

# 宏基站和微基站群之间的预测性信道调度算法

谢映海<sup>1,2</sup>, 姚若河<sup>2</sup>, 吴斌<sup>1</sup>

(1. 珠海中慧微电子博士后工作站, 广东 珠海 519085; 2. 华南理工大学电子与信息学院, 广东 广州 510640)

**摘 要:** 针对 5G 超蜂窝网络的宏基站和微基站群之间的非实时业务传输, 提出了一种新型预测性信道调度算法。首先, 利用静止通信体之间无线信道的随机平稳过程特征, 从古典概率论的角度为调度过程建立了一个离散信道状态概率空间并对其进行事件域分割; 然后, 通过各事件域的概率数值计算实现对多用户多非实时业务的高效调度。理论分析和仿真结果表明, 所提算法计算复杂度低, 和其他一些经典调度算法相比, 所提算法可在更长时间维度上进行业务优化传输, 在吞吐量性能上逼近最大信噪比算法, 重负荷情况下系统吞吐量提升了约 14%; 同时通过精确量化计算实现了业务期望速率和实际调度速率的自适应匹配。

**关键词:** 超蜂窝网络; 平稳随机过程; 离散信道状态概率空间; 预测性信道调度; 量化调度

**中图分类号:** TP393

**文献标识码:** A

**doi:** 10.11959/j.issn.1000-436x.2019217

## Predictive channel scheduling algorithm between macro base station and micro base station group

XIE Yinghai<sup>1,2</sup>, YAO Ruohu<sup>2</sup>, WU Bin<sup>1</sup>

1. Post-doctoral Workstation of Zhuhai Zhonghui Microelectronics, Zhuhai 519085, China

2. School of Electronics and Information, South China University of Technology, Guangzhou 510640, China

**Abstract:** A novel predictive channel scheduling algorithm was proposed for non-real-time traffic transmission between macro-base stations and micro-base stations in 5G ultra-cellular networks. First, based on the stochastic stationary process characteristics of wireless channels between stationary communication agents, a discrete channel state probability space was established for the scheduling process from the perspective of classical probability theory, and the event domain was segmented. Then, the efficient scheduling of multi-user, multi-non-real-time services was realized by probability numerical calculation of each event domain. The theoretical analysis and simulation results show that the algorithm has low computational complexity. Compared with other classical scheduling algorithms, the new algorithm can optimize traffic transmission in a longer time dimension, approximate the maximum signal-to-noise ratio algorithm in throughput performance, and increase system throughput by about 14% under heavy load. At the same time, the new algorithm is accurate. Quantitative computation achieves a self-adaption match between the expected traffic rate and the actual scheduling rate.

**Key words:** super cellular network, stationary stochastic process, discrete channel state probability space, predictive channel scheduling, quantitative scheduling

### 1 引言

超蜂窝网络作为 5G 系统的一项核心技术, 是

在传统宏蜂窝网络的基础上叠加不同种类的低功率基站或接入点, 例如家庭基站、Wi-Fi 接入点和中继节点等微基站<sup>[1-3]</sup>。新的网络架构面临诸多技术

收稿日期: 2019-03-14; 修回日期: 2019-09-24

基金项目: 广东省科技计划基金资助项目 (No.2016B010123004)

Foundation Item: The Guangdong Science and Technology Project (No.2016B010123004)

挑战,宏基站和微基站群之间的信道资源调度就是其中之一。比例公平算法是一种最经典的调度算法,该算法综合考虑了各个用户的业务公平服务和实时信道状态,成为蜂窝系统的一种常见调度机制<sup>[4-6]</sup>。在此基础上,文献[7]进一步将待传输数据的时延因素纳入信道调度决策过程,更好地满足了对时延比较敏感的业务服务质量要求<sup>[7]</sup>。但这些算法都仅考虑各个用户的当前信道状态,在时间维度上具有一定的局限性,无法实现系统调度性能的最优化。

5G 系统商用化时代即将到来,其中一个重要的标志就是 1 ms 的低时延。但除了自动驾驶等数据量小且对时延高度敏感的业务,占据网络流量的主要业务是内容分发等对时延不太敏感的业务(如点播式视频传输),甚至完全不敏感的业务(如文件下载),统称为非实时业务。如文献[8]所示,随着智能手机的普遍应用,2016 年无线网络中超过 50% 的业务量来自移动视频业务,预计到 2021 年这个比例将上升至 78%。

5G 系统中的基站数量、用户数量和业务传输量都快速增长,系统能耗急剧增加<sup>[9]</sup>,对业务传输过程进行优化来降低能耗成为研究焦点。通过使用预测性信息来提升非实时业务的无线传输效率引起了科研人员的关注。文献[10-11]认为人类行为具有可预测性,对于已经存储在内容服务器的非实时流媒体业务,利用预测的用户级情境信息(用户移动轨迹)和网络级情境信息(业务流量),在一个多用户的场景中使用一个低复杂度的资源规划与分配方法进行资源优化分配,考虑到用户业务对延迟的非敏感性,在密集基站应用场景下让移动用户在离基站较近时优先进行传输,提升了每个用户的服务质量和系统的综合性能。文献[12]根据预测的用户运动轨迹和射频信号地图来获得未来平均信道增益,并通过理论分析得出如下结论:通信系统利用未来平均信道增益或利用未来瞬时信道增益可以获得相同水平的能量效率。文献[13]假设已知未来的平均信道增益,在保障文件传输中断概率的前提下,研究了降低系统总能耗的资源分配策略。文献[14-15]在点播式视频开始下载和观看时,基于未来的平均信道增益对未来一段时间内的资源分配进行了规划,再根据瞬时信道增益进一步优化视频播放过程中的功率控制,得到一个双时间尺度的高能效传输策略。综合分析文献[10-15],这些算法的基本原理都是通过预测未来平均信道增益,从而更激进地

利用信道分集,获得了较大的系统性能增益。

大尺度衰落和小尺度衰落共同决定了无线信道的信号传输特性,超蜂窝网络的宏基站和微基站一般都是固定不动的,因此其大尺度衰落基本上不随时间发生较大变化,但通信路径上可能存在大量移动反射体,因此由多条信号路径的相长和相消干扰造成的小尺度衰落仍可能不断变化,尤其是 5G 系统中波长较短的毫米波宽带信号。与无线移动通信相比,这种静止通信体之间的信道虽然呈现一定的随机性,但其在时域和频域上的变化相对较缓慢,且在概率意义上呈现出较稳定的统计特征,因此可以认为其信道状态情况满足平稳随机过程的条件,在较大的时间维度上具有较强的可预测性。

