepsilon_{i,j}$	28, 40, 75

采样次数的增加, 后决策状态近似估计值函数由不断上升逐渐趋于稳定, 越逼近系统决策后状态值函数的实际值, 也可证所提算法的收敛性。由于外部时隙 t 用户到达的数据包数量 $a_{i,j}(t)$ 、基站分给新到达用户的无线资源量 $\beta_{ij}(t)$ 是非稳态的随机变量, 故而在同一次采样过程的不同时隙, 相应的近似估计值函数也就不同。

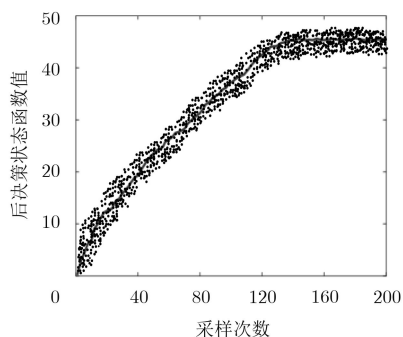


图2 近似估计值函数线性拟合图

图3为不同用户数下频谱效率和能量效率的关系。同一用户数下, 当SE增加, EE先增加并达到一个最大值, 后迅速减到一个较低点。当SE较小时, SE增加速度比系统总功耗增加速度要快得多。但EE有一个最大值点。在此之后, SE的增加速度变得比系统总功耗增加速度慢。当SE一定时, EE随用户数 J 的增加而增加, 且增加趋势越来越小。因为用户数增加时, 要达到同样SE, 功率消耗会减少。

图4刻画了连续3000个时隙上的数据队列积压, 用户数 J 为30, 时隙增加会引起数据队列积压的增大, 当时隙到一定值后, 数据队列积压将趋于其最大长度约束值, 说明了本文算法能保证平均时

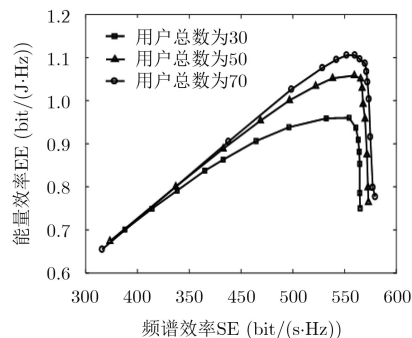


图3 频谱效率SE和能量效率EE的关系

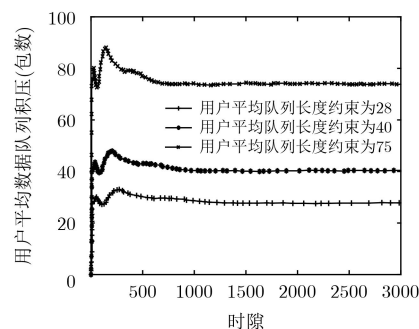
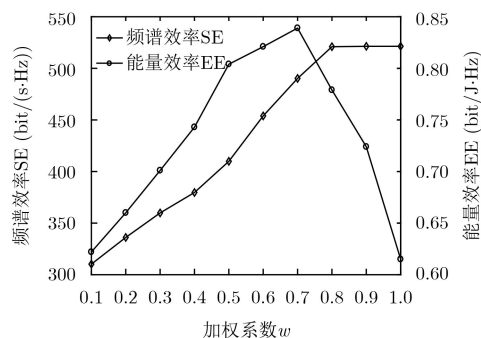


图4 用户在连续时隙上的平均数据队列积压

延约束。当时隙一定时, 数据积压队列长度随用户时延约束的增加而增加。

图5为频谱效率和能量效率在不同系数 ω 下的变化关系, 随着 ω 的增大, EE先达到峰值, 后降到一更低的值。原因是加权系数 ω 作为频效与能效的平衡系数, 当 ω 增加时, 系统更倾向于SE, 此时无线资源调度器会动态调整小基站资源分配策略, 用更高的系统传输功率换取SE的提升。但是增大传输功率的同时, 会增大系统的功耗, 会降低EE。

图6比较了4种算法下频谱效率和能量效率比较情况。当SE增加时, EE先增加到峰值后快速减少。当SE较小时, ADPDRA, VI算法的EE相差很大, 当SE超过一定值时, ADPDRA算法的EE已很逼近VI的EE, 同时, ADPDRA算法的复杂度要远

