

## 基于带缓存的云接入网络最优能效设计

孙 远 李春国 黄永明 杨绿溪\*

(东南大学信息科学与工程学院 南京 210096)

**摘要:** 在云接入网络(Cloud-RAN)中, 现有工作大多假定射频拉远头(RRH)不具备缓存功能。然而下一代通信网络具有以内容为中心的特性, 因此在Cloud-RAN中考虑带缓存的RRHs也变得有必要。该文考虑在Cloud-RAN中有效设计缓存方案, 并通过资源分配有效减轻前程链路负担。假设系统采用正交频分多址接入(OFDMA)技术, 通过联合优化子载波(SC)分配, RRH选择与传输功率, 最小化系统下行总功耗, 并通过拉格朗日对偶分解转化非凸问题, 获得最优分配方案。仿真结果表明, 比起其它缓存方案, 该文提出的优化算法可以有效地提升系统能效, 满足未来通信需求。

**关键词:** 云接入网络; 缓存; 能效; 算法设计

中图分类号: TN92

文献标识码: A

文章编号: 1009-5896(2019)07-1525-08

DOI: 10.11999/JEIT180722

## Optimal Energy-efficient Design for Cache-based Cloud Radio Access Network

SUN Yuan LI Chunguo HUANG Yongming YANG Lixi

(School of Information Science and Engineering, Southeast University, Nanjing 210096, China)

**Abstract:** In Cloud Radio Access Network (Cloud-RAN), most of the existing work assumes Remote Radio Heads (RRHs) could not cache content. To better adapt the content-centric feature of next-generation communication networks, it is necessary to consider caching function for RRHs in Cloud-RAN. Motivated by this, this paper intends to design the suitable caching schemes and reduce the burden of fronthaul link burden through resource allocation. It is assumed the system utilizes Orthogonal Frequency Division Multiple Access (OFDMA) technique. A jointly optimization scheme of SubCarrier (SC) allocation, RRH selection, and transmission power is proposed to minimize the total downlink power consumption. To transform the original non-convex problem, a Lagrange dual decomposition is utilized to design the optimal allocation scheme. The experimental results show that the proposed algorithms can effectively improve the energy efficiency of the system, which meets the requirements of green communication in the future.

**Key words:** Cloud Radio Access Network (Cloud-RAN); Cache; Energy Efficiency (EE); Algorithm design

### 1 引言

云接入网络(Cloud Radio Access Network, Cloud-RAN), 作为中国移动首先提出的技术, 有希望是未来5G标准的候补<sup>[1]</sup>。在Cloud-RAN中, 基带处理部分被聚集并且共享在一个虚拟的基带单元(Base Band Unit, BBU)池。作为软中继的射频拉远头(Remote Radio Heads, RRHs), 可以通过有

线/无线前程连接, 将移动用户设备(User Equipment, UE)接收来的信号压缩并转发至中央BBU池。与分布式天线架构<sup>[2]</sup>和大规模天线系统<sup>[3]</sup>不同, 在Cloud-RAN系统中, 分布在RRHs与BBU池中间的前程链路容量是有限的。为了在前程容量约束的条件下提升Cloud-RAN性能, 已有相关文献进行信号量化/压缩相关技术的研究<sup>[4,5]</sup>。但减轻前程链路负担的途径不止一条, 通过设计带缓存功能RRHs, 同样可以在前程链路容量受限的条件下, 提升系统性能。尽管Cloud-RAN中的绿色通信技术也有一些相关的工作<sup>[6,7]</sup>, 但大多数已有的工作集中于研究RRHs不带缓存功能场景。在Cloud-RAN研究领域, 在“RRHs带缓存”场景下, 如何有效设计资源分配方案, 提升系统性能, 现有的工作十分有限。

收稿日期: 2018-07-18; 改回日期: 2019-01-28; 网络出版: 2019-02-15

\*通信作者: 杨绿溪 lxyang@seu.edu.cn

基金项目: 国家863计划(2015AA01A703), 国家自然科学基金(61372101, 61671144)

Foundation Items: The National 863 Project of China (2015AA01A703), The National Natural Science Foundation of China (61372101, 61671144)

在现有带缓存的Cloud-RAN文献中,文献[8]首先从性能分析角度推导了无线信道的有效容量,并建立了嵌套联合形式博弈模型,进行资源的联合优化。文献[9]将稀疏多播波束成型技术引入带缓存的Cloud-RAN,建立了混合整数非线性规划问题,比较了不同传播、缓存方式的优劣。华为的美国研发团队与法国研发团队分别利用YouTube与优酷网的数据,从协作分层缓存<sup>[10]</sup>与主动缓存<sup>[11]</sup>两个不同的角度,设计了优化缓存方式,提升相关性能指标。文献[12]从更普遍的移动多媒体服务需求出发,研究了在线资源分配、内容放置与请求路径的联合优化问题。文献[13]的研究兴趣集中在文件传输,通过联合优化传输波束成型与文件传输策略,最小化用户下载文件的时延。文献[14]将文献[9]的工作进一步拓展,在多播前程链路中,研究了最优缓存大小的策略。文献[15]采用遗传算法作为优化工具,提出了权衡内容缓存策略以减少小区平均中断概率,减轻前程链路负担。

在以上相关工作中,文献[8]与文献[9]最早发表在2016年,文献[11]发表在2017年,文献[10],文献[12-15]均是2018年的工作。本文工作延续了以上文献的研究思路,但希望解决已有工作的一些缺陷:(1)干扰管理技术过于复杂。大多数文献采用复杂的预编码波束成型技术解决干扰问题,为了有效降低过于复杂的干扰管理,并与现有的4G网络兼容,考虑正交频分多址接入(OFDMA)技术引入带缓存的Cloud-RAN;(2)未能考虑联合优化多种因素。以上文献大多侧重某一方面变量优化,未曾联合考虑“缓存状态”“用户最低需求”“RRHs选择”等多种基本要素。为了使优化问题更加实际,需要进行多种因素的联合优化;(3)RRHs的节能需求未能突出。在Cloud-RAN中,RRHs作为功能弱化的处理转发节点,其续航能力也受到限制。因此,RRHs的能耗需求也成为一个需要关注的问题,而以上的文献并没有特别关注这一方面。

