

## MIMO 广播信道中联合多用户速率和延迟约束调度的研究

杨涛 胡波

(复旦大学电子工程系 上海 200433)

**摘要:** 该文给出了一种多入多出高斯广播信道中具有延迟约束的多用户无线数据包的传输方案。首先,在信道的QR分解及脏纸编码基础上,通过贪婪算法获取多用户分集与包延迟约束之间的有效结合。其次,将包到达速率、可达服务速率、用户数以及传输天线数构成一优化问题,得到最佳的用户组合及调度周期。最后,在不同用户数及发射天线数情况下对该方案进行了性能仿真,仿真结果表明:在满足包传递最小延迟等待要求的同时,得到了传输容量的最大化。

**关键词:** 多入多出; 高斯广播信道; 脏纸编码

中图分类号: TN934

文献标识码: A

文章编号: 1009-5896(2007)10-2504-04

## Joint Multiuser Rate and Delay Constraint Scheduling in MIMO Broadcast Channel

Yang Tao Hu Bo

(Department of Electronics Engineering, Fudan University, Shanghai 200433, China)

**Abstract:** In this paper a wireless packet transmission scheme over Multiple Input Multiple Output (MIMO) Gaussian Broadcast Channel (GBC) is proposed. Specifically, the transmission packet delay constraint and broadcast channel capacity are considered jointly, based on QR decomposition of channel and Dirty Paper Coding (DPC), a greedy algorithm is used to exploit multiuser diversity as well as transmission packet delay constraint in base station. It demonstrates that the packet arrival rate, achievable service rate, user number and transmitting antenna number can be formulated into an optimization problem. The simulation results show that the proposed scheme can maximize the system throughput, while providing guarantees on delay performance.

**Key words:** Multiple Input Multiple Output (MIMO); Gaussian Broadcast Channel (GBC); Dirty Paper Coding (DPC)

### 1 引言

多入多出(MIMO)天线的无线通信技术可有效提高频谱利用效率,因而在过去近十年中受到了学术界和产业界的广泛关注。其中,大量的研究是从信息论角度出发,研究如何使传输容量得到有效提高,而对用户传输的服务质量(QoS)关注较少。实际上,基于数据包的无线传输网络中,信源数据包通常以随机方式到达发射端,而发射端发送至移动接收端的数据包,常常会受到信道衰落及其他用户的干扰。在此情形下,若要获取较好的传输性能,需同时兼顾数据包的吞吐量和传输延迟两个要素。常规的网络定义中,包延迟性能定义在网络层,而包的吞吐量则定义在物理层,两者的研究通常是相互独立进行的<sup>[1]</sup>。为了得到实际多用户通信的理想性能,将延迟约束与吞吐量进行联合考虑即形成跨层分析则必不可少。

目前,联合容量和延迟约束调度问题的研究主要集中在

上行链路,即多址信道(MAC)的研究<sup>[1-3]</sup>,这是由于理论MAC容量问题已基本解决,将其与网络传输结合可以为实际传输性能提供准确的理论预期。而对于下行链路,即广播信道(BC),多传输天线的容量问题仍然未得到满意的解决。最近的研究表明<sup>[4]</sup>: BC容量可以通过脏纸编码(DPC)方案来获取,这为通过各种各样的信号处理技术求解容量问题,从而改善传输性能,开辟了一个新的思路。近期的一些典型研究包括:文献[5]给出了在多用户系统中,通过迫零(ZF)及空间预滤波构造的DPC传输方案达到最大可能获取的传输数率;延迟截至有限容量的定义则在文献[6]中给出,通过一个混合的延迟约束和非延迟约束服务模型来实施传输优化;文献[7]中,通过实施联合多用户分集和ZF-DPC方案近似达到了传输容量的Sato界。

本文在文献[5]和文献[8]的基础上,根据当前信道条件和发送端缓冲长度的约束,提出了新的调度模型。首先,根据等待时间对各用户数据包进行优先级排队,等待时间最长的数据包具有最高优先调度权,然后通过贪婪搜索过程为待调度的包找到最好的发送信道,由于数据包的延迟等待时间具

