

基于非正交多址接入的网络切片联合用户关联和功率分配算法

唐 伦 马润琳* 杨 恒 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: 为了满足网络切片多样化需求, 实现无线虚拟资源的动态分配, 该文提出在C-RAN架构中基于非正交多址接入的联合用户关联和功率资源分配算法。首先, 该算法考虑在不完美信道条件下, 以切片和用户最小速率需求及时延QoS要求、系统中断概率、前传容量为约束, 建立在C-RAN场景中最大化长时平均网络切片总吞吐量的联合用户关联和功率分配模型。其次, 将概率混合优化问题转换为非概率优化问题, 并利用Lyapunov优化理论设计一种基于当前时隙的联合用户调度和功率分配的算法。最后采用贪婪算法求得用户关联问题次优解; 基于用户关联的策略, 将功率分配的问题利用连续凸逼近方法将其转换为凸优化问题并采用拉格朗日对偶分解方法获得功率分配策略。仿真结果表明, 该算法能满足各网络切片和用户需求的同时有效提升系统时间平均切片总吞吐量。
关键词: 网络切片; 资源分配; 不完美信道状态信息; 非正交多址接入

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2019)09-2039-08

DOI: 10.11999/JEIT180770

Joint User Association and Power Allocation Algorithm for Network Slicing Based on NOMA

TANG Lun MA Runlin YANG Heng 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 satisfy the diversity of requirements for different network slices and realize dynamic allocation of wireless virtual resource, an algorithm for network slice joint user association and power allocation is proposed in Non-Orthogonal Multiple Access(NOMA) C-RAN. Firstly, by considering imperfect Channel State Information(CSI), a joint user association and power allocation algorithm is designed to maximize the average total throughput in C-RAN with the constraints of slice and user minimum required rate, outage probability and fronthaul capacity limits. Secondly, a joint user association and power allocation algorithm is designed according to the current slot by transforming the probabilistic mixed optimization problem into a non-probabilistic optimization problem and using Lyapunov optimization. Finally, for user association problem, a greedy algorithm is proposed to find a feasible suboptimal solution; The power allocation problem is transformed into a convex optimization problem by using successive convex approximation; Then a dual decomposition approach is exploited to obtain a power allocation strategy. Simulation results demonstrate that the proposed algorithm can effectively improve the average total throughput of system while guaranteeing the network slice and user requirement.

Key words: Network slicing; Resource allocation; Imperfect Channel State Information (CSI); Non-Orthogonal Multiple Access (NOMA)

收稿日期: 2018-08-23; 改回日期: 2019-03-20; 网络出版: 2019-05-30

*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)

1 引言

随着通信行业的飞速发展,通信业务规模不断扩大,用户需求的差异性越来越大,有限的频谱资源也变得越来越紧缺。因此,利用网络虚拟化技术而实现资源共享的网络切片技术应运而生。作为5G的一项关键技术,网络切片的实质是具有灵活性且相互独立的虚拟逻辑网络,其作用是为不同应用场景中具有差异化QoS需求的用户提供特定的服务。

网络切片资源分配已成为当下的研究热点。文献[1]为了使网络切片中基础设施供应商、服务提供商和用户都获得较高的收益,提出了一种基于乘法器的分布式资源分配算法,未考虑各切片业务的不同时延需求。文献[2]考虑网络切片中业务到达和无线信道的动态变化提出了基于OFDMA的无线虚拟网络中吞吐量最大化的资源分配方案,文中只考虑了时频资源,未充分利用频谱资源。文献[3]考虑每个网络切片的统计时延QoS需求,在下行传输系统中应用非正交多址接入(Non-Orthogonal Multiple Access, NOMA)技术提高了系统的有效容量,文献[4]在NOMA系统中考虑每个网络切片的最小预留速率来保证各网络切片间的隔离性,解决了在无线虚拟网络中最小化传输功率的资源分配问题,文献[5,6]在C-RAN架构下研究了网络切片资源分配机制,未考虑长期优化目标。

基于以上研究,本文创新点如下:(1)在C-RAN架构中结合NOMA开发功率域资源,建立优化系统长时平均吞吐量的联合用户关联和功率分配模型。本文考虑不完美CSI下系统长期平均吞吐量问题,将其建模为概率混合的非凸优化问题;(2)由于原问题为概率混合和时间平均的随机非凸优化问题,为了求解该优化问题,先将原问题转换为非概

率混合问题,然后利用Lyapunov优化理论将时间平均优化问题转换为单时隙上的优化问题;(3)提出了一种全局迭代算法,该算法包括用户关联和功率分配两个子问题。用户关联问题为0-1多重背包问题,为了降低复杂度本文采用贪婪算法求得次优解;基于用户关联的策略,将功率分配的问题利用连续凸逼近方法将其转换为凸优化问题,并采用拉格朗日对偶分解方法获得功率分配的策略。

2 系统模型和问题建模

2.1 基于NOMA的C-RAN架构模型

本文考虑如图1所示的C-RAN场景下基于NOMA下行链路网络切片资源分配模型。在该网络中共有 K 个切片,集合为 $\mathcal{K} = \{k | 1 \leq k \leq K\}$,支持 K 种不同类型的业务。用户 k_i 表示切片 k 中的第 i 个用户,切片 k 中的用户集合表示为 $U_k = \{k_i | 1 \leq k_i \leq N_k\}$ 。所有用户的集合表示为 U , $U = U_1 \cup U_2 \cup \dots \cup U_K$ 。

