Optimal Scheduling User Number in Massive MIMO with QoS Requirement

Lei Chen(✉) and Lu Zhang

Jiangsu Province Key Laboratory of Intelligent Industry Control Technology,
Xuzhou University of Technology, Xuzhou 221018, China
`chenlei@xzit.edu.cn`

Abstract. The Massive multiple-input multiple-output (MIMO) system can schedule dozens of end user equipment at each time slot, however, different quality-of-service (QoS) requirements needs different scheduling policy. Some QoS requirements of buffering services are related to the stability of long term transmit rate, and the instantaneous rate depends on the scheduling policy and channel state. Therefore it is difficult to build direct relationship between the QoS requirement and optimal scheduling user number at each time slot in Massive MIMO system. Based on the effective capacity (EC) theory, the relationship among the number of scheduling user, the QoS requirement and the effective transmit rate is built. The simulation result shows that EC can be described by a smooth function of the number of scheduled users and the QoS requirement.

Keywords: Massive MIMO · Quality of service · Quality of experience

1 Introduction

The huge differentiation among emerging mobile services poses the challenge to guarantee the quality-of-service (QoS). Some new technologies [1–3], such as compressive sensing and big data analysis, are proposed to predict the network traffic and user behavior. Based on network traffic analysis and user behavior analysis, new performance evaluation approaches [4] and routing schemes [5,6] are designed to guarantee the QoS and improve energy-efficiency. According to special QoS requirements, some refined network selection schemes [7] and user selection schemes [8] are designed in access network side. However, the instability of wireless access network is still the bottleneck to improve the end user experience. As two key technologies in the future 5G networks, Massive MIMO

This work is supported in part by the Natural Science Foundation of Jiangsu Province of China (No. BK20161165), the applied fundamental research Foundation of Xuzhou of China (No. KC17072), the Open Fund of the Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology. and the Ministry of Housing and Urban-Rural Development Science and Technology Planning Project (2016-R2-060).

H. Song et al. (Eds.): SIMUtools 2019, LNICST 295, pp. 627–632, 2019.
https://doi.org/10.1007/978-3-030-32216-8_61

and small cells are proposed to deal with increasing traffic data and diverse requirements. The base station adopting Massive multiple-input multiple-output (MIMO) technology is usually equipped with a few hundreds of antennas for simultaneously serving a large number of users. The researches have demonstrated that the large number of antennas can increase the spectral efficiency (SE), and effectively improve the end user experience [9]. Because the large number of users are scheduled simultaneously, The scheduling scheme is critical important to guarantee QoS. The first key problem in Massive MIMO scheduling should be the maximum number of user scheduled in a time slot under QoS constraint. In reference [10], a algorithm is proposed to compute the maximum number of user and the power allocation according to the QoS requirements for Massive MIMO. However, the QoS constraint mentioned in [10] is just related to the instantaneous rate. Generally, the QoS is affected by the jitter of a long term rate. Therefore, we analyze the relationship among the number of user, QoS constraint and achievable transmit rate in this paper.

2 Effective Capacity of Massive MIMO

We consider a massive MIMO cellular network where the BS of each cell equipped with an array of M antennas communicates with K single-antenna UEs at the time, out of a set of N UEs which have unlimited demand for data. Each cell is assigned an index in the set $\mathcal{L}$. The geographical position of UE $k \in \{1, ..., K\}$ in cell $l \in \mathcal{L}$ is given by $\mathbf{z}_{lk} \in \mathbb{R}^2$. The time-frequency resources are divided into equal frames whose time and bandwidth is smaller or equal to the coherence time and the coherence bandwidth of all UEs respectively. Thus all the channel are static within the frame. Let $\mathbf{h}_{jlk} \in \mathbb{R}^N$ denote the channel response between BS j and UE k in cell l, which are drown as realizations from zero-mean circularly symmetric complex Gaussian distributions [11]:

$$h_{jlk} \sim \mathcal{CN}\left(0, d_j\left(\mathbf{z}_{lk}\mathbf{I}_M\right)\right) \tag{1}$$

where $\mathbf{I}_M$ is the $M \times M$ identity matrix. The function $d_j(\mathbf{z})$ gives the variance of the channel attenuation from BS j to any UE position $\mathbf{z}$. Let S be the amount of symbols transmitted in each frame, B out of the S symbols are reserved for pilot signaling. Thus the remaining $S-B$ symbols are allocated for payload data. The symbols have transmit power $p_{lk} = \frac{\rho}{d_{j(\mathbf{z}_{lk})}}$, where ρ is a design parameter for the channel attenuation inversion policy. The policy make the average effective channel gain the same for all UEs:$\mathbb{E}\{p_{lk}\|\mathbf{h}_{llk}\|^2\} = M\rho$. The received download signal at UE k in cell j in a frame is given by:

$$y_{jk} = \sum_{l \in \mathcal{L}} \sum_{m=1}^{K} \mathbf{h}_{ljk}^T \mathbf{w}_{lm} x_{lm} + n_{jk} \tag{2}$$

where $(\cdot)^T$ denotes transpose, x_{lm} is the symbol transmitted to UE m in cell l, $w_{lm} \in \mathbb{C}^M$ is the corresponding precoding vector, and $\|w_{lm}\|^2$ is the allocated download transmit power. It can be expressed as

$$w_{lm} = \sqrt{\frac{p_{jk}}{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{jk}\|^2\}}}\mathbf{g}_{jk} \tag{3}$$

where the average transmit power $p_{jk} \geq 0$ is a function of the UE position, but not the instantaneous channel realizations. The vector $\mathbf{g}_{jk} \in \mathbb{C}^M$ defines the spatial directivity of the transmission and is based on the acquired CSI. The SNIR is given by (see reference [12] for the power control policy):

$$SINR_{jk} = \frac{p_{jk}\frac{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{jk}\mathbf{h}_{jjk}\|^2\}}{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{jk}\|^2\}}}{\sum_{l\in\mathcal{L}}\sum_{m=1}^{K} p_{lm}\frac{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{lm}\mathbf{h}_{ljk}\|^2\}}{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{lm}\|^2\}} - p_{jk}\frac{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{jk}\mathbf{h}_{jjk}\|^2\}}{\mathbb{E}_{\mathbf{h}}\{\|\mathbf{g}_{jk}\|^2\}} + \sigma^2}. \tag{4}$$