2006-03-27 收到, 2006-09-04 改回

航天863计划(2002AA742051)和上海市博士后科学基金(05R214110)资助课题

有不同的约束值, 故通过构建一优化模型, 进而得到每次调度的最佳周期, 最终取得容量最大化与延迟约束之间的有效折衷。

## 2 系统模型

考虑一  $M \times N$  的 MIMO-BC 多用户速率调度模型, 其中  $M, N$  分别为发送及接收天线个数, 满足  $M < N$ 。假定每个移动用户只有一个接收天线, 且各个接收天线上的高斯白噪声相互独立, 信道满足无记忆、准静态, 发送和接收端可以得到信道状态信息(CSI)。则接收模型可表示如下

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \quad (1)$$

其中  $\mathbf{H} = (h_{ij})$  为  $M \times N$  复矩阵中第  $i$  个发送天线与第  $j$  个接收天线之间的信道系数,  $N \times 1$  向量  $\mathbf{y}$  和  $\mathbf{n}$  分别为接收信号和循环对称的复高斯噪声, 且  $\mathbf{n}$  具有归一化的协方差矩阵。 $N \times 1$  向量  $\mathbf{x} = \sum_{i=1}^N x_i$  为发射符号, 其功率约束为  $P$ , 即有  $E(\text{Tr}(\mathbf{x}\mathbf{x}^T)) \leq P$ 。最后, 假设物理层上存在最优的编码, 这意味着一旦用户数据包被调度, 理论可达的和速率能够得到。

## 3 联合调度算法分析

对于一个基于缓存的包传输模型, 由于包服务速率和包到达速率的不匹配引起了传输延迟, 图 1 给出了到达速率  $\rho_i$ 、服务速率  $r_i$ 、缓冲长度  $q_i$ 、延迟等待时间  $t_{di}$  及最大延迟约束  $t_D$  之间的关系。其中图 1(a)、1(b)、1(c) 分别对应到达速率小于、等于以及大于服务速率时的传输模式。 $t_c$  表示在该时刻缓存刚好为空, 这意味着当  $t \geq t_c$  时该缓存不再被调度。图 1(a) 表明: 对于  $r_1 > \rho_1$ , 当延迟等待时间为  $t_c$  时, 缓存刚好为空, 数据包等待时间为  $t_{d1}$ , 对应的缓存长度为  $q_1$ , 且  $t_{d1}$ 、 $q_1$  在调度周期序列中将逐渐减小, 直至为零。显然, 图 1(a) 中最大的包延迟为  $t_1 = t_{d1}$ 。图 1(b) 中,  $r_1 = \rho_1$ , 这表明所有的数据包具有同样的延迟等待时间  $t_1 = t_{d1}$ 。图 1(a) 和 1(b) 给出的实际上是一种理想的到达及服务速率模型, 由于信源数据包的随机性及信道的时变特性,  $\rho_i$  和  $r_i$  通常会随时间变化, 因而图 1(c) 给出了一个更为一般的速率模型, 从图示可知, 不论当前缓存有多大, 若  $t_D$  必须小于  $\delta = t_D - t_1$ , 那么服务速率应该满足  $r_i \geq r_D$ 。在多用户环境下, 可达的服务速率与信道条件、用户数以及收发天线的配置密切相关, 从而为包延迟约束调度问题提供了可能的切入点。

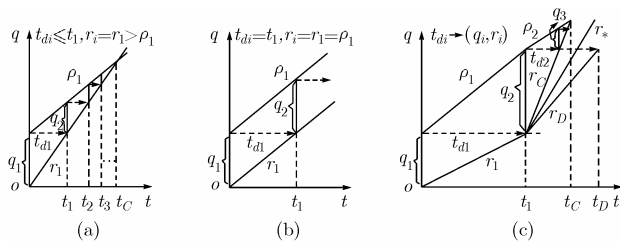


图 1  $\rho_1 < r_1$ ,  $\rho_1 = r_1$  及  $\rho_1 > r_1$  3 种条件下  $t_{di}$  与  $t_D$  之间的关系