切片用户被分配至不同资源块(Resource Block, RB),采用NOMA技术每个RB可以复用多个用户,用户之间使用不同的功率,共有 B 个RB,集合表示为 $\mathcal{B} = \{b | 1 \leq b \leq B\}$ 。C-RAN中共有 L 个射频拉远头(Remote Radio Head, RRH),集合表示为 $\mathcal{L} = \{l | 1 \leq l \leq L\}$,室内基带处理单元(Building Base band Unite, BBU)池将RB分配给各个RRH,每个RB带宽为 B_s 。BBU池通过fronthaul链路向RRH提供基带信号,设RRH l 的fronthaul链路容量为 $C_{l,\max}$ 。令 $a_{k_i,l}(t)$, $b_{k_i,l}^b(t)$ 分别表示一个二进制变量,用户 k_i 连接在RRH l 上时 $a_{k_i,l}(t) = 1$,反之为0。RB b 分配给了与RRH l 关联的用户 k_i 时 $b_{k_i,l}^b(t) = 1$,否则为0。

设RRH l 的最大发射功率为 $p_l^{\max}$, $p_{k_i,l}^b(t)$ 为RB b 分配给与RRH l 关联用户 k_i 的功率,则

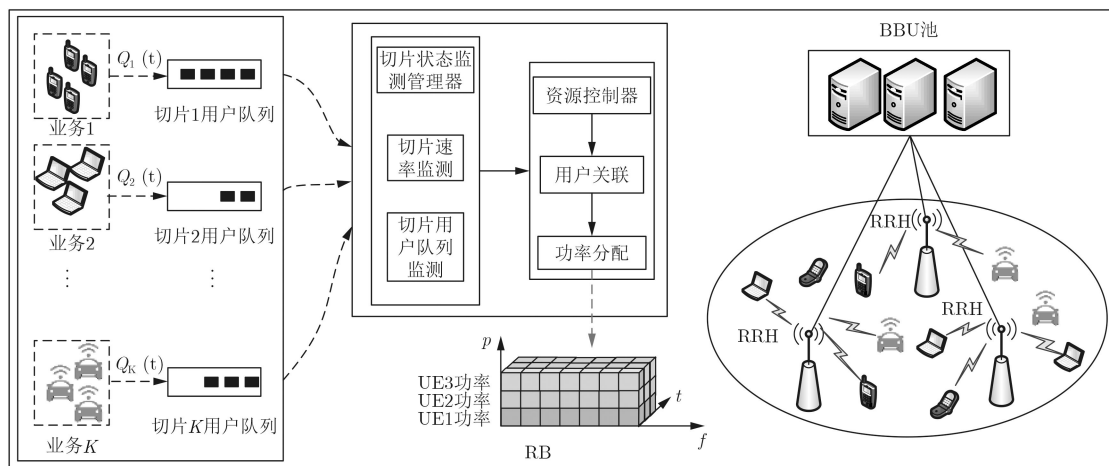


图1 基于NOMA的网络切片架构

$$\sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) p_{k_i,l}^b(t) \leq p_l^{\max}, \forall l \in \mathcal{L} \quad (1)$$

将复用在RB b 上的所有用户的信道增益按从小到大排列, 则复用在RB b 上且信道增益大于用户 k_i 的所有用户集合表示为 $M_{k_i,l}^b(t) = \{n_j | h_{n_j,l}^b(t) > h_{k_i,l}^b(t)\}$ 。在接收端采用串行干扰消除技术用户 k_i 接收到的信号为

$$y_{k_i}(t) = h_{k_i,l}^b(t) S_b(t) + z_{k_i}(t) = h_{k_i,l}^b(t) \sqrt{p_{k_i,l}^b(t)} x_{k_i}(t) + h_{k_i,l}^b(t) \sum_{n_j \in M_{k_i,l}^b(t)} \sqrt{p_{n_j,l}^b(t)} x_{n_j}(t) + z_{k_i}(t) \quad (2)$$

其中, $h_{k_i,l}^b(t) = L_{k_i,l}^b(t) g_{k_i,l}^b(t)$, $L_{k_i,l}^b(t) = \sqrt{d_{k_i,l}^b(t)}$ 为路径衰落系数, $d_{k_i,l}^b(t)$ 为用户 k_i 到RRH l 的距离; $g_{k_i,l}^b(t)$ 为瑞利衰落系数。 $z_{k_i}(t)$ 为均值为0, 方差为 σ_z^2 的复加性高斯白噪声。

2.1.1 信道模型

假设路径衰落系数 $L_{k_i,l}^b(t)$ 可以很好地被RRH估计[7]。通过使用最小均方误差信道估计模型[8], 本文将复用在RB b 的用户 k_i 与RRH l 之间的瑞利衰落系数建模为

$$g_{k_i,l}^b(t) = \bar{g}_{k_i,l}^b(t) + e_{k_i,l}^b(t) \quad (3)$$

其中, $\bar{g}_{k_i,l}^b(t)$ 为估计的瑞利衰落系数, $e_{k_i,l}^b(t)$ 为估计误差, 服从均值为0, 方差为 σ_e^2 的复高斯分布。

如果RRH能获得完美信道, 根据香农公式, 与RRH l 关联且复用在RB b 上的用户 k_i 的可达速率为

$$c_{k_i,l}^b(t) = B_s \log_2(1 + \Phi_{k_i,l}^b(t)) \quad (4)$$

其中, B_s 为RB的带宽, 信干噪比 $\Phi_{k_i,l}^b(t)$ 为

$$\Phi_{k_i,l}^b(t) = \frac{p_{k_i,l}^b(t) |h_{k_i,l}^b(t)|^2}{|h_{k_i,l}^b(t)|^2 \sum_{n_j \in M_{k_i,l}^b(t)} p_{n_j,l}^b(t) + \sigma_z^2} \quad (5)$$