本文利用超蜂窝网络的宏基站和微基站之间的静态无线信道的平稳随机特性,从古典概率论的角度出发,在信道调度过程中为调度用户建立了一个离散信道状态概率空间,然后对该空间进行互不相容的离散事件域分割,并计算出各个事件域发生的具体概率值。据此针对非实时业务,通过严格的数值计算在时域和频域上实现对信道资源的可量化调度。理论分析和仿真结果表明,所提算法充分利用了无线随机信道统计特性的平稳性,通过带有预测性的调度机制在更长的时间内优化信道资源的分配使用,在进行高谱效和高能效的信道资源利用的同时,实现了各用户的不同业务期望速率和其对应的调度速率的趋同,满足了调度算法的 3 个基本目标:信道利用效果高,多用户公平性强和业务用户体验好。

## 2 预测性信道调度算法

### 2.1 无线信道模型和信道调度模型

能量效率已经成为 5G 移动通信系统的一个重要设计目标,超蜂窝网络架构在传统的宏蜂窝网络的基础上叠加不同种类的低功率微基站,可以有效减少通信距离,大幅度降低单位比特的能量平均损耗,其架构如图 1 所示。

一般而言,这些微基站固定安装在建筑物墙壁、路边电线杆或小区楼顶等,因此和传统的无线移动通信信道模型不同,宏基站和这些微基站都是静止通信体,信道模型相对比较简单,其信道的随机时变性主要是由小尺度衰落而非大尺度衰落造成的。在 3GPP 系统下,2.0 GHz 载波频率的一个路径传输模型如式(1)所示。

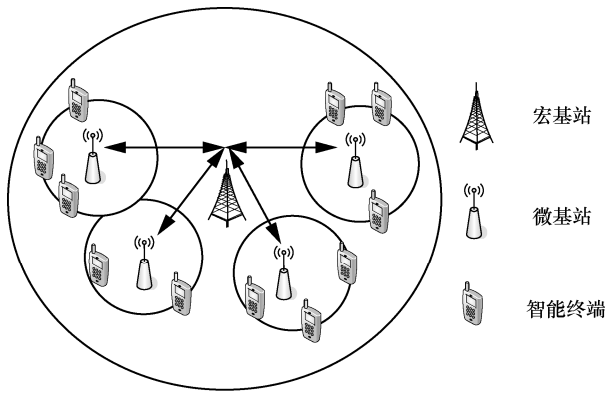


图 1 超蜂窝网络架构

$$2.0 \text{ GHz}: L(d) = 128.1 + 37.6 \lg(d) + X + Y \quad (1)$$

其中,  $L(d)$  表示路径损耗,  $d$  表示传输距离,  $X$  表示服从方差  $\delta^2$  的对数正态分布,  $Y$  表示服从均值为  $\mu$  的指数分布。

由式(1)的信道模型可知, 在传输距离不变的情况下, 信道情况仍然呈现出一定的随机性, 但满足平稳随机过程的条件, 其统计特性不会随着时间变化而变化, 因此宏基站侧借助所积累的历史数据, 在较长的时间维度上可以从统计角度对其自身和微基站之间的信道状态的分布情况进行较为精准的预测。

点播式视频等非实时流媒体业务将在 5G 网络中占据非常高比例的流量, 这些视频业务将存储在宏基站的内容服务器中, 大量使用智能手机的用户将基于微基站中继方式在同一个宏基站覆盖范围内进行不同视频内容的观看。保障这类业务服务质量的核心在于业务平均传输速率, 在初始一小段时间的视频内容缓存后, 只要系统在较长的时间维度上可以为该业务提供一个略微大于业务门限速率的业务平均传输速率, 用户就可以获取较为流畅的流媒体观看体验, 可使用户评价达到较高水平。而一旦达到某一门限速率后, 系统即使再提高业务传输速率, 也基本上不会提升用户评价了, 因此为这类业务产生传输需求时会向调度方提供一个业务期望速率值, 如果其在调度过程获得了不小于该值的平均传输速率就可以获得较高的用户评价。点播式视频业务用户观看评价和业务平均传输速率之间的关系如图 2 所示。

本文以点播式视频业务等非实时业务的调度为研究对象。下面给出宏基站侧的基于静止通信体之间的无线信道的业务调度综合模型说明。

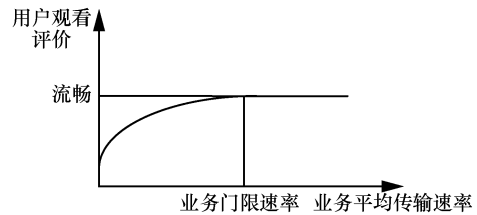


图 2 点播式视频业务用户观看评价和业务平均传输速率之间的关系

1) 宏基站需要支持  $M$  个静止微基站的下行非实时业务传输, 且考虑满负载业务模型, 即宏基站缓冲区始终有数据发送。

2)  $M$  个微基站业务期望速率分别为  $R_{\text{request}} = [R_1, R_2, \dots, R_M]$ , 单位为 bit/s。

3) 假设宏基站可以理想获取其自身和这  $M$  个微基站在数据传输信道上的信道状态信息, 且宏基站和任何一个微基站之间的无线信道状态虽是时变的, 但符合平稳随机过程的特征。

4) 信道调度机制的 3 个重要目标如下。

① 用户业务体验好。在下行容量允许范围内, 尽量多的微基站的业务期望速率获得满足。

② 信道利用效率高。各微基站的下行业务传输速率之和最大化。

③ 多用户公平性强。各用户对自身期望速率的满足水平趋同。

## 2.2 多用户离散信道状态概率空间及其事件域分割

在 OFDMA 系统中, 自适应调制编码技术可以基于不同的信道情况, 在各子载波上自适应地采用合适的调制编码方案, 从而在无线信道上获得最大的传输效率。设宏基站在一个时频调度资源块上可以支持  $N$  种不同速率的调制编码方案, 则对应的离散信道状态也为  $N$  个。基于无线信道的随机性,  $M$  个相互独立的随机变量  $\omega = \{\omega_1, \omega_2, \dots, \omega_M\}$  分别表示  $M$  个微基站在某一个子载波资源块上的信道状态, 且  $\omega_m, m = 1, \dots, M$  的取值范围为  $[1, N]$  中的正整数, 数值越大表示信道质量越好。

针对宏基站和  $M$  个微基站之间的多用户离散信道状态概率空间  $\Omega$  定义一个事件域集合, 如式(2)所示。

$$\{A_{n,m,k}, 1 \leq n \leq N, 1 \leq m \leq M, 1 \leq k \leq M - m + 1\} \quad (2)$$

其中, 事件域  $A_{n,m,k}$  表示如下 3 个条件同时获得满足的基本事件集合。

1)  $\max(\omega_1, \omega_2, \dots, \omega_{m-1}) < n$ 。

2)  $\omega_m = n$ 。

3) 有  $k$  个用户的信道状态值为  $n$ 。

即有

$$A_{m,n,k} = \left\{ \underbrace{\omega_1, \omega_2, \dots, \omega_{m-1}}_{<n}, \overbrace{\omega_m, \omega_{m+1}, \dots, \omega_M}^{=n, k-1 \text{ 个用户为 } n} \right\} \quad (3)$$