首先分析一下 MIMO 信道中可达的和速率。令  $\mathcal{L}_k = E(x_k x_k^T)$  为用户  $k$  的协方差矩阵,  $(\pi(1), \dots, \pi(M))$  为  $(1, \dots, M)$  上的任意固定的置换, 则基于 DPC 的可达速率为<sup>[4]</sup>

$$R_{\pi(k)} = \log \frac{\left| 1 + h_{\pi(k)} \left( \sum_{j \geq k} \mathcal{L}_{\pi(j)} \right) h_{\pi(k)}^\dagger \right|}{\left| 1 + h_{\pi(k)} \left( \sum_{j \geq k} \mathcal{L}_{\pi(j)} \right) h_{\pi(k)}^\dagger \right|}, \quad k=1, \dots, M \quad (2)$$

其中脏纸速率区域为协方差矩阵  $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_M$  上的速率向量联合, 当在所有的变换  $(\pi(1), \dots, \pi(M))$  上满足  $\text{Tr}(\mathcal{L}_1 + \dots + \mathcal{L}_M) \leq P$ , 则该区域为高斯广播信道的容量区域。然而, 欲通过简单方法获取该区域上任一点目前仍难以达到, 解决此问题的一种常规思路是采用迫零(ZF)方案。考虑  $\mathbf{H}$  的 QR 分解, 令  $\mathbf{H} = \mathbf{G}\mathbf{Q}$ , 这里  $\mathbf{G}$  为  $N \times M$  的下三角矩阵,  $\mathbf{Q}$  为行正交的  $M \times M$  矩阵, 酉阵  $\mathbf{Q}^H$  用来构造传输信号  $\mathbf{x} = \mathbf{Q}^H \mathbf{u}$ , 这里  $\mathbf{u}$  为构造的信息信号向量, 故有

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{G}\mathbf{Q}\mathbf{Q}^H \mathbf{u} + \mathbf{n} = \mathbf{G}\mathbf{u} + \mathbf{n} \quad (3)$$

故原来的信道转换为如下的一系列干扰子信道

$$\mathbf{y}_k = \mathcal{G}_{k,k} u_k + \sum_{j < k} \mathcal{G}_{k,j} u_j + n_k, \quad k=1, \dots, m \quad (4)$$

其中  $m$  为  $\mathbf{H}$  的秩, 满足  $m \leq M$ , 所以没有任何信息传给用户  $m+1, \dots, M$ 。由于  $\mathbf{Q}$  为酉阵, 故  $\mathbf{u}$  与  $\mathbf{x}$  的协方差矩阵具有相同的迹, 同理,  $\mathbf{u}$  与  $\mathbf{x}$  具有相同的功率约束。上述信道构建了一组并列的子信道, 最大的可达和速率为

$$R^{\text{ZF-DPC}} = \sum_{i=1}^m \log_2 (1 + P_i^* \mathcal{G}_{i,i}^2) \quad (5)$$

其中  $P_i^*$  为在并行信道上通过注水算法实施的最优功率分配, 满足

$$\left. \begin{aligned} P_i^* &= (\xi - 1 / \mathcal{G}_{i,i}^2)^+ \\ \text{s.t. } \sum_{i=1}^m P_i^* &= P, \quad i=1, \dots, m \end{aligned} \right\} \quad (6)$$

其中  $\xi$  为注水水平,  $(x)^+ = \max(x, 0)$ 。在联合调度问题中, 由于  $N > M$ , 故时分调度是可能的, 也就是说, 借助于多用户分集和包延迟约束, 合适的  $\mathbf{H}$  是可以找到的。同时, 可将调度的公平性一并考虑, 这意味着在一个调度周期, 所有用户的数据包按照令牌环方式进入调度, 而每次调度时间的分配受限于最大包延迟等待时间约束, 且满足如下优化过程

$$\begin{aligned} \max \quad & \sum_{i=1}^s \sum_{j=1}^{m(i)} t_i r_{ij} \\ \text{s.t. } \quad & \sum_{i=1}^s t_i \leq t_D \\ & \frac{q_{i_s}^{\max} + \sum_{u=1}^{v-1} t_u \rho_{un}}{t_D} \leq \min \{r_{i_s}\} \\ & \frac{V_{11} (q_{11} + t_1 \rho_{11})^\alpha}{r_{11}} = \dots = \frac{V_{sK} (q_{sK} + \sum_{u=1}^{s-1} t_u \rho_{uw})^\alpha}{r_{sK}}, \\ & v=2, \dots, s \quad m(i)=1, \dots, M \end{aligned} \quad (7) \quad (8) \quad (9)$$