其中, $|h_{k_i,l}^b(t)|^2 \sum_{n_j \in M_{k_i,l}^b(t)} p_{n_j,l}^b(t)$ 为复用在RB b 上且比用户 k_i 具有更高信道增益的用户产生的干扰。

实际情况下RRH只知道估计信道增益 $\bar{h}_{k_i,l}^b(t)$, 用户的数据速率为

$$r_{k_i,l}^b(t) = B_s \log_2(1 + \gamma_{k_i,l}^b(t)) \quad (6)$$

其中, $\gamma_{k_i,l}^b(t)$ 为信道增益为 $\bar{h}_{k_i,l}^b(t) = L_{k_i,l}^b(t) \bar{g}_{k_i,l}^b(t)$ 得到的信干噪比, 其表达式为

$$\gamma_{k_i,l}^b(t) = \frac{p_{k_i,l}^b(t) |\bar{h}_{k_i,l}^b(t)|^2}{|\bar{h}_{k_i,l}^b(t)|^2 \sum_{n_j \in M_{k_i,l}^b(t)} p_{n_j,l}^b(t) + \sigma_z^2} \quad (7)$$

因此系统总速率为

$$R(t) = \sum_{k \in \mathcal{K}} R_k(t) = \sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) \cdot b_{k_i,l}^b(t) r_{k_i,l}^b(t) \quad (8)$$

由于在实际计算用户数据速率时存在可能会超过实际可达的最大数据速率情况, 因此本文引入中断概率来衡量在不完美信道下得到的数据速率超过实际可达的最大数据速率时的性能。则系统总速率[8]重新表示为

$$\mathcal{R}(t) = \sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) b_{k_i,l}^b(t) \cdot r_{k_i,l}^b(t) P_r[r_{k_i,l}^b(t) \leq c_{k_i,l}^b(t) | \bar{g}_{k_i,l}^b(t)] \quad (9)$$

其中, $P_r[r_{k_i,l}^b(t) \leq c_{k_i,l}^b(t) | \bar{g}_{k_i,l}^b(t)]$ 为在估计瑞利衰落系数为 $\bar{g}_{k_i,l}^b(t)$ 条件下时, 用户 k_i 不中断的概率。

因此系统长期平均总吞吐量为

$$\bar{\mathcal{R}} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[\mathcal{R}(t)] \quad (10)$$

2.1.2 用户队列模型

在 t 时隙, 切片 k 中的第 i 个用户 k_i 的缓冲区中等待发送的数据包长度为 $Q_{k_i}(t)$, 其缓冲区的最大队列存储长度为 d_Q 。 $A_{k_i}(t)$ 表示在 t 时隙用户 k_i 的数据包到达过程。则用户 k_i 的队列更新过程为

$$Q_{k_i}(t+1) = \max\{Q_{k_i}(t) + A_{k_i}(t) - D_{k_i}(t), 0\} \quad (11)$$

其中, $D_{k_i}(t)$ 为用户 k_i 在 t 时隙内传输的数据包个数, 且 $D_{k_i}(t) = \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) r_{k_i,l}^b(t) / S$, S 为数据包大小。根据Little定理, 可知保证用户的时延性能可以刻画为保证队列的稳定性。本文定义为

$$\overline{Q_{k_i}} \triangleq \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} E[Q_{k_i}(t)] < \infty \quad (12)$$

2.1.3 问题建模

本文在保证用户队列时延, 满足切片最低平均速率需求和用户速率中断概率等约束下, 最大化系统总的平均吞吐量可数学建模为

$$\left. \begin{aligned}
& \text{P1: } \max_{a_{k_i,l}(t), b_{k_i,l}^b(t), p_{k_i,l}^b(t)} \bar{\mathcal{R}} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} [\mathcal{R}(t)] \\
& \text{s.t. C1: } \bar{Q}_{k_i} < \infty, \forall k_i \in U_k, \forall k \in \mathcal{K}; \\
& \text{C2: } \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \\
& \mathbb{E} \left[\sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) b_{k_i,l}^b(t) r_{k_i,l}^b(t) \right] \geq R_k^{\min}, \\
& \forall k \in \mathcal{K}; \\
& \text{C3: } P_r[c_{k_i,l}^b(t) < r_{k_i,l}^b(t) | \bar{g}_{k_i,l}^b(t)] \leq \varepsilon_{\text{out}}, \\
& \forall b \in \mathcal{B}, \forall l \in \mathcal{L}, \forall k_i \in U; \\
& \text{C4: } \sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) p_{k_i,l}^b(t) \leq p_l^{\max}, \forall l \in \mathcal{L}; \\
& \text{C5: } \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) r_{k_i,l}^b(t) \geq R_{k_i}^{\min}, \\
& \forall k_i \in U; \\
& \text{C6: } \sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) r_{k_i,l}^b(t) \leq C_{l,\max}, \\
& \forall l \in \mathcal{L}; \\
& \text{C7: } \sum_{l \in \mathcal{L}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) \leq M_b^{\max}, \forall b \in \mathcal{B}; \\
& \text{C8: } \sum_{l \in \mathcal{L}} a_{k_i,l}(t) = 1, \forall k_i \in U; \\
& \text{C9: } \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) = 1, \forall k_i \in U; \\
& \text{C10: } p_{k_i,l}^b(t) \geq 0, a_{k_i,l}(t), b_{k_i,l}^b(t) \in \{0, 1\}, \\
& \forall b \in \mathcal{B}, \forall l \in \mathcal{L}, \forall k_i \in U
\end{aligned} \right\} (13)$$