针对上述事件域  $A_{m,n,k}$  的定义, 给出一个基本结论, 如结论 1 所示。

**结论 1** 式(2)所定义的事件域集合  $\{A_{n,m,k}, 1 \leq n \leq N, 1 \leq m \leq M, 1 \leq k \leq M - m + 1\}$  对  $M$  个用户的离散信道状态概率空间  $\Omega$  进行了互不相容的事件域分割。

**证明** 首先证明集合中的任意 2 个不同的事件域一定互不相容, 即

$$\begin{aligned} & \text{if } (m_1, n_1, k_1) \neq (m_2, n_2, k_2), \\ & \text{then } A_{m_1, n_1, k_1} \cap A_{m_2, n_2, k_2} = \phi \end{aligned} \quad (4)$$

等价于证明

$$\text{if } \forall e \in A_{m_1, n_1, k_1}, \text{ then } e \notin A_{m_2, n_2, k_2} \quad (5)$$

其中,  $e$  表示多用户离散信道状态概率空间  $\Omega$  的一个基本事件, 即  $M$  个用户的信道状态值依次为  $e_1, e_2, \dots, e_M$ 。

$$e = [e_1, e_2, \dots, e_M], e_m = 1, \dots, N; m = 1, \dots, M$$

从式(3)对事件域定义可以看出, 如果某基本事件  $e \in A_{m_1, n_1, k_1}$ , 则  $e_1, e_2, \dots, e_M$  的数值同时满足如下 3 个条件。

- 1)  $\max(e_1, e_2, \dots, e_{m-1}) < n_1$ 。
- 2)  $e_{m_1} = n_1$ 。
- 3)  $e_1, e_2, \dots, e_M$  中有  $k_1$  个值为  $n_1$ 。

归纳上述 3 个条件可以看出, 在事件域  $A_{m,n,k}$  的 3 个下标中只要不少于一个下标的数值不同, 即  $(m_1, n_1, k_1) \neq (m_2, n_2, k_2)$ , 就可以证明如果  $\forall e \in A_{m_1, n_1, k_1}$ , 则  $e \notin A_{m_2, n_2, k_2}$ 。

下面证明事件域集合  $\{A_{n,m,k}, 1 \leq n \leq N, 1 \leq m \leq M, 1 \leq k \leq M - m + 1\}$  对概率空间  $\Omega$  进行了分割, 即

$$\Omega = \bigcup_{n=1}^N \bigcup_{m=1}^M \bigcup_{k=1}^{M-m+1} A_{n,m,k} \quad (6)$$

式(6)的证明等价于证明对于  $\forall e \in \Omega$ , 则  $\exists A_{m,n,k}$  使  $e \in A_{m,n,k}$ 。显然, 对于  $\forall e = [e_1, e_2, \dots, e_M] \in \Omega$ , 可以根据数值情况直接计算其所归属的事件域  $A_{m,n,k}$  的 3 个下标, 具体为

$$n = \max([e_1, e_2, \dots, e_M])$$

$$k = \{[e_1, e_2, \dots, e_M] \text{ 中数值为 } n \text{ 的用户数量}\}$$

$$m = \{[e_1, e_2, \dots, e_M] \text{ 中第一个数值为 } n \text{ 的用户编号}\}$$

即任意的一个基本事件, 都存在一个对应的其所归属的事件域。

综合式(4)和式(6)的证明, 结论 1 的内容成立。证毕。

下面给出一个具体案例来辅助说明事件域集合对概率空间的分割情况, 其中,  $M = 3, N = 3$ , 即概率空间  $\Omega$  含有  $3^3 = 27$  个基本事件, 则这些基本事件和事件域集合之间的归属情况如下。

$$\begin{aligned} A_{1,1,3} &= \{[1,1,1]\}; A_{1,2,1} = \{[2,1,1]\}; \\ A_{1,2,3} &= \{[2,2,2]\}; A_{3,2,1} = \{[1,1,2]\}; \\ A_{1,2,2} &= \{[2,2,1], [2,1,2]\}; \\ A_{2,2,1} &= \{[1,2,1]\}; A_{2,2,2} = \{[1,2,2]\}; \\ A_{1,3,3} &= \{[3,3,3]\}; A_{2,3,2} = \{[1,3,3], [2,3,3]\}; \\ A_{1,3,2} &= \{[3,3,1], [3,3,2], [3,1,3], [3,2,3]\}; \\ A_{1,3,1} &= \{[3,1,1], [3,1,2], [3,2,1], [3,2,2]\}; \\ A_{3,3,1} &= \{[1,1,3], [1,2,3], [2,1,3], [2,2,3]\}; \\ A_{2,3,1} &= \{[1,3,1], [1,3,2], [2,3,1], [2,3,2]\} \end{aligned}$$

### 2.3 事件域概率的计算

下面基于结论 1 的内容, 给出式(2)所定义的事件域集合  $\{A_{n,m,k}\}$  中每个事件域的概率值的具体计算过程和计算式。

1) 设  $M$  个用户离散信道状态概率空间  $\Omega$  的信道状态概率矩阵为

$$P = \begin{bmatrix} P(e_1 = 1) & P(e_1 = 2) & \dots & P(e_1 = N) \\ P(e_2 = 1) & P(e_2 = 2) & \dots & P(e_2 = N) \\ \vdots & \vdots & \ddots & \vdots \\ P(e_M = 1) & P(e_M = 2) & \dots & P(e_M = N) \end{bmatrix} \quad (7)$$

其中, 元素  $P(e_m = n)$  表示第  $m$  个用户的信道状态值为  $n$  的概率大小,  $m = 1, \dots, M, n = 1, \dots, N$ 。

2) 基于式(7)的概率分布矩阵, 分别建立如下  $N$  个阶数都为  $M$  的多项式。

$$\begin{aligned} f_1(x) &= (a_{1,1}x + b_{1,1}) \times (a_{2,1}x + b_{2,1}) \times \dots \times (a_{M,1}x + b_{M,1}) \\ f_2(x) &= (a_{1,2}x + b_{1,2}) \times (a_{2,2}x + b_{2,2}) \times \dots \times (a_{M,2}x + b_{M,2}) \\ &\vdots \\ f_N(x) &= (a_{1,N}x + b_{1,N}) \times (a_{2,N}x + b_{2,N}) \times \dots \times (a_{M,N}x + b_{M,N}) \end{aligned} \quad (8)$$

其中, 多项式系数定义为

$$a_{m,n} = P(e_m = n), 1 \leq m \leq M, 1 \leq n \leq N$$

$$b_{m,n} = \begin{cases} 0, 1 \leq m \leq M, n = 1 \\ \sum_{l=1}^{n-1} P(e_m = l), 1 \leq m \leq M, 2 \leq n \leq N \end{cases} \quad (9)$$