因此,本文采用OFDMA的基于缓存的Cloud-RAN中下行传输过程中,在考虑用户需求与RRHs缓存状态的情形下,通过联合优化用户子载波分配、RRH选择、子载波传输功率分配,最小化RRHs的传输功率。由于问题是非凸的,因此提出拉格朗日对偶算法,有效地转化问题为凸问题,并给出了优化的资源分配方案。仿真表明,比起其他缓存方案,本文提出的优化算法可以有效地提升系统能效,性能有明显提升。由于篇幅限制,暂不考虑有关RRH调度的传输时延问题。

## 2 系统模型

如图1所示,考虑Cloud-RAN单个聚类中下行传输中,存在数量为 $M$ 单天线的RRHs,定义为 $\mathcal{M} = \{1, 2, \dots, M\}$ ;  $K$ 个单天线用户,定义为 $\mathcal{K} = \{1, 2, \dots, K\}$ ; 所有内容数目为 $T$ ,定义为 $\mathcal{T} = \{1, 2, \dots, T\}$ ,每个RRH的存储数量 $L$ 不能超过上界 $T$ ,即 $L \leq T$ 。当内容 $t$ 存储到第 $m$ -th RRH时,定义 $i_{m,t} = 1$ ,否则 $i_{m,t} = 0$ 。对于用户 $k$ 请求内容 $t$ ,也有类似的定义指示 $\theta_{k,t} = 1$ ,否则 $\theta_{k,t} = 0$ 。设定云中心已知所有RRHs的缓存内容集合 $\{i_{m,t}\}$ ,与所有用户的请求内容集合 $\{\theta_{k,t}\}$ 。假定每个用户在同一时刻只能请求一份内容,因此满足 $\sum_{t=1}^T \theta_{k,t} = 1, \forall k \in \mathcal{K}$ ,但同一份内容可以被多个用户请求, $t_k \in \mathcal{T}$ 为用户 $k$ 所请求的内容。

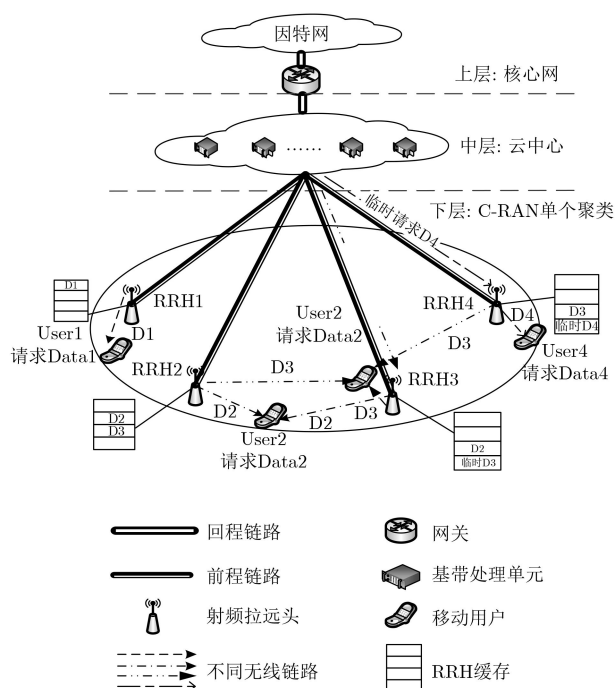


图1 带缓存Cloud-RAN下行传输结构图

在RRH中缓存的业务信息受欢迎的程度满足文献[15]中所提及的Zipf分布,并通过OFDMA方式,根据请求内容具体的大小,在单个或者多个间隔内传输。本文假定在单个间隔内完成资源分配优化。定义 $\varphi_{k,n}$ 指示子载波 $SC_n$ 是否分配给用户 $k$ ,当子载波 $SC_n$ 被分配给用户 $k$ 时, $\varphi_{k,n} = 1$ ,否则 $\varphi_{k,n} = 0$ 。同时定义 $\phi_n \triangleq [\varphi_{1,n}, \varphi_{2,n}, \dots, \varphi_{K,n}]^T$ 为在子载波 $SC_n$ 上的用户分配。根据OFDMA的定义,设定任何子载波 $SC_n$ 同时最多只能分配给单个用户,即满足 $\mathbf{1}^T \phi_n \leq 1, \forall n \in \mathcal{N}$ 。分配给用户 $k$ 的子载波集合定义为 $\mathcal{N}_k \in \mathcal{N}$ ,当 $\varphi_{k,n} = 1$ 时,

$\mathcal{N}_k = \{n\}$ , 并满足  $\mathcal{N}_{k1} \cap \mathcal{N}_{k2} = \emptyset, \forall k1 \neq k2, k1, k2 \in \mathcal{K}$ .

前程链路容量受限时, RRH一般仅会向中央云请求用户的未缓存内容, 并通过OFDMA下行传输给用户。因此定义RRH调用子载波传输的指示因子为  $\chi_{m,n}$ , 当RRH  $m$  调用子载波  $n$  进行传输时,  $\chi_{m,n} = 1$ , 否则  $\chi_{m,n} = 0$ 。定义通过子载波  $n$  传输的RRHs子集为  $\mathcal{X}_n = \{m \in \mathcal{M} | \chi_{m,n} = 1\}, n \in \mathcal{N}$ 。因此,  $\mathcal{X}_n$  中的RRHs可以在子载波  $n$  上协作传输内容给用户  $k$ , 即  $\varphi_{k,n} = 1$ 。