其中, C1为切片用户队列时延和稳定性约束; C2为各切片长期平均最小速率约束; C3为中断概率约束; C4为RRH发射功率约束; C5为切片用户的最小保证比特率约束; C6为RRH与BBU池的fronthaul链路容量约束; C7为一个RB上最大的功率复用用户个数约束; C8表示每个切片用户只能与1个RRH关联; C9表示每个切片用户只能复用在1个RB上; C10为保证功率非负性和二进制变量的约束。

2.2 优化问题转化

2.2.1 概率混合问题转换

混合概率问题P1需转换为无概率的问题, 则将中断概率约束嵌入到系统平均总速率公式中^[8], 为

$$\bar{\mathcal{R}} = \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} \left[\sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} (1 - \varepsilon_{\text{out}}) \cdot a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \right] \quad (14)$$

其中, $\tilde{r}_{k_i,l}^b(t) = B_s \log_2(1 + \gamma_{k_i,l}^b(t))$, 且

$$\gamma_{k_i,l}^b(t) = \frac{\varepsilon_{\text{out}} F^{-1} \left(\frac{\varepsilon_{\text{out}}}{2} \right) (L_{k_i,l}^b(t))^2 p_{k_i,l}^b(t)}{2 |h_{k_i,l}^b(t)|^2 \sum_{n_j \in M_{k_i,l}^b(t)} p_{n_j,l}(t) + \varepsilon_{\text{out}} \sigma_z^2} \quad (15)$$

F^{-1} 为卡方分布累计分布函数的反函数, 因此, P1问题可改写为

$$\left. \begin{aligned}
& \text{P2: } \max_{a_{k_i,l}(t), b_{k_i,l}^b(t), p_{k_i,l}^b(t)} \bar{\mathcal{R}} \\
& \text{s.t. } \tilde{\text{C1:}} \bar{Q}_{k_i} < \infty, \forall k_i \in U_k, \forall k \in \mathcal{K}; \\
& \tilde{\text{C2:}} \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} \\
& \mathbb{E} \left[\sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \right] \\
& \geq R_k^{\min}, \forall k \in \mathcal{K}; \\
& \tilde{\text{C3:}} \sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) p_{k_i,l}^b(t) \leq p_l^{\max}, \forall l \in \mathcal{L}; \\
& \tilde{\text{C4:}} \sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \leq C_{l,\max}, \\
& \forall l \in \mathcal{L}; \\
& \tilde{\text{C5:}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \geq R_{k_i}^{\min}, \\
& \forall k_i \in U; \\
& \tilde{\text{C6:}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U} a_{k_i,l}(t) b_{k_i,l}^b(t) \leq M_b^{\max}, \forall b \in \mathcal{B}; \\
& \tilde{\text{C7:}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) = 1, \forall k_i \in U; \\
& \tilde{\text{C8:}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) = 1, \forall k_i \in U; \\
& \tilde{\text{C9:}} p_{k_i,l}^b(t) \geq 0, a_{k_i,l}(t), b_{k_i,l}^b(t) \in \{0, 1\}, \\
& \forall b \in \mathcal{B}, \forall l \in \mathcal{L}, \forall k_i \in U
\end{aligned} \right\} (16)$$

可看出问题P2是一个长期时间平均优化问题, 其约束条件有两种时间尺度约束: 瞬时约束、连续时隙上平均约束。因此, 用Lyapunov理论设计一种权衡网络吞吐量和功率分配及时延间的约束关系。

2.2.2 Lyapunov 优化理论分析

为便于处理问题P2的 $\tilde{\text{C2}}$, 将其转换为切片虚拟队列 $Z_k(t)$, 其更新为

$$Z_k(t+1) = \max \{ Z_k(t) + R_k^{\min} - \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) \cdot b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t), 0 \}, \forall k \in \mathcal{K} \quad (17)$$

其虚拟队列并不具有任何实际传输的意义。

文中 t 时隙时系统的队列向量则可表示为 $\Omega(t) = [Q(t), Z(t)]$ ，定义Lyapunov函数为

$$L(\Omega(t)) = \frac{1}{2} \left[\sum_{k \in \mathcal{K}} \sum_{k_i \in U_k} (Q_{k_i}(t))^2 + \sum_{k \in \mathcal{K}} (Z_k(t))^2 \right] \quad (18)$$

定义单时隙Lyapunov偏移为^[9]

$$\Delta(\Omega(t)) = E\{L(\Omega(t+1)) - L(\Omega(t)) | \Omega(t)\} \quad (19)$$

Lyapunov偏移与加权总吞吐量之差定义为

$$\begin{aligned} & \Delta(\Omega(t)) - VE\{\mathcal{R}(t) | \Omega(t)\} \\ &= E\{L(\Omega(t+1)) - L(\Omega(t)) | \Omega(t)\} \\ & \quad - VE\{\mathcal{R}(t) | \Omega(t)\} \\ & \leq B + E\left\{ \sum_{k \in \mathcal{K}} \sum_{k_i \in U_k} \{Q_{k_i}(t)[A_{k_i}(t) - D_{k_i}(t)] | \Omega(t)\} \right\} \\ & \quad + E\left\{ \sum_{k \in \mathcal{K}} \left\{ Z_k(t)[R_k^{\min} - \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) \right. \right. \\ & \quad \cdot b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) | \Omega(t) \} \} \\ & \quad - VE\left\{ \sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} (1 - \varepsilon_{\text{out}}) a_{k_i,l}(t) \right. \\ & \quad \cdot b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) | \Omega(t) \} \} \end{aligned} \quad (20)$$