其中式(7)为总的传输延迟约束, 式(8)表示对于每个子调度

周期,最短服务时间与给定服务速率之间的关系,  $q_{\text{min}}^{\text{max}}$  为子调度周期中期望的最大权重缓存。式(9)保证了 BC 信道中调度是稳定的<sup>[9]</sup>,  $\{q_m + \sum_{u=1}^{v-1} t_u \rho_{um}\}$  为第  $v$  个子调度中第  $n$  个用户的缓存长度,  $V = [V_{11}, \dots, V_{sK}]$  为正常数集合,  $\alpha \geq 1$ ,  $\mathbf{r} = [r_{11}, \dots, r_{sK}]$  为容量区域  $\mathcal{R}$  界上的向量。若从上述优化中得到可靠的解  $t_i$ , 则可在速率最大化与最大延迟约束之间取得最佳折衷。对于给定的  $\{r_{ij}, \rho_{ij}, t_D\}$  及约束式(9), 上述优化实质上为一线性规划(LP)问题。当上述优化无解, 可通过松弛  $t_D$  进而得到次优的  $t_i$ , 实际上式(8)为一个紧约束, 松弛可能会导致某些用户的延迟约束得不到满足, 尽管如此, 通过上述优化可以观察到吞吐量与包延迟随时间变化的趋势。

下面给出两个命题, 量化地描述包延迟与用户数、平均系统吞吐量以及传输天线个数之间的关系。

**命题 1** 令  $\mathcal{N}(\rho)$  为平稳的系统用户数分布, 数据包到达速率为  $\rho$ , 定义包延迟为最小的信道使用, 从而保证所有  $K$  个用户在一个调度周期中可成功接收到  $n$  个数据包, 那么当传输天线数为  $M$  时, 平均的数据包延迟  $t_{av\_d}$  满足

$$\frac{1}{n} \sum_{i=1}^{\beta \exp(\exp(\rho))} \min[r_{i1}, r_{i2}, \dots, r_{iM}] \leq t_{av\_d} \leq \frac{1}{n} \sum_{i=1}^{\sigma \exp(\exp(\rho))} r_i \quad (10)$$

**证明** 考虑两种极端情形, 即机会的单用户时分调度和多用户复用调度, 两者分别为延迟等待时间的上界和下界。对于单用户调度, 如果第  $i$  个用户的服务速率为  $r_i$ , 那么完成  $Kn$  个包传输的总时间为  $(1/n) \sum_{i=1}^K r_i$ , 对任意  $i \neq j$ , 有  $r_i \neq r_j$ , 故对不同用户完成  $n$  个包的传输时间是不相同的。对于复用调度, 最理想的方法是在每个子调度中,  $M$  个用户完成  $Mn$  个包的发送, 显然, 最大服务时间对应于最小服务速率用户, 故可得最小传输时间和为  $(1/n) \cdot \sum_{i=1}^{[K/M]} \min[r_{i1}, r_{i2}, \dots, r_{iM}]$ 。另外, 对任何机会(单用户时分或多用户复用)调度, 当  $K \rightarrow \infty$  时, 有  $\rho \rightarrow \infty$ 。同时, 平均用户数满足<sup>[10]</sup>

$$\bar{K} \geq \Phi^{-1}(\rho) = \exp(\exp(\rho)) \quad (11)$$

故可找到正常数  $\sigma$  和  $\beta$  满足  $\bar{K} = \sigma \exp(\exp(\rho))$  以及  $[\bar{K}/M] = \beta \exp(\exp(\rho))$ , 使得命题式(10)成立。