定义  $h_{k,m,n}$  为从RRH  $m$  到用户  $k$  在子载波  $n$  上的无线信道增益系数, 而云端已知所有信道状态信息。在  $\mathcal{X}_n$  中所有RRHs同时传输, 参考文献[8], 分配到子载波  $n$  的用户  $k$  上的接收信噪比(SNR)可以表示为

$$\begin{aligned} \gamma_{k,n}(\chi_n, \mathbf{p}_n) &= |\mathbf{h}_{k,n}^T \mathbf{x}_n|^2 / \sigma^2 \\ &= \left( \sum_{m=1}^M |h_{k,m,n}| \chi_{m,n} \sqrt{p_{m,n}} \right)^2 / \sigma^2 \end{aligned} \quad (1)$$

其中,  $\mathbf{x}_n$  代表SC  $n$  向用户  $k$  传输信号的矢量,  $\sigma^2$  是接收端加性高斯白噪声的功率, 所有用户端噪声相同。在SC  $n \in \mathcal{N}_k$  上的可达速率可表示为

$$R_{k,n}(\chi_n, \mathbf{p}_n) = B \log_2(1 + \gamma_{k,n}(\chi_n, \mathbf{p}_n)) / N \quad (2)$$

当  $\chi_n$  固定时,  $R_{k,n}(\chi_n, \mathbf{p}_n)$  对变量  $\{p_{m,n}\}$  是凹函数, 由于篇幅限制, 证明略。

图1列举了几种用户请求内容的典型模式, 具体如下所示:

(1) User1模式: User1请求的内容Data1(D1)仅在某个RRH的缓存中, 比如RRH1。此时不需要通过前程链路从BBU池调度数据, 由RRH1直接下行传输内容。

(2) User2模式: User2请求的内容Data2(D2)同时存储在多个RRH的缓存中, 比如RRH2, RRH3。此时不需要通过前程链路从BBU池调度数据, 由RRH2与RRH3联合协同下行传输内容。

(3) User3模式: User3请求的内容Data3(D3)只有部分存储在多个RRH的缓存中, 比如RRH2, RRH4。需要由离User3最近的RRH3通过前程链路从BBU池调度缺失数据, 再由RRH2, RRH3, RRH4联合协同下行传输内容。在多个RRHs协同发送数据给用户3时, 可以参照现有文献[16–18]设计一种联合的时频同步算法, 保证传输信息不出现乱序。本文默认传输同步已经完成。

(4) User4模式: User4请求的内容Data4(D4)没有存储在任何RRH的缓存中。需要由离User4最

近的RRH4通过前程链路从BBU池调度完整数据, 由RRH4直接下行传输内容。

在实际中, 由于多个用户可能都请求最受欢迎的内容, RRHs将受欢迎的内容按照比例放在缓存中, 并按照实际请求从BBU池调度一些独特的内容。无论如何, 由于前程链路容量是受限的。因此RRH  $m$  通过OFDMA方式, 向  $N$  SCs上所有用户传输非缓存数据的速率(即从BBU池调度下行的数据), 不应当超过前程链路容量  $F_m^{\text{tl}}$  (tl是total的简写), 即需满足限制条件

$$\sum_{t=1}^T (1 - i_{m,t}) \max_{k \in \mathcal{K}} \left\{ \theta_{k,t} \sum_{n=1}^N \chi_{m,n} \varphi_{k,n} \cdot R_{k,n}(\chi_n, \mathbf{p}_n) \right\} \leq F_m^{\text{tl}} \quad (3)$$

### 3 优化问题建立

基于第2节系统模型的定义, 本节将建立具体的优化模型。优化目标为最小化所有RRHs在SCs上的传输功率, 优化变量为用户-子载波分配  $\{\phi_n\}_{n \in \mathcal{N}}$ , RRHs选择  $\{\chi_n\}_{n \in \mathcal{N}}$ , 功率分配  $\{\mathbf{p}_n\}_{n \in \mathcal{N}}$ , 约束条件为前程链路容量限制与用户最小速率限制。优化问题(P1)为

$$\text{P1} \quad \min_{\{\mathbf{p}_n, \chi_n, \phi_n\}_{n \in \mathcal{N}}} P_{\text{RRHs}}^{\text{tl}} \quad (4)$$

s.t.

$$\sum_{t=1}^T (1 - i_{m,t}) \max_{k \in \mathcal{K}} \left\{ \theta_{k,t} \sum_{n=1}^N \chi_{m,n} \varphi_{k,n} \cdot R_{k,n}(\chi_n, \mathbf{p}_n) \right\} \leq F_m^{\text{tl}} \quad (5)$$

$$\sum_{n=1}^N \varphi_{k,n} \frac{B}{N} \log_2(1 + \gamma_{k,n}(\chi_n, \mathbf{p}_n)) \geq R_k^{\text{min}}, \forall k \in \mathcal{K} \quad (6)$$

$$\chi_{m,n} \in \{0, 1\}, \forall m \in \mathcal{M}, \forall n \in \mathcal{N} \quad (7)$$

$$\mathbf{1}^T \phi_n \leq 1, \varphi_{k,n} \in \{0, 1\}, \forall k \in \mathcal{K}, n \in \mathcal{N} \quad (8)$$

$$P_{\text{RRHs}}^{\text{tl}} = \sum_{m=1}^M \sum_{n=1}^N p_{m,n} + P_c, \forall m \in \mathcal{M}, n \in \mathcal{N} \quad (9)$$

$P_c$  为传输固定功率损耗。需要特别指出的是, 式(9)中的变量  $p_{m,n}$  的物理意义是RRH  $m$  在子载波  $n$  上所分配的功率, 是一个非负值, 而式(10)中的优化变量  $\mathbf{p}_n = [p_{1,n}, p_{2,n}, \dots, p_{M,n}]^T \in \mathbb{R}_+^{M \times 1}$  是一个矩阵。

式(5)是第2节所提及的前程链路容量约束, 具

体可见式(3);式(6)是满足用户需求的最小速率,即传统网络所提及的QoS(Quality of Service)限制;式(7)是RRH $m$ 是否在SC $n$ 传输的指示因子, $\chi_{m,n} = 1$ 代表RRH $m$ 选择在SC $n$ 上进行传输;式(8)是OFDMA设定的SC分配方式, $\varphi_{k,n} = 1$ 代表载波SC $n$ 被分配给用户 $k$ , $\mathbf{1}^T \boldsymbol{\phi}_n \leq 1$ 代表任何子载波SC $n$ 同时最多只能分配给一个用户。