图5 频谱效率SE和能量效率EE随加权系数 ω 的变化

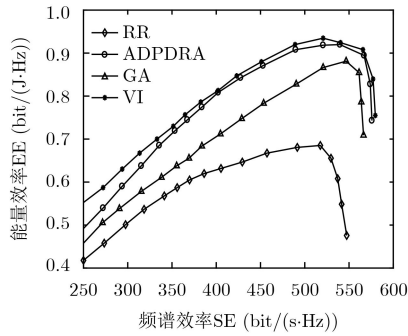


图6 4种算法的频谱效率SE和能量效率EE比较

小于VI算法。ADPDRA, VI和GA这三个方案随着SE的增加,在一定范围内EE都会提升,因为VI是基于队列状态信息动态分配资源的算法,ADPDRA是在一定时延约束下,为任一动态到达的用户需求分配最佳的接入与资源,GA在动态资源分配时,总是优先服务队列积压长度最长的用户服务。而相比较另外两种方案,因为ADPDRA采用了基于近似动态规划的资源分配算法,且考虑了用户数据包到达的随机性、数据包传输的时延性。而RR是周期性地把资源分配给每个用户,未考虑小基站资源剩余量对资源分配的影响,也未考虑用户时延约束。因此,在同SE下,RR算法的EE整体值最小,且队列稳定性最差。

从图7可看出,平均到达率增加会引起平均队列积压增加。在GA,ADPDRA和RR3个算法中,ADPDRA队列积压增加的速度最慢,且ADPDRA队列积压最接近VI的平均队列积压。原因在于RR方案没考虑用户时延约束和小基站资源剩余量对资源分配的影响。算法GA忽略了当前所做的决策对后续到达用户需求分配资源的影响。VI算法是一种联合队列状态信息、数据包传输时延特性的算法,但其复杂度高。而本文提出的ADPDRA算法联合考虑当前请求接入基站的资源使用和用户需求,做出此时的资源分配决策。

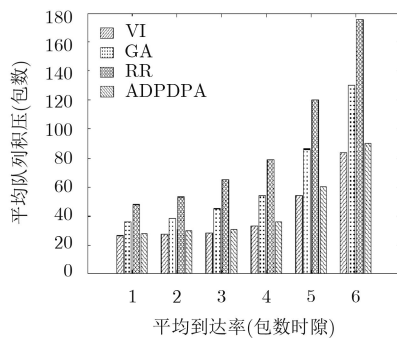


图7 4种算法的平均队列积压比较

5 结束语

针对无线接入与回传一体化小基站场景下相对长期的频效和能效最大化问题,本文提出了一种基于近似动态规划的接入与回传一体化小基站接入控制与资源联合动态优化算法(ADPDRA),该算法联合考虑当前基站的资源使用和功率配置情况并使用马尔科夫决策过程建立最大化频效和最小化功耗的多目标优化模型,然后运用切比雪夫理论将原问题转化为单目标问题,并使用拉格朗日对偶分解法进一步转化为非受限的MDP问题,最后,本文提出基于近似动态规划的无线接入与回传一体化小基站资源动态分配算法求出每一时隙的策略。仿真结果表明,本文算法能在保证平均时延约束、小基站回传速率约束和传输功率约束的同时最大化长期平均频效和能效。