**命题 2** 令各用户包到达速率  $\rho_i, i \in \{1, \dots, K\}$  服从独立同分布(i.i.d), 若发射天线  $M$  足够大, 满足  $M \geq O(\log K)$ , 服务速率为  $r$ ,  $t_D \geq (nK/r)$ , 那么传输和速率的最大化和延迟约束可同时满足。

**证明** 首先给出 i.i.d 分布  $\rho_i$  的含义, 假如所有用户服务速率相同, 那么在给定分布条件下, 所有用户的缓存长度相同(即具有相同的延迟等待时间)。由文献[11]可知, 当  $M \geq O(\log K)$  时, 在实施多用户分集的过程中, 选择每个用户的概率独立于 SNR, 并且接近  $1/K$ , 这意味着所有用户具有同一服务速率。其次, 考虑极端情形, 对于单用户时分调度, 用户是按照延迟权重逐一被调度的, 每个用户的调度时间是  $t_D/K$ , 那么当服务速率为  $r$  时, 需要至少完成  $n$

个包的传输, 故有  $t_D \geq (nK/r)$ 。

#### 4 仿真及性能分析

假定 MIMO 信道满足瑞利衰落, 信道矩阵循环对称且元素服从零均值、单位方差的复高斯分布。定义包延迟等待时间为包到达至离开缓存所持续的时间。图 2 给出了平均延迟与平均到达速率之间的关系。仿真表明: 对于给定的传输天线数  $M = 7, 10, 12$ , 当用户数从  $K = 10$  增加到  $K = 20$ , 包到达速率  $\rho = 50$  kbps 以及  $M = 7$  时, ZF-DPC 方案的平均包延迟增加 600ms, 而 JD-ZF-DPC 方案的包延迟增加为 95ms。对于  $K = 20$ , 当  $\rho$  的增加超过其阈值时, 平均延迟将会迅速增加, 如对于 ZF-DPC 方案, 当  $M = 7$  时,  $\rho \geq 50$  kbps。对于 JD-ZF-DPC 方案, 当  $M = 12$  时,  $\rho \geq 80$  kbps。这表明, 当用户数增加超过系统所能支撑的极限时, 缓存需要接纳越来越多的数据包, 那么对调度时间的要求就会变得更加苛刻。

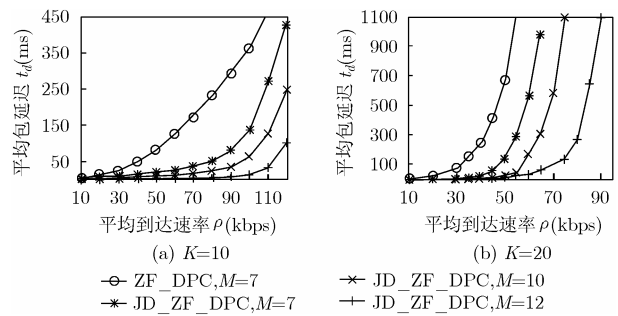


图 2 不同用户数情况下的平均延迟性能比较

图 3 给出了不同传输方案下的遍历和容量比较。在瑞利平坦衰落信道中, 当  $t \geq 4$  时, ZF-DPC 方案低于 Sato 界大约  $1.5 \text{bps/t/Hz}$ , 而对于 JD-ZF-DPC 方案,  $t_d = 60 \text{ms}$  和  $t_d = 100 \text{ms}$  之间的和速率大约相差  $2.5 \text{bps/t/Hz}$ , 当  $t \geq 8$  时, 这种速率差将保持恒定。总之, 随着  $t_d$  的增长, JD-ZF-DPC 的和速率将逐渐接近 ZF-DPC。

最后, 图 4 给出了总的传输和速率与包延迟之间的关系。

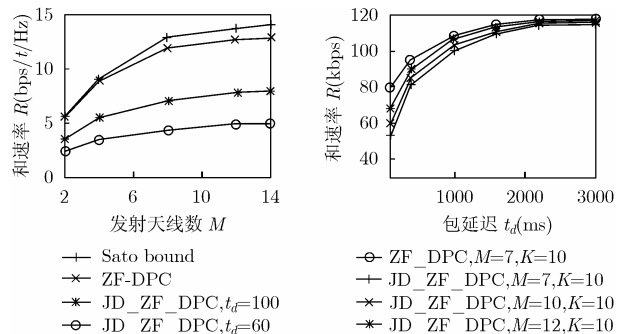


图 3 不同  $t_d$  时的遍历和容量比较 ( $K=20$ , SNR=6dB)