为了简化约束条件式(5),考虑引入辅助变量

$$\xi_{m,t} = \max_{k \in \mathcal{K}} \left\{ \theta_{k,t} \sum_{n=1}^N \chi_{m,n} \varphi_{k,n} R_{k,n}(\boldsymbol{\chi}_n, \mathbf{p}_n) \right\} / F_m^{\text{tl}} \quad (10)$$

优化问题(P1)可以等效转化为问题(P2)

$$\text{P2} \quad \min_{\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}, \{\xi_{m,t}\}} P_{\text{RRHs}}^{\text{tl}} \quad (11)$$

s.t.

$$\frac{1}{F_m^{\text{tl}}} \theta_{k,t} (1 - i_{m,t}) \sum_{n=1}^N \chi_{m,n} \varphi_{k,n} R_{k,n}(\boldsymbol{\chi}_n, \mathbf{p}_n) \leq \theta_{k,t} (1 - i_{m,t}) \xi_{m,t} \quad (12)$$

$$\sum_{t=1}^T (1 - i_{m,t}) \xi_{m,t} \leq 1, \forall m \in \mathcal{M} \quad (13)$$

式(6),式(7),式(8),式(9)

当RRHs已经缓存用户请求的内容,即 $i_{m,t} = 1$ 时,约束条件式(12),式(13)无意义。由于式(2)中定义的 $R_{k,n}(\boldsymbol{\chi}_n, \mathbf{p}_n)$ 对于 $\{p_{m,n}\}$ 是凹函数,且(P2)中约束条件是非凸的,可以得知问题(P2)整体上是而非凸的。

#### 4 优化问题解决方案

文献[19]分析了非凸优化问题的对偶间隙,当目标与约束在SCs上可分离,SCs数目趋向无限时,对偶间隙趋向于0。在优化问题(P2)中,目标式(11)与约束式(12)在SCs上可分离,约束式(13)在SCs上虽然不可分离,但不影响整体问题(P2)的凹凸性。因此,当 $N$ 趋向于无穷时,问题(P2)的对偶界限趋向于0,因此可以利用拉格朗日对偶方法<sup>[20]</sup>解决问题(P2)。

设定约束式(12)的对偶因子为 $\alpha_{m,k,t} \geq 0$ , $m \in \mathcal{M}$ , $k \in \mathcal{K}$ , $t \in \mathcal{T}$ ;约束式(6)的对偶因子为 $\beta_k \geq 0$ , $k \in \mathcal{K}$ 。如果将 $\mathcal{M}_{t_k}^i \triangleq \{m | i_{m,t_k} = 0\}$ , $k \in \mathcal{K}$ 定义为缓存中没有被用户 $k$ 请求内容 $t_k$ 的RRHs合集,则在约束式(12)中存在 $I \triangleq \sum_{k=1}^K |\mathcal{M}_{t_k}^I|$ 。

定义向量 $\boldsymbol{\alpha} \in \mathbb{R}_+^{I \times 1}$ , $\boldsymbol{\beta} \in \mathbb{R}_+^{K \times 1}$ 为对偶变量的合集,仅有约束式(12)与约束式(6)的问题(P2)的拉格朗日变换为

$$\begin{aligned} & L(\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}, \{\xi_{m,t}\}, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ &= \sum_{n=1}^N L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta}) - \sum_{m=1}^M \sum_{k=1}^K \sum_{t=1}^T \alpha_{m,k,t} \\ & \quad \cdot \theta_{k,t} (1 - i_{m,t}) \xi_{m,t} + \sum_{k=1}^K \beta_k \end{aligned} \quad (14)$$

其中,子函数 $L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta})$ 定义为

$$\begin{aligned} & L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ & \triangleq \sum_{m=1}^M p_{m,n} + \sum_{m=1}^M \sum_{k=1}^K \sum_{t=1}^T \chi_{m,n} \varphi_{k,n} \theta_{k,t} (1 - i_{m,t}) \frac{\alpha_{m,k,t}}{F_m^{\text{tl}}} \\ & \quad \cdot R_{k,n}(\boldsymbol{\chi}_n, \mathbf{p}_n) - \sum_{k=1}^K \frac{\beta_k}{R_k^{\text{min}}} \varphi_{k,n} R_{k,n}(\boldsymbol{\chi}_n, \mathbf{p}_n) \end{aligned} \quad (15)$$

对应的拉格朗日对偶函数可以表示为

$$\begin{aligned} g(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \min_{\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}, \{\xi_{m,t}\}, \boldsymbol{\alpha}, \boldsymbol{\beta}} L(\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}, \\ & \quad \{\xi_{m,t}\}, \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \text{s.t.} & \text{式(13),式(7) - 式(9)} \end{aligned} \quad (16)$$

参照约束式(12),约束式(6)的对偶函数可以表示为

$$g(\boldsymbol{\alpha}, \boldsymbol{\beta}) = g_1(\boldsymbol{\alpha}, \boldsymbol{\beta}) + g_2(\boldsymbol{\alpha}, \boldsymbol{\beta}) + \sum_{k=1}^K \beta_k \quad (17)$$

$$\begin{aligned} g_1(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \min_{\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}} \sum_{n=1}^N L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta}), \\ \text{s.t.} & \text{式(7) - 式(9)} \end{aligned} \quad (18)$$

$$\begin{aligned} g_2(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \min_{\{\xi_{m,t}\}} - \sum_{m=1}^M \sum_{t=1}^T \sum_{k=1}^K \theta_{k,t} (1 - i_{m,t}) \alpha_{m,k,t} \xi_{m,t}, \\ \text{s.t.} & \text{式(13)} \end{aligned} \quad (19)$$