由于  $M$  个用户的信道状态是相互独立的, 则事件域  $A_{m,n,k}$  的发生概率  $P(A_{m,n,k})$  的计算式为

$$P(A_{m,n,k}) = \prod_{r=1}^{m-1} P(e_r < n) P(e_m = n) \cdot P(\text{后}(M-m)\text{个用户状态值中有}(k-1)\text{个达到最大值}n) = \begin{cases} a_{1,1} a_{2,1} \cdots a_{M,1}, n = 1 \\ b_{1,n} b_{2,n} \cdots b_{m-1,n} a_{m,n} ((a_{m+1,n}, b_{m+1,n}) \otimes (a_{m+2,n}, b_{m+2,n}) \otimes \cdots \otimes (a_{M,n}, b_{M,n}))_{k-1}, n > 1 \end{cases} \quad (10)$$

其中,  $((a_{m+1,n}, b_{m+1,n}) \otimes (a_{m+2,n}, b_{m+2,n}) \otimes \cdots \otimes (a_{M,n}, b_{M,n}))_{k-1}$  表示式(8)中挑选出的部分多项式系数进行卷积计算后的阶数为  $(k-1)$  的系数数值, 即令

$$f(x) = (a_{m+1,n}x + b_{m+1,n}) \times (a_{m+2,n}x + b_{m+2,n}) \times \cdots \times (a_{M,n}x + b_{M,n}) = c_{M-m}x^{M-m} + c_{M-m-1}x^{M-m-1} + \cdots + c_{k-1}x^{k-1} + \cdots + c_1x^1 + c_0 \quad (11)$$

则

$$c_{k-1} = ((a_{m+1,n}, b_{m+1,n}) \otimes (a_{m+2,n}, b_{m+2,n}) \otimes \cdots \otimes (a_{M,n}, b_{M,n}))_{k-1} \quad (12)$$

## 2.4 预测性信道调度算法

宏基站和微基站之间的物理层基本参数和业务调度需求说明如下。

1) 系统采用 OFDMA 体制, 但不考虑使用信道功率分配机制, 并以基本时频资源块为单位进行信道调度, 其中资源块的时间长度为  $T$ , 带宽为  $B$ , 数据传输信道含  $L$  个资源块, 每个资源块在一次调度中只能承载一个微基站的下行业务数据。

2) 数据传输信道共支持  $N$  种不同的调度编码方案, 各方案的单位频谱利用率依次为  $r = [r_1, r_2, \cdots, r_N]$ , 单位为 bit/Hz。

3)  $M$  个微基站的业务期望速率分别为  $R_{\text{request}} = [R_1, R_2, \cdots, R_M]$ , 单位为 bit/s。

传统调度算法的基本思想是基于一个当前权重值的调度过程, 属于短时调度机制, 无法从长时角度更充分地利用各用户信道状态的随机抖动。而基于上述物理层模型和业务调度需求, 本文的调度机制的核心思想为: 由于各用户的信道状态值符合平稳随机条件, 因此基于信道先验统计信息并依照

2.2 节和 2.3 节的内容进行事件域的切割和发生概率值计算, 然后基于各个用户的业务期望速率值来计算出一个匹配的长时信道调度策略, 最终实现统计意义上的理论调度速率和业务期望速率相同, 从而分别满足各用户的业务调度需求。具体的信道调度机制如下。

### 2.4.1 业务期望速率和信道状况的预适配

本节基于当前信道综合状况, 对待调度的业务期望速率合理修正, 从而达到期望速率和信道状况的预适配。具体机制如下。

由于信道变化为平稳随机过程, 因此基于历史数据, 基站在开始进行信道调度前已经获得了这  $M$  个微基站在数据传输信道上的  $L$  个基本时频资源块的信道统计信息矩阵, 具体如式(13)所示。

$$P_l = \begin{bmatrix} P(e_{l,1} = 1) & P(e_{l,1} = 2) & \cdots & P(e_{l,1} = N) \\ P(e_{l,2} = 1) & P(e_{l,2} = 2) & \cdots & P(e_{l,2} = N) \\ \vdots & \vdots & \ddots & \vdots \\ P(e_{l,M} = 1) & P(e_{l,M} = 2) & \cdots & P(e_{l,M} = N) \end{bmatrix} \quad (13)$$

其中, 元素  $P(e_{l,m} = n)$  表示在第  $l$  个基本时频资源块上第  $m$  个用户的信道状态值为  $n$  的概率大小,  $l = 1, \cdots, L, m = 1, \cdots, M, n = 1, \cdots, N$ 。

如果调度过程中总是把基本时频资源块分配给某一个信道状态值最好的用户, 则实际的业务吞吐量等于信道容量。考虑到在下行用户数达到一定数量且在地理上满足均匀分布时, 在一些部署较为密集的环境下, 任意时刻所有用户在每个基本时频资源块上的信道状态最大值为  $N$  的概率将趋于 1, 因此, 本文将基站的下行数据信道吞吐量上参考值设定为

$$\bar{C} = 0.95BLr_N \quad (14)$$

如果此时  $M$  个微基站的下行业务期望速率值之和  $R_{\text{all}}$  大于式(14)所定义的信道吞吐量上限参考值的 0.95, 即

$$R_{\text{all}} = \sum_{m=1}^M R_m > 0.95\bar{C} \quad (15)$$

则说明业务的整体期望速率值过大, 基站很有可能无法同时满足这些业务的服务质量要求, 因此将启用业务控制措施, 根据事先规定的业务优先级来逐步降低一些低等级业务的期望速率, 甚至放弃部分业务, 直到所有业务的期望速率值之和降低至信道容量理论值的 0.95 左右。

如果此时  $M$  个微基站的下行业务期望速率值之和  $R_{\text{all}}$  小于式(14)所定义的信道吞吐量上限参考值的 0.8, 即

$$R_{\text{all}} = \sum_{m=1}^M R_m < 0.8\bar{C} \quad (16)$$

则说明业务的整体期望速率值过小, 很有可能造成下行链路资源浪费现象。此时将启用业务控制措施, 一定程度上提高那些速率非受限业务(如 FTP 下载业务)的期望速率, 直到业务期望速率之和提高至信道容量理论值的 0.9 左右。

本文研究重点在于信道调度机制, 因此这里不讨论业务控制措施的具体细节。不失一般性地, 设经过上述业务控制措施后,  $M$  个微基站的业务期望速率值调整为  $R_{\text{request}} = [\bar{R}_1, \bar{R}_2, \dots, \bar{R}_M]$ 。

### 2.4.2 信道调度

完成 2.4.1 节的业务期望速率和信道状况的预适配后, 本节开始进行信道调度, 具体目标是使  $M$  个微基站中尽可能多的微基站分别获得和期望值趋同的实际业务传输速率。具体机制如下。