其中， $V > 0$ 是权衡偏移与代价函数的控制参数， B 为有限的正常量。由Lyapunov优化理论得，P2问题可转换为在每一时隙上最小化不等式(20)的右侧

$$\begin{aligned} \text{P3: } & \min_{a_{k_i,l}(t), b_{k_i,l}^b(t), p_{k_i,l}(t)} -V \sum_{k \in \mathcal{K}} \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} (1 - \varepsilon_{\text{out}}) \\ & \cdot a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \\ & + \sum_{k \in \mathcal{K}} \sum_{k_i \in U_k} \left\{ Q_{k_i}(t) \left[A_{k_i}(t) \right. \right. \\ & \quad \left. \left. - \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) / S \right] \right\} \\ & + \sum_{k \in \mathcal{K}} \left\{ Z_k(t) \left[R_k^{\min} \right. \right. \\ & \quad \left. \left. - \sum_{b \in \mathcal{B}} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) b_{k_i,l}^b(t) \tilde{r}_{k_i,l}^b(t) \right] \right\} \\ & \text{s.t. } \tilde{\text{C3}} \sim \tilde{\text{C9}} \end{aligned} \quad (21)$$

3 联合用户关联和功率分配机制

本文所提C-RAN场景下基于NOMA下行链路网络切片资源分配的算法如算法1所示。

3.1 用户关联问题

定义用户 k_i 与RRH l 关联的宽带信干噪比(wide-band SINR) $\Psi_{k_i,l}(t)$ 为

$$\Psi_{k_i,l}(t) = \frac{p_l^{\max} |h_{k_i,l}(t)|^2}{\sum_{s \in \mathcal{L} \setminus \{l\}} p_s^{\max} |h_{k_i,s}(t)|^2 + \sigma_z^2} \quad (22)$$

其中， $h_{k_i,l}(t)$ 为 t 时隙用户 k_i 与RRH l 间的信道增益。则本文用户关联问题建模为

$$\max_{a_{k_i,l}(t)} \sum_{l \in \mathcal{L}} \sum_{k_i \in U_k} a_{k_i,l}(t) \Psi_{k_i,l}(t) \quad (23)$$

$$\text{s.t. } \sum_{b \in \mathcal{B}} \sum_{k_i \in U} a_{k_i,l}(t) R_{k_i}^{\min} \leq C_{l,\max}, \quad \forall l \in \mathcal{L} \quad (23a)$$

$$\sum_{l \in \mathcal{L}} a_{k_i,l}(t) = 1, \quad \forall k_i \in U \quad (23b)$$

$$a_{k_i,l}(t) \in \{0, 1\}, \quad \forall l \in \mathcal{L}, \forall k_i \in U \quad (23c)$$

式(23a)为RRH上的fronthaul链路容量至少能承载与之关联用户所需最小数据速率，式(23b)为用户 k_i 只与一个RRH关联，式(23c)表明关联因子为二进制变量。

用户关联问题为0-1多重背包问题，通过所提贪婪算法寻求一个可行的次优解，具体见算法2。

算法1：基于NOMA的网络切片资源分配全局算法

(1) 初始化：设置控制参数 V ，设所有用户队列 $Q_{k_i}(0) = 0$ ，切片虚拟队列 $Z_k(0) = 0$ ，设时隙长度 $T_{\max}$ ；

(2) For $t = 0, 1, \dots, T_{\max} - 1$ do；

(3) 观察该时隙的每个用户的队列状态 $Q_{k_i}(t)$ ，切片虚拟队列 $Z_k(t)$ 以及信道状态信息；

(4) 执行算法2获得用户与RRH关联次最优用户关联策略 $\mathbf{a}$ ；

(5) 在用户与RRH关联的基础上执行算法3获得与RRH关联的用户RB复用策略 $\mathbf{b}$ 和功率分配方案 $\mathbf{p}$ ；

(6) 输出当前时隙的资源分配方案 $\mathbf{a}, \mathbf{b}, \mathbf{p}$ ；

(7) 由式(11)，式(17)更新下一时隙 $Q_{k_i}(t+1)$ ， $Z_k(t+1)$ ；

(8) $t = t+1$ ；

(9) End

算法2：用户关联贪婪算法

(1) 初始化：设参数 $\bar{\mathcal{L}} = \mathcal{L}$ 和 $\bar{C}_l = C_{l,\max}$ ， $a_{k_i,l}(t) = 0, f_l(t) = 0, \forall l \in \mathcal{L}, \forall k_i \in U$ ；

(2) for all $k_i \in U, l \in \mathcal{L}$ do；

(3) 计算 $f_l(t) = \Psi_{k_i,l}(t)$ ；

(4) 选择 $l = \arg \max_{p \in \bar{\mathcal{L}}} f_p(t)$ ；

(5) If $\bar{C}_l \geq R_{k_i}^{\min}$ then；

(6) $a_{k_i,l}(t) = 1$ ，更新 $\bar{C}_l = \bar{C}_l - R_{k_i}^{\min}$

(7) Else

(8) $\bar{\mathcal{L}} = \bar{\mathcal{L}} \setminus \{l\};$

(9) end if

(10) end for

3.2 功率分配问题

在得到用户关联策略的基础上, 功率分配问题是一个非凸优化问题, 采用连续凸逼近^[10]和参数转换方法来解决, 且令 $p = e^{\tilde{p}}$, 则

$$\gamma_{k_i,l}^b(t) = \frac{\varepsilon_{\text{out}} F^{-1} \left| \frac{h_{k_i,l}^b(t)}{g_{k_i,l}^b(t)} \right|^2 (\varepsilon_{\text{out}}/2) (L_{k_i,l}^b(t))^2 e^{\tilde{p}_{k_i,l}^b(t)}}{2 |h_{k_i,l}^b(t)|^2 \sum_{n_j \in M_{k_i,l}^b(t)} e^{\tilde{p}_{n_j,l}^b(t)} + \varepsilon_{\text{out}} \sigma_z^2} \quad (24)$$

$$p_{k_i,l}^{b*}(t) = e^{\tilde{p}_{k_i,l}^{b*}(t)} = \left[\frac{B_s A_{k_i,l}^b(t) \frac{1}{\ln 2}}{\frac{\rho_l}{V(1 - \varepsilon_{\text{out}}) + Q_{k_i}(t)/S + Z_k(t) - v_l + \delta_{k_i}} - \sum_{f_m \in \bar{M}_{k_i,l}^b(t)} E_{f_m,l}^b(t)} \right]^+ \quad (26)$$