式(22)中的最小化问题可以分解成 $N$ 个平行的子问题,每个子问题对应一个单独的SC $n \in \mathcal{N}$ 。所有子问题可以用相似的优化问题表示。

$$\left. \begin{aligned} & \min_{\{\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n\}_{n \in \mathcal{N}}} \sum_{n=1}^N L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta}) \\ \text{s.t.} & \mathbf{p}_n \geq \mathbf{0}, \boldsymbol{\chi}_n \in \{0, 1\}^{M \times 1}, \\ & \mathbf{1}^T \boldsymbol{\phi}_n \leq 1, \boldsymbol{\phi}_n \in \{0, 1\}^{K \times 1} \end{aligned} \right\} \quad (20)$$

式(20)中, $L_n(\mathbf{p}_n, \boldsymbol{\chi}_n, \boldsymbol{\phi}_n, \boldsymbol{\alpha}, \boldsymbol{\beta})$ 与式(15)中的定义一致。下面的内容将研究如何解决优化问题式(20)。

定义SC $n$ 上用户关联策略固定 $\varphi_n = \tilde{\varphi}_n$ ,当 $\tilde{\varphi}_n = 0$ 时,SC $n$ 不被任何用户占用。由于 $\mathbf{p}_n \geq \mathbf{0}$ ,式(20)中的目标函数仅在 $\mathbf{p}_n = \mathbf{0}$ 时实现最小值,与



RRHs选择变量 $\chi_n$ 无关。因此，如果没有用户被分配SC $n$ ，所有RRHs的功率分配 $\mathbf{p}_n = \mathbf{0}$ 时，可以不失一般性地假设 $\chi_n = \mathbf{0}$ 。定义 $\tilde{k}_n \in \mathcal{K}$ 是分配SC $n$ 的用户，因此 $\tilde{\varphi}_{\tilde{k}_n, n} = 1, \tilde{\varphi}_{k, n} = 0$ ；定义 $\tilde{t}_{k_n} \in \mathcal{T}$ 是用户 $\tilde{k}_n$ 请求内容，满足 $\theta_{\tilde{k}_n, \tilde{t}_{k_n}} = 1, \theta_{\tilde{k}_n, t} = 0, \forall t \neq \tilde{t}_{k_n}$ 。式(20)可以转化为

$$\begin{aligned} \min_{\mathbf{p}_n, \chi_n} & - \left( \beta_{\tilde{k}_n} / R_{\tilde{k}_n}^{\min} - \sum_{m=1}^M (1 - i_{m, \tilde{k}_n}) \chi_{m, n} \alpha_{m, \tilde{k}_n, \tilde{t}_{k_n}} / F_m^{\text{tl}} \right) \\ & \cdot R_{\tilde{k}_n, n}(\chi_n, \mathbf{p}_n) + \sum_{m=1}^M p_{m, n}, \\ \text{s.t. } & \mathbf{p}_n \succeq \mathbf{0}, \chi_n \in \{0, 1\}^{M \times 1} \end{aligned} \quad (21)$$

对于给定的RRH选择方案 $\tilde{\chi}_n$ ，问题式(21)的最优功率分配方案 $\hat{\mathbf{p}}_n$ 由以下推论给出。

**推论1** 假定 $\chi_n = \tilde{\chi}_n$ ，问题式(21)的最优功率分配方案 $\hat{\mathbf{p}}_n$ 如式(26)

$$\hat{p}_{m, n} = \Pi(\tilde{\chi}_n) |h_{\tilde{k}_n, m, n}^-|^2 \tilde{\chi}_{m, n} / \left[ (\Theta_{\tilde{k}_n, n}(\tilde{\chi}_n))^2 \sigma^2 \right] \quad (22)$$

其中，

$$\Pi(\tilde{\chi}_n) = \left[ \frac{B \Omega_{\tilde{k}_n, n}(\tilde{\chi}_n) \Theta_{\tilde{k}_n, n}(\tilde{\chi}_n)}{N \ln 2} - 1 \right]^+ \quad (22a)$$

$$\Omega_{\tilde{k}_n, n}(\tilde{\chi}_n) = \frac{\beta_{\tilde{k}_n}}{R_{\tilde{k}_n}^{\min}} - \sum_{m=1}^M \frac{(1 - i_{m, \tilde{t}_{k_n}}) \chi_{m, n} \alpha_{m, \tilde{k}_n, \tilde{t}_{k_n}}}{F_m^{\text{tl}}} \quad (22b)$$

$$\Theta_{\tilde{k}_n, n}(\tilde{\chi}_n) = \sum_{m=1}^M \frac{\chi_{m, n} |h_{\tilde{k}_n, m, n}^-|^2}{\sigma^2} \quad (22c)$$

推论1说明，对于给定用户关联策略 $\tilde{k}_n$ 与RRH选择策略 $\tilde{\chi}_n$ ，最优功率分配策略是阈值结构。当 $\Omega_{\tilde{k}_n, n}(\tilde{\chi}_n) \Theta_{\tilde{k}_n, n}(\tilde{\chi}_n) \leq (N \ln 2) / B$ ，所有RRHs在SC $n$ 上分配零功率；当 $\Omega_{\tilde{k}_n, n}(\tilde{\chi}_n) \Theta_{\tilde{k}_n, n}(\tilde{\chi}_n) > (N \ln 2) / B$ ，RRH $m$ 上分配的功率取决于无线信道增益 $|h_{\tilde{k}_n, m, n}^-|$ ， $\beta_{\tilde{k}_n}$ 等变量。

在式(22)中，如果所有RRHs均缓存用户请求的内容 $\tilde{t}_{k_n}$ ，即 $i_{m, \tilde{t}_{k_n}} = 1$ ，功率分配表达式可简化为

$$\begin{aligned} \hat{p}_{m, n} & = \tilde{\chi}_{m, n} |h_{\tilde{k}_n, m, n}^-|^2 \left[ B \beta_{\tilde{k}_n} \Theta_{\tilde{k}_n, n}(\tilde{\chi}_n) / R_{\tilde{k}_n}^{\min} N \ln 2 - 1 \right]^+ \\ & / \left[ (\Theta_{\tilde{k}_n, n}(\tilde{\chi}_n))^2 \sigma^2 \right] \end{aligned} \quad (23)$$