1) 在决策某次  $L$  个基本时频资源块的分配结果时, 首先统计  $M$  个微基站在本次调度前的先验信息区(1 000 次, 如图 3 所示)中信道状态值情况  $\bar{P} = [\bar{P}_1, \bar{P}_2, \dots, \bar{P}_L]$ , 具体如式(17)所示。

$$\bar{P}_l = \begin{bmatrix} \bar{P}(e_{l,1} = 1) & \bar{P}(e_{l,1} = 2) & \dots & \bar{P}(e_{l,1} = N) \\ \bar{P}(e_{l,2} = 1) & \bar{P}(e_{l,2} = 2) & \dots & \bar{P}(e_{l,2} = N) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{P}(e_{l,M} = 1) & \bar{P}(e_{l,M} = 2) & \dots & \bar{P}(e_{l,M} = N) \end{bmatrix} \quad (17)$$

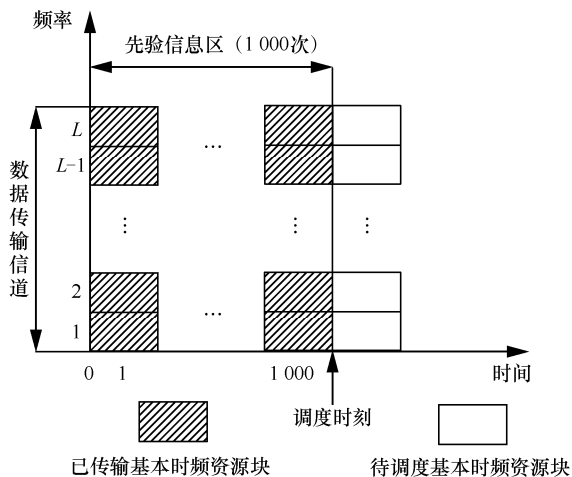


图 3 基于先验信息的调度过程

其中, 元素  $\bar{P}(e_{l,m} = n)$  表示在第  $l$  个基本时频资源块上第  $m$  个用户在前 1 000 次调度的信道状态值为  $n$  的比例值,  $l = 1, \dots, L, m = 1, \dots, M, n = 1, \dots, N$ 。

2) 对  $M$  个用户在基本时频资源块上的优先级进行排序, 并根据其优先级制定信道分配策略。

$$\theta_{l,m,n} = P(e_{l,m} = n) \frac{\prod_{k=1}^{n-1} P(e_{l,k} = r)}{\sum_{r=1}^{n-1} P(e_{l,m} = r)} \quad (18)$$

其中,  $\theta_{l,m,n}$  表示在第一个时频资源块上第  $m$  个用户的信道状态为  $n$ , 且其余用户的信道状态值都小于该值的概率大小;  $m = 1, \dots, M, n = 2, \dots, N$ 。

$$\lambda_{l,m} = \frac{\sum_{n=1}^N \theta_{l,m,n} r_n B}{TR_m} \quad (19)$$

其中,  $\lambda_{l,m}$  表示在第一个时频资源块上第  $m$  个用户在自己信道状态值大于其他所有用户情况下, 即可使用该资源块进行数据传输时, 所获得的调度速率和其业务期望速率的比值。 $\lambda_{l,m}$  的值越大, 一定程度上说明第  $m$  个用户在第一个时频资源块上信道条件越好或其业务期望速率越低, 因此满足其业务调度要求的难度越低。

基于  $\lambda_{l,m}$  的大小对用户进行优先级排序, 数值越小(调度难度越高)的用户优先级越高, 根据排序结果对式(17)中的第一个矩阵  $\bar{P}_1$  进行位置调整, 调整后矩阵如式(20)所示。

$$\bar{P}_1 = \begin{bmatrix} \bar{P}(e_{1,1} = 1) & \bar{P}(e_{1,1} = 2) & \dots & \bar{P}(e_{1,1} = N) \\ \bar{P}(e_{1,2} = 1) & \bar{P}(e_{1,2} = 2) & \dots & \bar{P}(e_{1,2} = N) \\ \vdots & \vdots & \ddots & \vdots \\ \bar{P}(e_{1,M} = 1) & \bar{P}(e_{1,M} = 2) & \dots & \bar{P}(e_{1,M} = N) \end{bmatrix} \quad (20)$$

3) 如果在调度过程中, 时频资源块总是承载信道状态值最好的用户的业务, 则实际的业务吞吐量等于信道容量。因此, 根据  $\lambda_{l,m}$  的优先级排序结果调度机制为:  $M$  个用户中信道状态值达到最大值, 且优先级最高的唯一用户将获得第一个时频资源块的使用权。

上述调度机制综合考虑了时频资源利用效率、用户信道质量和用户期望速率这 3 个因素, 在同等信道状况下让调度难度较高的用户优先获得信道使用权, 从而获得了均衡调度的效果。下面利用 2.3 节所

提供的事件域概率式计算  $M$  个用户在第一个时频资源块上所获得的概率意义上的调度速率  $V_{1,m}$ 。

首先，参考式(3)定义事件域  $\bar{A}_{1,m,n,k} = \{\text{在第一个资源块上, 前 } (m-1) \text{ 个用户状态值小于 } n, \text{ 第 } m \text{ 个用户状态值为 } n, \text{ 其后面 } (M-m) \text{ 个用户中共 } k \text{ 个用户状态值为 } n\}$ 。

然后，基于式(20)，令

$$\begin{aligned} a_{1,m,n} &= \bar{P}(e_{1,m} = n), 1 \leq m \leq M, 1 \leq n \leq N \\ b_{1,m,n} &= \begin{cases} 0, 1 \leq m \leq M, n = 1 \\ \prod_{l=1}^{n-1} \bar{P}(e_{1,m} = l), 1 \leq m \leq M, 2 \leq n \leq N \end{cases} \quad (21) \end{aligned}$$

基于式(10)，则事件域  $\bar{A}_{1,m,n,k}$  的发生概率  $P(\bar{A}_{1,m,n,k})$  为

$$\begin{aligned} P(\bar{A}_{1,m,n,k}) &= \begin{cases} a_{1,1,1} a_{1,2,1} \cdots a_{1,M,1}, n = 1 \\ b_{1,1,n} b_{1,2,n} \cdots b_{1,m-1,n} a_{1,m,n} ((a_{1,m+1,n}, b_{1,m+1,n}) \otimes \\ (a_{1,m+2,n}, b_{1,m+2,n}) \otimes \cdots \otimes (a_{1,M,n}, b_{1,M,n}))_{k-1}, n > 1 \end{cases} \quad (22) \end{aligned}$$

如果归属于事件域  $\bar{A}_{1,m,n,k}$  中的某个基本事件出现，则根据调度机制可知第  $m$  个用户获得了第一个资源块的使用权，因此第  $m$  个用户在第一个资源块上所获得的概率意义上的调度速率为

$$V_{1,m} = \frac{B}{T} \sum_{n=1}^N r_n P(\bar{A}_{1,m,n,k}) \quad (23)$$

4) 对于第二个资源块，如果有用户在第一个资源块上所获得的概率意义上的调度速率不小于其调整后的业务期望速率，则认为这些用户的业务期望速率已经获得满足，放弃对第二个资源块以及后续的资源块的竞争。