参考文献

- [1] BENCHABENE Y, BOUJNAH N, and ZARAI F. 5G cellular: Survey on some challenging techniques[C]. Proceedings of the 18th International Conference on Parallel and Distributed Computing, Applications and Technologies, Taipei, China, 2017: 348-353. doi: [10.1109/PDCAT.2017.00064](https://doi.org/10.1109/PDCAT.2017.00064).
- [2] MESODIAKAKI A, ADELANTADO F, ALONSO L, et al. Energy- and spectrum-efficient user association in millimeter-wave backhaul small-cell networks[J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(2): 1810-1821. doi: [10.1109/TVT.2016.2565539](https://doi.org/10.1109/TVT.2016.2565539).
- [3] OGAWA H, TRAN G K, SAKAGUCHI K, et al. Traffic adaptive formation of mmWave meshed backhaul networks[C]. Proceedings of 2017 IEEE International Conference on Communications Workshops, Paris, France, 2017: 185-191. doi: [10.1109/ICCW.2017.7962655](https://doi.org/10.1109/ICCW.2017.7962655).
- [4] LI Yuzhou, SHI Yan, SHENG Min, et al. Energy-efficient transmission in heterogeneous wireless networks: A delay-aware approach[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(9): 7488-7500. doi: [10.1109/TVT.2015.2472578](https://doi.org/10.1109/TVT.2015.2472578).
- [5] LIU Yanping and FANG Xuming. Joint user association and resource allocation for self-backhaul ultra-dense networks[J]. *China Communications*, 2016, 13(2): 1-10. doi: [10.1109/CC.2016.7405718](https://doi.org/10.1109/CC.2016.7405718).
- [6] YANG H H, GERACI G, and QUEK T Q S. Energy-efficient design of MIMO heterogeneous networks with wireless backhaul[J]. *IEEE Transactions on Wireless Communications*, 2016, 15(17): 4914-4927. doi: [10.1109/TWC.2016.2549529](https://doi.org/10.1109/TWC.2016.2549529).
- [7] MASOUDI M, ZAEFARANI H, MOHAMMADI A, et al.

- Energy and spectrum efficient resource allocation in two-tier networks: A multiobjective approach[C]. Proceedings of 2017 IEEE Wireless Communications and Networking Conference, San Francisco, USA, 2017: 1–6. doi: [10.1109/WCNC.2017.7925663](https://doi.org/10.1109/WCNC.2017.7925663).
- [8] LI Yi, FAN Pingzhi, LIU Lingjia, *et al.* Distributed MIMO precoding for in-band full-duplex wireless backhaul in heterogeneous networks[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(3): 2064–2076. doi: [10.1109/TVT.2017.2713413](https://doi.org/10.1109/TVT.2017.2713413).
- [9] TANG Lun, WEI Yannan, CHEN Wan, *et al.* Delay-aware dynamic resource allocation and ABS configuration algorithm in HetNets based on Lyapunov optimization[J]. *IEEE Access*, 2017, 5: 23764–23775. doi: [10.1109/ACCESS.2017.2761863](https://doi.org/10.1109/ACCESS.2017.2761863).
- [10] ZORDAN D, MELODIA T, and ROSSI M. On the design of temporal compression strategies for energy harvesting sensor networks[J]. *IEEE Transactions on Wireless Communications*, 2016, 15(2): 1336–1352. doi: [10.1109/TWC.2015.2489200](https://doi.org/10.1109/TWC.2015.2489200).
- [11] DEVRAJ A M and MEYN S P. Differential TD learning for value function approximation[C]. Proceedings of 2016 IEEE 55th Conference on Decision and Control, Las Vegas, USA, 2016: 6347–6354. doi: [10.1109/CDC.2016.7799246](https://doi.org/10.1109/CDC.2016.7799246).
- [12] XIANG Lin, NG D W K, SCHOBBER R, *et al.* Secure video streaming in heterogeneous small cell networks with untrusted cache helpers[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(4): 2645–2661. doi: [10.1109/TWC.2018.2800005](https://doi.org/10.1109/TWC.2018.2800005).
- [13] MI Xiang, ZHAO Ming, XIAO Limin, *et al.* Delay-aware resource allocation and power control for device-to-device communications[C]. Proceedings of 2015 IEEE Wireless Communications and Networking Conference Workshops, New Orleans, USA, 2015: 115–122. doi: [10.1109/WCNCW.2015.7122573](https://doi.org/10.1109/WCNCW.2015.7122573).
- 唐 伦: 男, 1973年生, 教授, 博士, 研究方向为新一代无线通信网络、异构蜂窝网络、软件定义无线网络等。
- 马润琳: 女, 1993年生, 硕士生, 研究方向为5G网络切片、无线资源分配。
- 刘云龙: 男, 1992年生, 硕士生, 研究方向为5G无线自回传网络中的资源分配问题。
- 王耀玮: 男, 1991年生, 硕士生, 研究方向基于卷积神经网络的车辆识别问题。
- 陈前斌: 男, 1967年生, 教授, 博士生导师, 研究方向为个人通信、多媒体信息处理与传输、下一代移动通信网络等。