如果 $\tilde{\chi}_n = \mathbf{0}$ ，即没有RRH选择， $\hat{\mathbf{p}}_n = \mathbf{0}$ 。对于聚类仅有一个RRH的特例，功率分配表达式可简化为

$$\begin{aligned} \hat{p}_n & = \left[ B \left( \beta_{\tilde{k}_n} / R_{\tilde{k}_n}^{\min} - (1 - i_{\tilde{t}_{k_n}}) \alpha_{\tilde{k}_n, \tilde{t}_{k_n}} / F^{\text{tl}} \right) \right. \\ & \cdot (N \ln 2) - \sigma^2 / |h_{\tilde{k}_n, n}^-|^2 \left. \right]^+ \end{aligned} \quad (24)$$

式(24)与已知的注水算法形式相似，主要区别在于在不同SC $n \in \mathcal{N}$ 上的水位不同。如果 $\left( \beta_{\tilde{k}_n} / R_{\tilde{k}_n}^{\min} \right) \leq (1 - i_{\tilde{t}_{k_n}}) \alpha_{\tilde{k}_n, \tilde{t}_{k_n}} / F^{\text{tl}}$ ，SC $n$ 上不分配任何功率。

对于任何给定对偶变量 $\alpha, \beta$ ，式(21)可以按以下步骤，通过推论1得到最优解

步骤1 对 $\tilde{k}_n \in \mathcal{K}$ ，固定SC $n$ 上的用户；

步骤2 对于 $2^M$ 种RRH选择可能，利用式(22)计算最优功率分配策略 $\hat{\mathbf{p}}_n$ ，并将式(22)代入式(21)，通过使式(21)目标函数最小化，得到最优RRH选择策略 $\tilde{\chi}_n$ ；

步骤3 基于最优功率分配策略 $\hat{\mathbf{p}}_n$ ，最优RRH选择策略 $\tilde{\chi}_n$ ，通过使式(20)目标函数最小化，得到最优用户SC分配策略 $\hat{\phi}_n$ 。

**推论2** 问题式(19)的最优解如式(25)所示

$$\hat{\xi}_{m, t} = \begin{cases} 1, & t = \arg \max_{l \in \mathcal{T}} \left\{ \sum_{k=1}^K \theta_{k, t} \right. \\ & \left. \cdot (1 - i_{m, l}) \alpha_{m, k, l} \right\} \\ 0, & \text{其它} \end{cases} \quad (25)$$

推论2说明式(19)可以通过搜寻 $\sum_{k=1}^K \theta_{k, t} (1 - i_{m, l}) \alpha_{m, k, l}$ 的最大值内容，并将其与用户请求的内容比较，如果两者一致，则 $\hat{\xi}_{m, t} = 1$ 。通过分析，可知寻找式(19)最优解的最大复杂度为 $\mathcal{O}(MT)$ ，发生在至少 $T$ 个用户请求不同内容，且所有内容都没有在任何RRHs端缓存。

此时，问题(P1)的对偶问题可表示为

$$\max_{\alpha \succeq \mathbf{0}, \beta \succeq \mathbf{0}} g(\alpha, \beta) \quad (26)$$

式(30)是一个凸问题，可以通过经典凸优化方法，如椭球法得到最优对偶变量 $\hat{\alpha}, \hat{\beta}$ 。基于推论1与推论2得到的最优解 $\hat{\mathbf{p}}_n, \hat{\chi}_n, \hat{\phi}_n, \hat{\xi}_{m, t}$ ，与最优对偶变量 $\hat{\alpha}, \hat{\beta}$ ，本文提出解决式(4)中问题(P1)基于功耗最小的联合优化算法步骤如下：

步骤1 相关参数初始化, 设置 $\alpha, \beta$ , 最大迭代次数 $I_{\max}$ , 收敛条件 $\varepsilon$ 等;

步骤2 给定分配SC $n$ 的用户 $k_n$ , RRHs选择SCs策略 $\chi_n$ , 基于推论1得出最优功率分配 $\hat{p}_n$ ;

步骤3 将 $\hat{p}_n$ 代入式(21), 通过使式(21)目标函数最小化, 得到最优RRH选择SCs策略 $\hat{\chi}_n$ ;

步骤4 将 $\hat{p}_n, \hat{\chi}_n$ , 通过使式(20)目标函数最小化, 得到最优用户SC分配策略 $\hat{\phi}_n$ ;

步骤5 基于推论2, 解决式(19)中的优化问题, 得到最优解 $\hat{\xi}_{m,t}$ ;

步骤6 利用椭球法, 更新对偶变量 $\hat{\alpha}, \hat{\beta}$ ;

步骤7 重复步骤2—步骤6, 直至算法收敛。

通过分析, 可知解决式(4)中优化问题(P1)的计算复杂度是 $\mathcal{O}(NK2^M)$ ; 式(20)的计算复杂度是 $\mathcal{O}(K2^M)$ ; 式(16)的计算复杂度最小值 $\mathcal{O}(M)$ , 最大值 $\mathcal{O}(MT)$ ; 算法1的计算复杂度最小值为 $\mathcal{O}(2NK + 1)$ , 最大值 $\mathcal{O}(NK2^M + MT)$ 。

### 5 仿真结果与分析

仿真中, 设定一个小区聚类中无线带宽为 $B=20$  MHz, 均分为32个SCs。假设用户所经历的是独立同分布6路径瑞利衰落信道, 背景噪声的功率谱密度设为 $-169$  dBm/Hz。RRHs个数 $M=4$ , 在半径为100 m范围内均匀分布, 用户个数 $K=8$ , 在半径为100 m范围内随机分布。满足用户最低传输需求 $R_k^{\min}=20$  Mbps, 用户请求所有内容数目 $T=40$ , 满足Zipf分布<sup>[9]</sup>, 用户请求内容 $t$ 的概率为 $\rho_t = t^{-\lambda} / \sum_{t=1}^T t^{-\lambda}, t \in \mathcal{T}, \lambda$ 为分布参数, 设定 $\lambda=0.9$ 。每个RRH前程容量约束 $F_m^{\text{ul}}=60$  Mbps, 最大缓存内容 $S=4$ 。