其中, $[x]^+ = \max(0, x)$,

$$E_{f_m,l}^b(t) = - \frac{B_s A_{f_m,l}^b(t) 2 |h_{f_m,l}^b(t)|^2}{\ln 2 \varepsilon_{\text{out}} F^{-1} \left| \frac{h_{f_m,l}^b(t)}{g_{f_m,l}^b(t)} \right|^2 (\varepsilon_{\text{out}}/2) (L_{f_m,l}^b(t))^2 e^{\tilde{p}_{f_m,l}^b(t)}} \gamma_{f_m,l}^b(t),$$

$\rho_l, v_l, \delta_{k_i}, \pi_b, \lambda_{k_i}$ 为非负拉格朗日乘子。

经分析上式的1阶偏导是单调递增的, 使用二分搜索法找关于 $\tilde{p}_{k_i,l}^b(t)$ 求偏导后精度较高的数值解。

对 $b_{k_i,l}^b(t)$ 求偏导且满足KKT条件, 可推出

$$b_{k_i,l}^{b*}(t) = \begin{cases} 1, (k_i^*, l^*) = \arg \min_{b \in \mathcal{B}} H_{k_i,l}^b \\ 0, \text{其它} \end{cases}, \quad \forall k_i \in U, \forall l \in \mathcal{L} \quad (27)$$

其中,

$$H_{k_i,l}^b = (-V(1 - \varepsilon_{\text{out}}) - Q_{k_i}(t)/S - Z_k(t) + v_l - \delta_{k_i}) r_{k_i,l}^b(t) + \rho e^{\tilde{p}_{k_i,l}^b(t)}(t) \quad (28)$$

功率分配子问题的求解如算法3。

算法3: 基于连续凸逼近的功率分配算法

(1) 在时隙 t , 观察每个用户的队列状态 $Q_{k_i}(t)$, 切片虚拟队列 $Z_k(t)$ 及信道状态信息;

(2) 初始化: 近似值向量 $\mathbf{A} = \mathbf{1}$, $\mathbf{B} = \mathbf{0}$, 最大迭代次数 N_{max}^1 , N_{max}^2 和收敛条件 $\varepsilon_1, \varepsilon_2, \boldsymbol{\rho}^0, \mathbf{v}^0, \boldsymbol{\delta}^0$;

(3) While $m \leq N_{\text{max}}^1$ or $|F^{m+1}(t) - F^m(t)| \geq \varepsilon_1$ do;

(4) While $n \leq N_{\text{max}}^2$ or $|L^{n+1}(t) - L^n(t)| \geq \varepsilon_2$ do;

(a) 假设 $b_{k_i,l}^b(t) = 1$, 通过KKT条件和二分搜索法求得功率分配策略 $p_{k_i,l}^{b*}(t)$;

因此与RRH关联且功率复用在RB b 上的用户 k_i 的数据速率下界可表示为

$$r_{k_i,l}^{*b}(t) = B_s [A_{k_i,l}^b(t) \log_2(\gamma_{k_i,l}^b(t)) + B_{k_i,l}^b(t)] \quad (25)$$

其中, $A_{k_i,l}^b(t) = \gamma_{k_i,l}^b(t)/(1 + \gamma_{k_i,l}^b(t))$, $B_{k_i,l}^b(t) = \log_2(1 + \gamma_{k_i,l}^b(t)) - A_{k_i,l}^b(t) \log_2 \gamma_{k_i,l}^b(t)$ 。

根据文献^[10], 将 $b_{k_i,l}^b(t)$ 松弛为 $0 < b_{k_i,l}^b(t) < 1$, 则RB复用和功率分配优化问题可进行重写。且该问题是一个凸优化问题^[11], 因此用拉格朗日对偶方法求解, 利用KKT条件, 对拉格朗日函数求 $\tilde{p}_{k_i,l}^b(t)$ 偏导得切片 k 中用户 k_i 在 t 时隙的功率资源量为

(b) 由式(27)求切片用户复用RB策略;

(c) 若 $b_{k_i,l}^{b*}(t) = 0$, 则 $p_{k_i,l}^{b*}(t) = 0$;

(5) 根据次梯度算法更新拉格朗日乘子 $\boldsymbol{\rho}^n, \mathbf{v}^n, \boldsymbol{\delta}^n$;

(6) $n = n+1$;

(7) End while

(8) 更新近似值向量, 使用户速率接近它的下界;

(9) $m = m+1$;

(10) End while

4 性能仿真与结果分析

为验证本文最大化系统平均切片总吞吐量而提出的基于NOMA的网络切片动态资源分配算法的有效性, 考虑一个在C-RAN场景下为3个网络切片服务用户的覆盖区域为500 m × 500 m的物理网络, 由5个RRH构成。并将其与OFDMA中进行的用户关联和功率分配的算法(OFDMA)、文献^[12]中NOMA下进行用户关联和功率分配的匹配算法(NOMA-Match)、文献^[13]在NOMA系统中在RB之间平均分配功率后再进行的资源分配算法(NOMA-Equal)进行了对比。仿真参数设置见表1。