如果所有用户在第一个资源块上所获得的概率意义上的调度速率都小于其调整后的业务期望速率，则基于式(17)中的  $\bar{\mathbf{P}}_2$  中的元素进行计算，如式(24)所示。

$$\theta_{2,m,n} = P(e_{2,m} = n) \frac{\prod_{k=1}^M \sum_{r=1}^{n-1} P(e_{2,k} = r)}{\sum_{r=1}^{n-1} P(e_{2,m} = r)} \quad (24)$$

其中， $\theta_{2,m,n}$  表示在第二个时频资源块上第  $m$  个用户的信道状态为  $n$ ，且其余用户的信道状态值都小

于该值的概率大小； $m = 1, \dots, M, n = 2, \dots, N$ 。

$$\lambda_{2,m} = \frac{\sum_{n=1}^N \theta_{2,m,n} r_n B}{T \left( \bar{R}_m - \sum_{p=1}^1 \bar{V}_{p,m} \right)} \quad (25)$$

参考  $\lambda_{1,m}$  的定义可知， $\lambda_{2,m}$  表示在第二个时频资源块上第  $m$  个用户在自己信道状态值大于其他所有用户情况下，即可使用该资源块进行数据传输时，所获得的调度速率和其剩余业务期望速率的比值，其中剩余业务期望速率指业务期望速率减去在前面资源块上所获得的调度速率值之和。同理， $\lambda_{2,m}$  的值越大，一定程度上说明第  $m$  个用户在第二个时频资源块上信道条件越好或其剩余业务期望速率越低，因此满足其业务调度要求的难度越低。

与在第一个时频资源块上的方法相同，基于  $\lambda_{2,m}$  的大小对用户进行优先级排序，数值越小的用户优先级越高，根据排序结果对式(17)中的第二个矩阵  $\bar{\mathbf{P}}_2$  进行位置调整，且在第二个时频资源块上的调度机制同样为： $M$  个用户中信道状态值达到最大值，且优先级最高的唯一用户将获得第二个时频资源块的使用权。然后，采用和第一个时频资源块相同的方法，即式(23)计算在这种调度规则下第  $m$  个用户在第二个资源块上所获得的概率意义上的调度速率  $V_{2,m}$ 。

5) 对于第三个资源块及其后续资源块而言，如果有用户在前面资源块上所获得的概率意义上的调度速率之和不小于其调整后的业务期望速率，则认为这些用户的业务期望速率已经获得满足，从而退出当前及后续的资源块的调度竞争；而对于那些剩余业务期望速率大于 0 的用户，继续进行优先级的计算和排序工作，并将该资源块分配给信道状态值达到最大值且优先级最高的用户。持续资源块的分配过程，直到所有用户的要求速率都获得满足或者所有的资源块都分配完毕。

6) 在资源块调度过程中，如果某一个用户在前面所有资源块上所获得的概率意义上的调度速率之和尚未超过其业务期望速率，但加入当前资源块后就超过其业务期望速率的 3% 时，为避免该用户出现资源过度调度，设超过部分的比例为  $x$ ，当该用户在当前资源块获得调度机会后，产生一个在区间  $[0,1]$  均匀分布的随机数  $\alpha$ ，如果  $\alpha \leq \frac{1}{1+x}$ ，则该

用户获得该资源块的使用权；如果  $\alpha > \frac{1}{1+x}$ ，则调度机会由其他用户公平竞争。

### 2.5 信道调度算法的性能分析

由于信道调度算法是一个多维优化问题，不存在性能最优方案，因此很难用单一指标来定性衡量算法的性能高低，需要从算法收敛性、复杂度、资源利用效率、调度公平性等多个方面来均衡分析。

1) 算法收敛性。本文算法不是采用传统优化问题的迭代求解机制，而是一种基于信道统计特征的一种可量化的调度机制，在无线信道满足平稳随机性的前提下，调度结果是可以明确计算的，因此不存在不收敛的风险。

2) 算法复杂度。综合分析上述调度机制可知，本文算法的调度过程为，调度需求产生—调度策略计算—调度策略执行—调度新情况产生—新调度策略计算—…。因此宏基站侧在一次调度需求产生后，仅进行一次调度策略计算后就不再需要计算，后续较长时间的调度过程仅需要依照计算结果执行即可，直至当前有业务结束传输需求或有新的传输需求进入。

在第一次的初始调度过程中，分析整个调度过程可知，算法计算量主要集中在对式(22)的事件域发生概率值的计算过程中，其中事件域数量一共为  $LN \frac{M(M+1)}{2}$  个，而每一个事件域发生概率值  $P(A_{i,m,n,k})$  的计算需要约  $m + (k-2)C_{M-m}^{k-1}$  次的实数乘法；当有业务结束传输或有新传输需求进入时，需要重新进行调度策略计算。综合分析式(22)的计算过程可知，不管是当前业务传输结束或新用户业务进入，新的计算结果和上一次计算结果都存在广义线性递归关系。虽然第一次初始调度过程的计算复杂度比较高，但是在后续调度过程中，宏基站不需要重复进行初始调度时的复杂计算过程，可在上一次计算结果的基础上采用类似于递归计算的方式，通过更换计算式中的少量元素即可完成对新调度策略的计算，即使系统频繁产生新业务传输需求，调度策略的更新所需的计算量是几乎可忽略不计的，因此调度算法的整体复杂度较低。

3) 资源利用效率。本文的调度机制是一种新型的长时调度机制，可以在更长的时间段中更充分地利用无线信道在时频二维上的时变性，可以尽可能地将信道资源的使用权分配给众多用户中信道状

态最好的用户，因此资源利用效率高，信道吞吐量可以逼近在这个指标上的性能上限，即基于最大信噪比的调度算法。

4) 调度公平性。传统狭义的调度公平性是指各用户调度速率的趋同，但本文采用多用户多速率的业务模型，因此采用广义概念的公平性，具体为用户期望速率满足程度的趋同（调度速率和期望速率的比值）。本文算法是一种以满足用户期望速率为导引目标的精确量化调度算法，可以证明在信道统计特征稳定的前提下，各用户可以获得概率意义上的业务调度速率和业务期望速率的匹配，基本满足整个系统对调度公平性的要求。

## 3 调度算法性能仿真

为验证上述调度机制的信道资源分配效果，本文利用如下模型进行性能仿真。设一个下行数据传输信道包含  $L = 40$  个基本时频资源块的宏基站，需要同时支持 80 个均匀分布在小区内的静止微基站的下行业务需求。系统在每个调度时隙上都可以根据用户的理想反馈采用表 1 中的调制编码方案来获得最大的传输速率。将数据传输信道带宽和传输时间归一化后，宏基站分别采用这 7 个调制编码方案的下行数据速率，依次为  $\frac{1}{4}$ 、 $\frac{1}{2}$ 、1、 $\frac{3}{2}$ 、2、3 和 4。