在仿真中, 除了本文提出的算法, 还涉及以下对比方案:

方案1 最大概率缓存。每个RRH根据内容的流行概率进行缓存, 直到缓存空间全部占用。

方案2 无缓存。RRHs没有缓存功能, 用户请求任何数据, 均需要通过前程链路向BBU池请求。

方案3 平均功率分配。RRHs端在所有子载波上平均分配功率, 不进行任何优化。

方案4 单RRH选择。聚类中仅有1个RRH, 不存在RRH选择的问题。

图2给出了随着RRH总数变大, 不同参数下本文算法的最大复杂度变化趋势。当SCs总数 $N$ 与需求内容总数 $T$ 一致时, 用户数目 $K$ 越大, 算法最大复杂度越大。同样的趋势, 对参数SCs总数 $N$ 与需求内容总数 $T$ 也同时成立。但必须指出, RRHs个

数 $M$ 对复杂度影响最大, 因为运算中 $M$ 位于指数位置, 其它变量均为线性乘积。

图3给出了随着用户总数变大, 不同参数下本文算法的最大复杂度变化趋势。当用户数目 $K$ 线性增长, 本文算法最大复杂度也线性增长。在图2中, 当RRHs个数 $M$ 线性增长, 本文算法最大复杂度呈指数增长。这说明不同参数对复杂度影响不同, RRHs个数 $M$ 占主导作用。

图4给出了随着前程容量门限变化时, 多种方案功耗对比。比起方案1—方案4, 本文算法可以使单个RRH平均功率损耗最低, 即保持最优性能。本文算法性能优于基于最大概率缓存的方案1, 因为在大多数场景存在多个RRHs各有用户请求一部分内容, 并联合发送给用户的状况, 方案1更适应RRHs独立为用户缓存内容的场景。当前程容量门

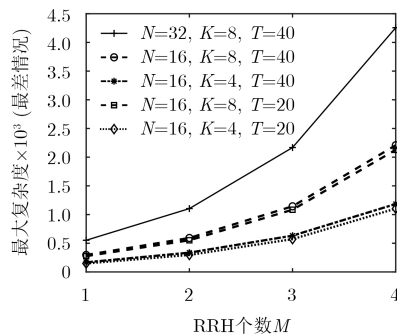


图2 RRH个数变化, 本文算法最大复杂度变化趋势

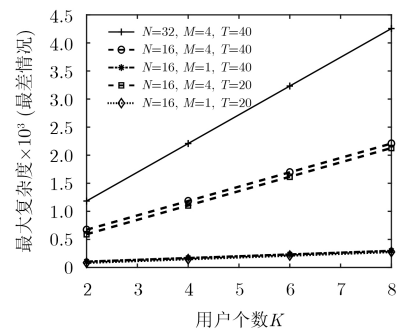


图3 用户个数变化, 本文算法最大复杂度变化趋势

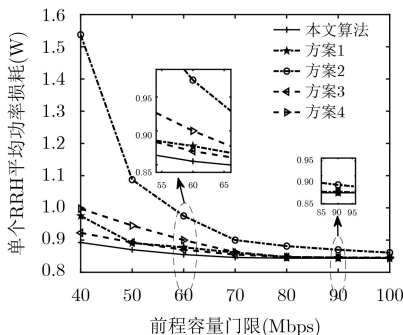


图4 前程容量门限变化, 多种方案功耗对比

限增大时, 本文算法的优势逐渐变小。  $F_m^{\text{tl}}=90$  Mbps 时, 除了不做任何缓存的方案2, 本文算法与方案1, 3, 4性能几乎一致。因为随着前程容量门限增大, 说明前程链路的负担已经被大大减轻, 用户可以直接从BBU池请求发送下行数据而不必顾虑前程约束, RRHs缓存作用则受到削弱。这说明, 缓存优化在前程链路严重受限场景更有必要。

图5给出了随着单个RRH最大缓存数目变化时, 多种方案功耗对比。与图4结论类似, 本文算法可以一直实现单个RRH最低功率损耗, 但算法优势随着单个RRH最大缓存数目变大而变小。因为当单个RRH最大缓存数目变大时, 用户更倾向于从更少数目的RRHs请求内容, 避免了向所有RRHs或者BBU池发送内容请求, 从整体上节约能耗。这说明, RRHs增加缓存功能具有必要性, 在节约能耗上有着非常明显的作用。

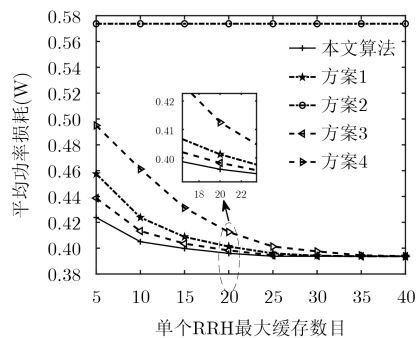


图5 单个RRH最大缓存数目变化, 多种方案功耗对比

## 6 结束语

本文针对带缓存的Cloud-RAN, 以最小化RRHs能耗为目标, 对SCs分配、RRHs选择与RRHs下行功率等变量进行了优化设计。通过拉格朗日对偶分解, 原始的非凸问题转化为几个可以解决的独立子问题, 基于子问题的最优解, 本文提出了联合设计优化算法。仿真结果验证了所提算法可以有效地降低RRHs功耗, 提升系统能效, 符合未来绿色通信的需求。本文并没有涉及RRHs端前程容量与缓存大小动态分配的问题, 在未来工作中需要进一步研究。