图2为3个切片共60个用户时, 切片平均总吞吐量走势。可看出, 时间平均总吞吐量随 V 的增加不断提升并趋于平稳, 而数据包队列长度也随 V 的增加以近似 $O(V)$ 线性递增。这是由于控制参数 V 为系统平均吞吐量和时延之间的折中因子, V 越大, 时间平均总吞吐量更大, 但会以更大的数据队列积压为代价, 因此需选择一个合理的控制参数。

图3为不同切片的用户平均队列长度在连续时

表 1 仿真参数

参数	数值
RB数量	35
RB最大复用用户数	3
RB带宽	180 kHz
RRH功率 $p_t^{\max}$	30 dBm
fronthaul容量 $C_{l,\max}$	100 Mb/s
切片用户最低速率需求	500 kb/s, 1 Mb/s, 2 Mb/s
用户数据包到达率	3 packets/slot
路径损耗衰落模型	$157.4+32\lg(d)(d[\text{km}])$
噪声功率谱密度 N_0	-174 dBm/Hz
时隙长度, $\sigma_e^2, \varepsilon_{\text{out}}$	5 ms, 0.05, 0.10

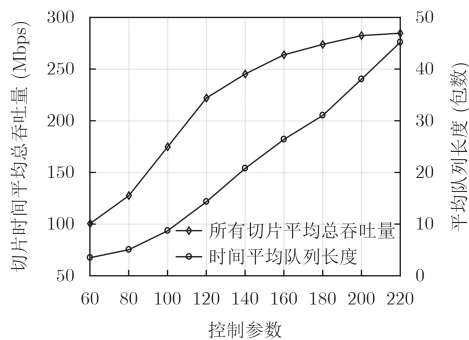


图 2 切片平均总吞吐量和平均队列长度

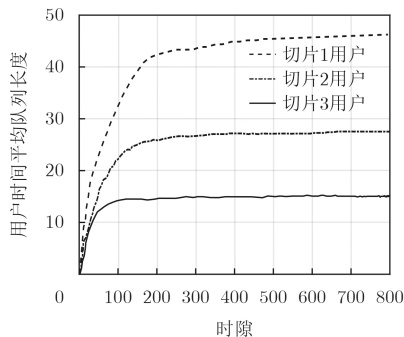


图 3 不同切片用户在连续时隙上的队列变化

隙上的收敛过程，其中任意选取了3个分别属于不同切片的用户。可看出，用户平均队列长度先随时隙的增长逐渐增大，最后趋于平稳。此图证明了本文所提基于Lyapunov的优化方法能够有效保证队列的稳定性，同时也体现了本文所提算法的有效性。

由图4可见，用户数据包到达率越高，时延越低，原因是当用户数据包到达率较高时，为保证队列稳定性用户数据包处理速率会相应地增大，需要给该切片用户分配较多的资源。当用户数据包处理速率的增长率大于用户队列长度的增加速率时，就会出现图4所示其平均等待时延反而越小的情况。

图5为系统平均总吞吐量随切片中用户数增多的变化情况。在4种算法下系统平均总吞吐量都是

随切片中用户数量的增多而增大，但是增长速率越来越慢，这是由于系统中功率资源有限，不能无限接入用户并满足他们的需求。从图中可看出，本文所提算法明显优于其他3种算法，这是因为本文联合考虑了切片状态、用户队列和信道质量的状态，能动态地在每一时隙实施用户关联和资源分配决策。NOMA系统的1个RB可功率复用多个用户，而OFDMA系统1个RB只能分给1个用户，未充分利用频谱资源。因此OFDMA算法的性能最差。

图6描述了不同算法中时间平均队列积压与数据包到达率的关系，对于任意的数据包到达率，本文NOMA算法的队列积压最少，这是由于本文所提算法在每个时隙上都考虑了切片的状态以及切片用户的队列缓存长度，然后动态地为用户分配资源。