表 1 数据传输的调制编码方案

调制方案	编码码率	频率利用率/(bit·Hz <sup>-1</sup> )	要求信噪比/dB
BPSK	$\frac{1}{4}$	$\frac{1}{4}$	-2.9
BPSK	$\frac{1}{2}$	$\frac{1}{2}$	-0.2
QPSK	$\frac{1}{2}$	1	2.2
8PSK	$\frac{1}{2}$	$\frac{3}{2}$	5.2
8PSK	$\frac{2}{3}$	2	8.4
64QAM	$\frac{1}{2}$	3	11.8
64QAM	$\frac{3}{2}$	4	15.1

考虑到静态通信体之间的无线信道满足平稳随机过程的条件，即 80 个微基站在 40 个基本时频

资源块上的信道状态值的统计值保持稳定。为避免其他因素对调度算法性能的影响，宏基站采用发送功率均匀分配机制，且在每个资源块上的下行发射功率都为 48，仿真初始化时 80 个用户的大尺度衰落系数在区间 $[-36, -30]$ 上满足均匀分布，而小尺度衰落系数则均为满足均值为 1 的瑞利衰落分布。由于本文仅利用了无线信道的统计平稳特性，为简化仿真模型，在不影响仿真结果的情况下，本文忽略信道在频域上的相关性，假设不同资源块之间的信道状态是相互独立的。仿真时间持续了 3 000 个调度时隙，其中，前 1 000 个时隙不产生任何业务传输需求，用于宏基站积累和各微基站之间的信道状态值的历史数据；后 2 000 个时隙则正式开始进行业务的调度和传输过程。

目前的一些经典信道分配算法都是基于短时信息进行信道分配，其中最大信噪比算法以当前用户信道质量为唯一考虑因素，可以达到吞吐量的上限值，可用于新算法吞吐量性能的比较；而其他一些改进型比例公平算法则是对当前用户信道质量、用户当前业务速率、业务时延等因素进行加权计算来进行信道的分配，但无法实现多用户不同业务速率的自适应匹配调度。因此本文采用一种基于业务期望速率满足度的改进型比例公平算法作为调度算法性能的对比对象，具体为累积平均速率和其业务期望速率之间的比值，如式(26)所示。

$$\text{priority}_{m,l} = \arg \max \frac{U_{m,l}(t)}{\Delta_m(t)}, m=1, \dots, M, l=1, \dots, L \quad (26)$$

其中， $U_{m,l}(t)$  表示  $t$  时刻第  $m$  个用户在第  $l$  个基本时频资源块上的信道速率， $\Delta_m(t)$  表示  $t$  时刻第  $m$  个用户的期望速率的满足度。

本文提供的信道调度算法的主要优点就是实现量化调度过程，可以在尽可能地最大化下行吞吐量的同时，自适应地满足不同用户的不同的业务速率要求。假设 80 个微基站中共存在 4 种不同的下行非实时业务，业务期望速率比值为 1:2:4:8，分别记为 A、B、C 和 D，这 4 种业务的用户数量分别为 30、30、10 和 10 个。基于式(14)可知，宏基站的下行数据信道吞吐量上参考值为  $\bar{C} = 0.95 \times 4 = 3.8$ 。基于式(15)的要求，将这 80 个微基站的下行业务期望速率之和设定为

$$R_{\text{all}} = \sum_{m=1}^{80} R_m = 0.94 \bar{C} = 3.57, \text{ 则归一化后 A、B、C}$$

和 D 类业务对应的业务期望速率依次为 0.017、0.034、0.068 和 0.136。

下面基于仿真结果，分别从信道资源利用效率、调度公平性和用户业务体验的角度比较本文算法和式(26)的改进型比例公平算法的性能。首先，图 4 分别给出了宏基站使用最大信噪比算法、本文算法和改进型比例公平算法在整个调度过程中平均下行吞吐量情况(每 20 个调度时隙求一次平均)，对比 3 条曲线可以看出，本文算法的吞吐量性能逼近最大信噪比算法(两者相差约 3%)，比改进型比例公平算法的吞吐量性能提升了约 14%，表明本文算法的信道资源利用效率已经趋于性能上限值。

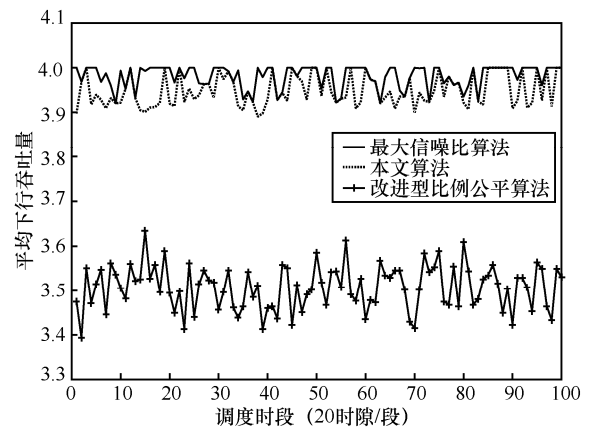


图 4 3 种算法在仿真过程中的系统平均下行吞吐量

在仿真过程中，宏基站一共给出了  $40 \times 2\,000 = 8 \times 10^4$  次的基本时频块的调度结果，图 5 进一步给出了调度过程中本文算法和改进型比例公平算法的基本时频资源块的利用情况统计(最大信噪比算法全部为最优利用，因此这里不再进行比较)。“最优”表示将该资源块分配了所有用户中信道质量达到最大值中的某个用户，“次优”表示分配给了比最大值小一个档次的用户，“其他”表示其他调度情况。从图 5 可以看出，本文算法可以从更长时间维度利用信道时变性，因此，“最优”“次优”和“其他”调度情况的比例值分别为 94%、4%和 2%，下行吞吐量逼近上限值，即采用基于最大信噪比的调度算法所获得的吞吐量；而比例公平算法属于短时调度机制，资源使用权有更大的概率被授予那些信道状态不是最好甚至较差的用户，因此 3 种调度情况的比例值依次为 82%、11%和 7%，这也是其平均下行吞吐量

差于本文调度算法的原因。

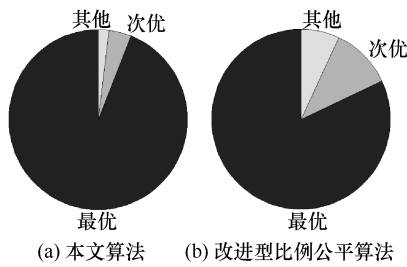


图 5 本文算法和改进型比例公平算法在仿真过程中时频资源块的利用情况统计

图 6 给出了 2 种算法仿真过程中 80 个微基站共 4 种业务 (A、B、C、D) 分别获得的平均调度速率。从图 6 可以看出, 2 种算法的调度目标都以满足用户期望速率为导向, 因此不同用户的信道情况和业务需求虽然存在差异, 但从广义公平性的角度出发, 不同用户在相同业务上所获得的实际调度速率差异性较小, 所有用户的期望速率的满足度 (累积平均速率和期望速率的比值) 都是基本趋同的, 因此 2 种算法在广义公平性的性能是基本相当的。但是从最终调度效果来看, 由于本文算法对信道资源的利用效率更高, 因此在各业务期望速率之和到达系统下行吞吐量 94% 的高业务负荷情况下, 改进型比例公平算法中各用户的期望速率的满足度水平较差, 大部分用户仅有约 85%; 而本文算法的满足度水平较高, 大部分用户达到了约 98%, 即大部分用户都获得了和其业务期望速率相当的调度速率。因此, 本文算法在重业务负荷场景下对各业务的期望速率的满足能力明显优于改进型比例公平算法。