## 参考文献

- [1] CHECKO A, CHRISTIANSEN H L, YAN Yan, *et al.* Cloud RAN for mobile networks—A technology overview[J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(1): 405–426. doi: 10.1109/COMST.2014.2355255.
- [2] WANG Dongming, WANG Jiangzhou, YOU Xiaohu, *et al.* Spectral efficiency of distributed MIMO systems[J]. *IEEE Journal on Selected Areas in Communications*, 2013, 31(10): 2112–2127. doi: 10.1109/JSAC.2013.131012.
- [3] WANG Yi, LI Chunguo, HANG Yongming, *et al.* Energy-efficient optimization for downlink massive MIMO FDD systems with transmit-side channel correlation[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(9): 7228–7243. doi: 10.1109/TVT.2015.2483519.
- [4] LIU Liang, BI Suzhi, and ZHANG Rui. Joint power control and fronthaul rate allocation for throughput maximization in OFDMA-based cloud radio access network[J]. *IEEE Transactions on Communications*, 2015, 63(11): 4097–4110. doi: 10.1109/TCOMM.2015.2469687.
- [5] ZHOU Yuhua and YU Wei. Optimized backhaul compression for uplink cloud radio access network[J]. *IEEE Journal on Selected Areas in Communications*, 2014, 32(6): 1295–1307. doi: 10.1109/JSAC.2014.2328133.
- [6] LI Chunguo, SONG Kang, WANG Dongming, *et al.* Optimal remote radio head selection for cloud radio access networks[J]. *Science China Information Sciences*, 2016, 59(10): 102315. doi: 10.1007/s11432-016-0060-y.
- [7] LUO Shixin, ZHANG Rui, and LIM T J. Downlink and uplink energy minimization through user association and beamforming in C-RAN[J]. *IEEE Transactions on Wireless Communications*, 2015, 14(1): 494–508. doi: 10.1109/TWC.2014.2352619.
- [8] ZHAO Zhongyuan, PENG Mugen, DING Zhiguo, *et al.* Cluster content caching: An energy-efficient approach to improve quality of service in cloud radio access networks[J]. *IEEE Journal on Selected Areas in Communications*, 2016, 34(5): 1207–1221. doi: 10.1109/JSAC.2016.2545384.
- [9] TAO Meixia, CHEN Erkai, ZHOU Hao, *et al.* Content-centric sparse multicast beamforming for cache-enabled Cloud RAN[J]. *IEEE Transactions on Wireless Communications*, 2016, 15(9): 6118–6131. doi: 10.1109/TWC.2016.2578922.
- [10] TRAN T X, LE D V, YUE Guosen, *et al.* Cooperative hierarchical caching and request scheduling in a cloud radio access network[J]. *IEEE Transactions on Mobile Computing*, 2018, 17(12): 2729–2743. doi: 10.1109/TMC.2018.2818723.
- [11] CHEN Mingzhe, SAAD W, YIN Changchuan, *et al.* Echo state networks for proactive caching in cloud-based radio access networks with mobile users[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(6): 3520–3535. doi: 10.1109/TWC.2017.2683482.
- [12] PU Lingjun, JIAO Lei, CHEN Xu, *et al.* Online resource allocation, content placement and request routing for cost-efficient edge caching in cloud radio access networks[J]. *IEEE Journal on Selected Areas in Communications*, 2018, 36(8): 1751–1767. doi: 10.1109/JSAC.2018.2844624.
- [13] YANG Xiaolong, ZHENG Jianchao, FEI Zesong, *et al.*

- Optimal file dissemination and beamforming for cache-enabled C-RANs[J]. *IEEE Access*, 2018, 6: 6390–6399. doi: [10.1109/ACCESS.2017.2775198](https://doi.org/10.1109/ACCESS.2017.2775198).
- [14] DAI Binbin, LIU Yafeng, and YU Wei. Optimized base-station cache allocation for cloud radio access network with multicast backhaul[J]. *IEEE Journal on Selected Areas in Communications*, 2018, 36(8): 1737–1750. doi: [10.1109/JSAC.2018.2844638](https://doi.org/10.1109/JSAC.2018.2844638).
- [15] YE Zhun, PAN Cunhua, ZHU Huiling, *et al.* Tradeoff caching strategy of the outage probability and fronthaul usage in a cloud-RAN[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(7): 6383–6397. doi: [10.1109/TVT.2018.2797957](https://doi.org/10.1109/TVT.2018.2797957).
- [16] GOLNARI A, SHABANY M, NEZAMALHOSSEINI A, *et al.* Design and implementation of time and frequency synchronization in LTE[J]. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 2015, 23(12): 2970–2982. doi: [10.1109/TVLSI.2014.2387861](https://doi.org/10.1109/TVLSI.2014.2387861).
- [17] ABDZADEH-ZIABARI H and SHAYESTEH M G. Robust timing and frequency synchronization for OFDM systems[J]. *IEEE Transactions on Vehicular Technology*, 2011, 60(8): 3646–3656. doi: [10.1109/TVT.2011.2163194](https://doi.org/10.1109/TVT.2011.2163194).
- [18] KRÄMER J and SEEGER B. Semantics and implementation of continuous sliding window queries over data streams[J]. *ACM Transactions on Database Systems*, 2009, 34(1): Article No.4. doi: [10.1145/1508857.1508861](https://doi.org/10.1145/1508857.1508861).
- [19] YU Wei and LUI R. Dual methods for nonconvex spectrum optimization of multicarrier systems[J]. *IEEE Transactions on Communications*, 2006, 54(7): 1310–1322. doi: [10.1109/TCOMM.2006.877962](https://doi.org/10.1109/TCOMM.2006.877962).
- [20] BOYD S and VANDENBERGHE L. *Convex Optimization*[M]. Cambridge: Cambridge University Press, 2004: 32–40.
- 孙 远: 男, 1988年生, 博士生, 研究方向为云接入网络资源分配.
- 李春国: 男, 1983年生, 教授, 博士生导师, 研究方向为新一代无线通信理论与技术, 基于人工智能的视频信息处理, 水下无线通信.
- 黄永明: 男, 1977年生, 教授, 博士生导师, 研究方向为下一代移动通信技术, 毫米波MIMO.
- 杨绿溪: 男, 1964年生, 教授, 博士生导师, 研究方向为下一代移动通信技术, 信号与信息处理.