图7为在4种算法下平均切片总吞吐量与fronthaul

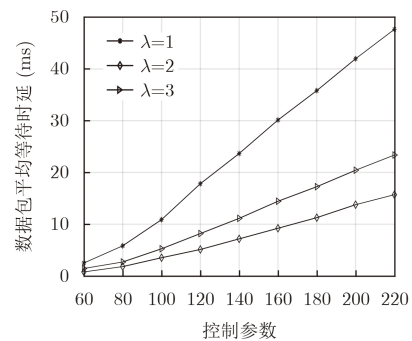


图 4 平均时延与控制参数 V 的关系

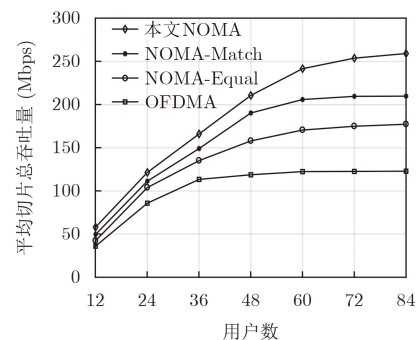


图 5 不同资源分配方案平均切片吞吐量的比较

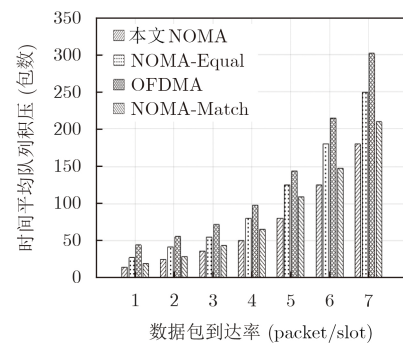


图 6 不同资源分配方案的平均队列积压比较

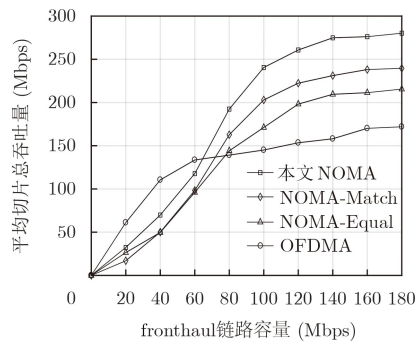


图7 平均切片总吞吐量与fronthaul链路容量

链路容量关系比较, 可看到随着fronthaul链路容量增加, 系统平均总吞吐量不断提升, 但增加速率逐渐变得缓慢, 这是由于在功率资源足够的时候增加fronthaul链路容量会相应的增大所有切片总吞吐量, 但是当功率资源使用量达到极限时, 尽管再增加fronthaul链路容量也不会无限提升切片的平均吞吐量。当fronthaul链路容量较小时, OFDMA系统的总吞吐量高于本文NOMA系统, 因为基于本文NOMA的资源分配方案1个RB能复用多个用户, 但是fronthaul链路容量较小不能满足每个切片用户的需要, 吞吐量性能反而不如OFDMA系统。但是随着fronthaul链路容量的增加, 本文NOMA系统的资源分配方案吞吐量逐渐超过了OFDMA系统。

5 结束语

本文基于网络切片提出了一种最大化时间平均切片总吞吐量的联合用户关联和功率资源分配算法。在NOMA系统下行链路信道条件不完美情况下, 将概率混合问题转换为非概率问题, 并利用Lyapunov理论设计了一种基于当前切片虚拟队列状态、用户队列状态和信道状态的联合用户关联和功率分配的算法。同时采用了有较低复杂度的贪婪算法解决切片中用户关联问题求得次优解, 然后将Lyapunov优化中的非凸优化问题利用连续凸逼近方法将其转换为凸优化, 并采用拉格朗日对偶分解法获得功率分配策略。仿真结果表明, 该算法可保证网络切片用户时延, 提高时间平均切片总吞吐量。

参考文献

- [1] LIU Gang, YU F R, JI Hong, *et al.* Distributed resource allocation in virtualized full-duplex relaying networks[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(10): 8444–8460. doi: [10.1109/TVT.2015.2513070](#).
- [2] YIN Lei, QIU Ling, and CHEN Zheng. Throughput-maximum resource provision in the OFDMA-based wireless virtual network[C]. *IEEE 85th Vehicular Technology Conference (VTCSpring)*, Sydney, Australia, 2017: 1–6. doi: [10.1109/VTCSpring.2017.8108502](#).
- [3] SINAIE M, NG D W K, and JORSWIECK E A. Resource allocation in NOMA virtualized wireless networks under statistical delay constraints[J]. *IEEE Wireless Communications Letters*, 2018, 7(6): 954–957. doi: [10.1109/LWC.2018.2841852](#).
- [4] DAWADI R, PARSAEFARD S, DERAKHSHANI M, *et al.* Power-efficient resource allocation in NOMA virtualized wireless networks[C]. *IEEE Global Communications Conference (GLOBECOM)*, Washington, USA, 2016: 1–6. doi: [10.1109/GLOCOM.2016.7842162](#).
- [5] LEE Y L, LOO J, CHUAH T C, *et al.* Dynamic network slicing for multitenant heterogeneous cloud radio access networks[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(4): 2146–2161. doi: [10.1109/TWC.2017.2789294](#).
- [6] HA V N and LE L B. End-to-end network slicing in virtualized OFDMA-based cloud radio access networks[J]. *IEEE Access*, 2017, 5: 18675–18691. doi: [10.1109/ACCESS.2017.2754461](#).
- [7] IKKI S S and AISSA A. Two-way amplify-and-forward relaying with Gaussian imperfect channel estimations[J]. *IEEE Communications Letters*, 2012, 16(7): 956–959. doi: [10.1109/LCOMM.2012.050912.120103](#).
- [8] WANG Xiaoming, ZHENG Fuchun, ZHU Pengcheng, *et al.* Energy-efficient resource allocation in coordinated downlink multicell OFDMA systems[J]. *IEEE Transactions on Vehicular Technology*, 2016, 65(3): 1395–1408. doi: [10.1109/TVT.2015.2413950](#).
- [9] XIANG Xudong, LIN Chuang, CHEN Xin, *et al.* Toward optimal admission control and resource allocation for LTE-A femtocell uplink[J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(7): 3247–3261. doi: [10.1109/TVT.2014.2351837](#).
- [10] PAPANDRIOPOULOS J and EVANS J S. SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks[J]. *IEEE Transactions on Information Theory*, 2009, 55(8): 3711–3724. doi: [10.1109/TIT.2009.2023751](#).
- [11] FOOLADIVANDA D and ROSENBERG C. Joint resource allocation and user association for heterogeneous wireless cellular networks[J]. *IEEE Transactions on Wireless Communications*, 2013, 12(1): 248–257. doi: [10.1109/TWC.2012.121112.120018](#).
- [12] ZHU Jianyue, WANG Jiaheng, HUANG Yongming, *et al.* On optimal power allocation for downlink non-orthogonal multiple access systems[J]. *IEEE Journal on Selected Areas in Communications*, 2017, 35(12): 2744–2757. doi: [10.1109/JSAC.2017.2725618](#).
- [13] PARIDA P and DAS S S. Power allocation in OFDM based NOMA systems: A DC programming approach[C]. *IEEE GLOBECOM Workshops*, Austin, USA, 2015: 1026–1031.

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, 网络功能虚拟化, 无线资源分配等。

杨 恒: 男, 1993年生, 硕士生, 研究方向为切片及虚拟网络。

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。