图 4 和速率与包延迟关系比较 (SNR=6dB)

仿真结果表明,当  $t \geq 2200\text{ms}$  时,发射天线数  $M$  将不会对和速率产生影响,两种方案的和速率接近  $R = 110\text{kbps}$ ,从而验证了当  $t_d \rightarrow \infty$  时,JD-ZF-DPC 和 ZF-DPC 两种方案具有同样的和速率。

## 5 结束语

本文给出了一种 MIMO 广播信道中联合包速率和包延迟约束的调度方案,当每个用户只有一个接收天线且总的用户数大于发射天线数时,提出了一种贪婪搜索算法,该算法将多用户分集和包延迟约束联合考虑,旨在满足传输延迟性能的同时最大化系统传输容量。仿真结果表明,在一定的用户数和发射天线数的配置下,联合调度将在和速率与包延迟性能之间取得有效的折衷,从而使系统的整体传输性能得到优化。

## 参 考 文 献

- [1] Berry R A and Yeh E M. Cross-layer wireless resource allocation[J]. *IEEE Signal Processing Mag.*, 2004, 21(5): 59-68.
- [2] Boche H and Wiczanowski M. Optimal scheduling for high speed uplink packet access- a cross-layer approach. in Proc. Vehicular Technology Conference, VTC, Los Angeles, USA, September 2004, 5: 2575-2579.
- [3] Boche H and Wiczanowski M. Stability region of arrival rates and optimal scheduling for MIMO-MAC-a cross-layer approach. in Proc. International Zurich Seminar on Comm., ETH Zurich, Switzerland, 2004: 18-21.
- [4] Jindal N and Goldsmith A. Dirty-paper coding versus TDMA for MIMO broadcast channels[J]. *IEEE Trans. on Inform. Theory*, 2005, 51(5): 1783-1794.
- [5] Samardzija D and Mandayam N. Multiple antenna transmitter optimization schemes for multiuser systems. in Proc. Vehicular Technology Conference, VTC, Orlando, FL, USA, 2003, 1: 399-403.
- [6] Liang Ying-Chang and Zhang Rui. Transmit optimization for MIMO channels with mixed delay-constrained and no-delay-constrained services. in Proc. Acoustics, Speech, and Signal Processing. ICASSP, Montreal, Canada, 2004, 4: 17-21.
- [7] Taesang Yoo and Goldsmith A. Optimality of zero-forcing beamforming with multiuser diversity. in Proc. IEEE International, Conference on Communication, ICC, Seoul, Korea, 2005, 1: 542-546.
- [8] Jiang J, Buehrer R M, and Tranter W H. High-speed downlink packet transmission with spatial multiplexing and scheduling. in Proc. Wireless Communications and Networking Conference, WCNC, Atlanta, Georgia USA, 2004, 4: 21-25.
- [9] Eryilmaz A, Srikant R, and Perkins J R. Stable scheduling policies for broadcast channels. in Proc. IEEE International Symposium on Information Theory, Lausanne, Switzerland, 2002: 382-382.
- [10] Airy M, Shakkattai S, and Heath R W. Spatially greedy scheduling in multi-user MIMO wireless systems. The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, Pacific Grove, CA, USA, 2003, 1: 982-986.
- [11] Sharif M and Hassibi B. On the capacity of MIMO broadcast channel with partial side information[J]. *IEEE Trans. on Inform. Theory*, 2005, 51(2): 506-522.

杨 涛: 男, 1970 年生, 博士后, 研究方向为宽带无线通信理论及信号处理.

胡 波: 男, 1968 年生, 教授, 博士生导师, 研究方向为无线通信中的信号处理、图像压缩以及嵌入式系统的开发.