### 4 结束语

本文针对宏基站和微基站之间下行非实时业务, 利用静态通信体之间的无线信道的平稳随机性, 从更长的时间维度上研究信道调度算法, 从离散概率空间的角度精确计算各事件域发生的概率大小, 据此利用各个用户的信道状态统计值进行预测性和可量化的调度过程, 比起传统的比例公平类的短时调度类算法, 这种长时调度算法可以更加充分地利用多用户的频率和时间分集所带来的调度增益, 在业务吞吐量逼近系统上限值的同时, 满足了多用户多业务的复杂调度需求, 实现了信道资源的高效利用和自适应按需分配。

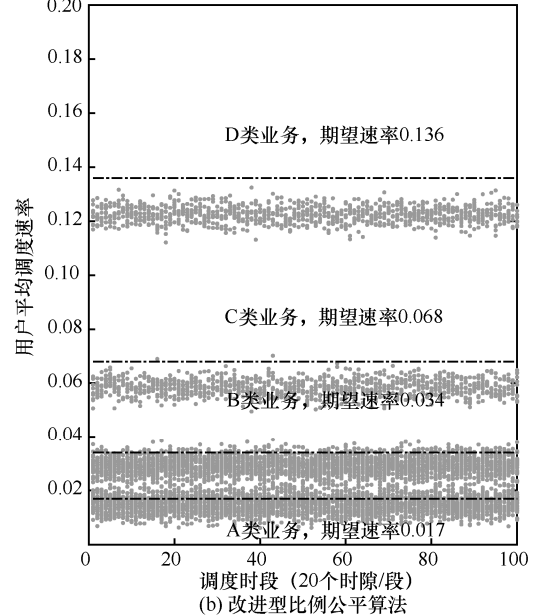
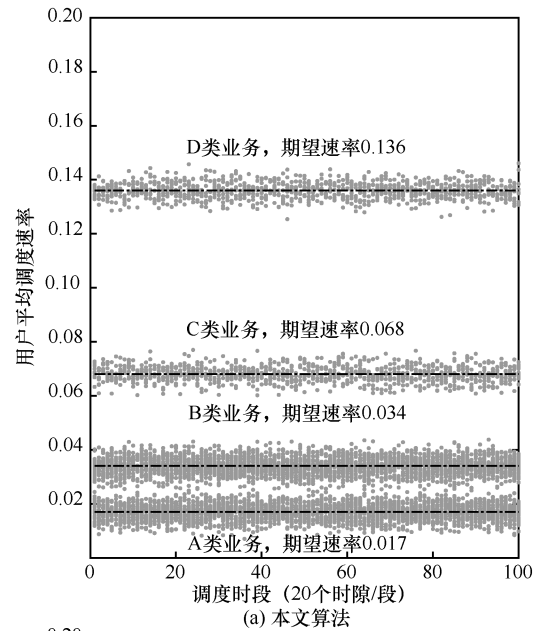


图 6 本文算法和改进型比例公平算法的 4 种业务在不同时段中所获得的平均调度速率

### 参考文献:

[1] YOU X H, PAN Z W, GAO X Q, et al. The 5G mobile communication: the development trends and its emerging key techniques[J]. Scientia Sinica Informationis, 2014, 44(5): 551-563.

[2] 程日涛, 张海涛, 王乐. 5G 无线网部署策略[J]. 电信科学, 2018, 34(S1): 1-8.

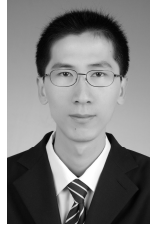
[3] CHENG R T, ZHANG H T, WANG L. Deployment strategy of 5G wireless network[J]. Telecommunications Science, 2018, 34(S1): 1-8.

王祖阳, 杨传祥, 张进, 等. 5G 无线网技术特征及部署应对策略分析[J]. 电信科学, 2018, 34(S1): 9-16.

WANG Z Y, YANG C X, ZHANG J, et al. Analysis on technology characteristics and deployment strategies of 5G wireless network[J]. Telecommunications Science, 2018, 34(S1): 9-16.

- [4] JALALI A. Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system[C]// IEEE Vehicular Technology Conference. IEEE, 2000: 1854-1858.
- [5] WESTPHAL C. Monitoring proportional fairness in cdma2000 high data rate networks[C]//IEEE Global Telecommunications Conference. IEEE, 2004: 3866-3871.
- [6] HOLTZMAN J M. Asymptotic analysis of proportional fair algorithm[J]. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2001,2: 33-37.
- [7] ZHANG X . Study on delay of proportional fair algorithm with controlled channel quality feedback[C]//International Conference on Communication Technology. IEEE, 2003:966-969.
- [8] Cisco. Cisco visual networking index: global mobile data traffic forecast update(2016-2021)[R]. (2019-01-18)[2019-03-14].
- [9] HAN D S, ZHENG B, CHEN Z S, et al. Cost efficiency in coordinated multiple-point system based on multi-source power supply[J]. IEEE Access, 2018, 6(1): 71994-72001.
- [10] YAO C, YANG C, XIONG Z. Energy-saving predictive resource planning and allocation[J]. IEEE Transactions on Communications, 2016, 64:5078-5095.
- [11] ABOU-ZEID H, HASSANEIN H. Predictive green wireless access: exploiting mobility and application information[J].IEEE Wireless Communications, 2013, 20: 92-99.
- [12] YAO C, YANG C. Role of large scale channel information on predictive resource allocation [C]//The IEEE Wireless Communications and Networking Conference. IEEE, 2016: 1-6.
- [13] HU Y, HAN S, YANG C. Context-aware energy saving with proactive power allocation[C]//The IEEE Global Conference on Signal and Information Processing. IEEE, 2015: 53-57.
- [14] SHE C, YANG C. Context aware energy efficient optimization for video on-demand service over wireless networks[C]//The IEEE/CIC International Conference on Communications. IEEE, 2015: 1-6.
- [15] SHE C, YANG C. Energy efficient predictive resource allocation for VoD and real-time services[J]. IEEE Wireless Communications, 2017(7): 111-115.

## [作者简介]



谢映海（1983- ），男，福建仙游人，博士，珠海中慧微电子博士后工作站在站博士后，主要研究方向为数字宽带通信。



姚若河（1961- ），男，博士，华南理工大学教授、博士生导师，主要研究方向为电子器件及应用。



吴斌（1971- ），男，广东深圳人，博士，珠海中慧微电子博士后工作站总工程师，主要研究方向为智能物联网技术。