In reference [12], the achievable Spectral Efficiency in download of cell j can be written by:

$$SE_j = K\left(1 - \frac{B}{S}\right)\log_2\left(1 + \frac{1}{I_j^{scheme}}\right) \tag{5}$$

where

$$\begin{aligned} I_j^{scheme} = &\sum_{l\in\mathcal{L}_{|}(\beta)\backslash\{j\}}\left(\mu_{jl}^{(2)} + \frac{\mu_{jl}^2 + \left(\mu_{jl}^{(1)}\right)^2}{G^{scheme}}\right) \\ &+ \frac{\left(\sum_{l\in\mathcal{L}}\mu_{jl}^{(1)}Z_{jl}^{scheme} + \frac{\sigma^2}{\rho}\right)\left(\sum_{\ell\in\mathcal{L}_j(\beta)}\mu_{jl}^{(1)} + \frac{\sigma^2}{B\rho}\right)}{G^{scheme}} \end{aligned} \tag{6}$$

where the G^{scheme} and Z_{jl}^{scheme} depends on the different receive combining schemes, $G^{MR} = M$ and $Z_{jl}^{scheme} = K$ with MR combining, while $G^{ZF} = M - k$ and

$$Z_{jl}^{ZF}\begin{cases} K\left(1 - \frac{\mu_{jl}^{(1)}}{\sum_{\ell\in\mathcal{L}_j(\beta)}\mu_{jl}^{(1)} + \frac{\sigma^2}{B\rho}}\right) & \text{if } l \in \mathcal{L}_j(\beta), \\ K & \text{if } l \notin \mathcal{L}_j(\beta). \end{cases} \tag{7}$$

The following notation was used:

$$\mu_{jl}^{(w)} = \mathbf{E}_{\mathbf{z}_{lm}}\left\{\left(\frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})}\right)^w\right\} \text{ for } w = 1, 2. \tag{8}$$

The QoS requirement is often formulated by

$$Pr\left(\max_{1\leq i\leq N} Q_i(0) > B\right) \leq \epsilon. \tag{9}$$

According to reference [13], for some large buffer size B, given the QoS constraint ϵ and by choosing $\theta = -\log(\epsilon)/B$, the QoS requirement can be expressed as an effective capacity (EC) problem:

$$\lambda \leq \min_{1\leq j\leq N} C_k(\theta), \tag{10}$$

where

$$C_k(\theta) = \frac{1}{\theta}\lim_{n\to\infty}\frac{-1}{n}\ln\mathbb{E}\left(e^{-\theta\sum_{t=1}^{n} r_k(t)}\right), \tag{11}$$

the $r_k(t)$ is the rate allocated to user k in cell j at time t. We assume that the scheduling scheme at the base station stochastically picks the K users out of a set of the N active users for transmission, thus the r_k can be written as:

$$r_k(t) = \begin{cases} \frac{N_f}{K}\left(1-\frac{B}{S}\right)\log_2\left(1+\frac{1}{I_j^{scheme}}\right), & \text{w.p. } \frac{K}{N}, \\ 0, & \text{w.p. } 1-\frac{K}{N}. \end{cases} \tag{12}$$

Let $\nu = N_f\left(1-\frac{B}{S}\right)\log_2\left(1+\frac{1}{I_j^{scheme}}\right)$, the EC of user k is rewritten as:

$$\begin{aligned} C_k(\theta) &= \frac{1}{\theta}\lim_{n\to\infty}\frac{-1}{n}\ln\sum_{\tau=0}^{n}\left(e^{-\theta\tau\frac{\nu}{K}}P\left\{\left(\sum_{t=1}^{n} r_k(t)\right)=\tau\frac{\nu}{K}\right\}\right) \\ &= \frac{1}{\theta}\lim_{n\to\infty}\frac{-1}{n}\ln\sum_{\tau=0}^{n}\left(e^{-\theta\tau\frac{\nu}{K}}\binom{n}{\tau}\left(\frac{K}{N}\right)^{\tau}\left(1-\frac{K}{N}\right)^{(n-\tau)}\right) \\ &= \frac{1}{\theta}\lim_{n\to\infty}\frac{-1}{n}\ln\sum_{\tau=0}^{n}\left(\binom{n}{\tau}\left(e^{-\theta\frac{\nu}{K}}\frac{K}{N}\right)^{\tau}\left(1-\frac{K}{N}\right)^{(n-\tau)}\right) \\ &= \frac{1}{\theta}\lim_{n\to\infty}\frac{-1}{n}\ln\left[1-\frac{K}{N}\left(1-e^{-\theta\frac{\nu}{K}}\right)\right]^{n} \\ &= \frac{-1}{\theta}\ln\left[1-\frac{K}{N}\left(1-e^{-\frac{\theta\nu}{K}}\right)\right]. \end{aligned} \tag{13}$$

While $\theta \to \infty$ the EC is 0, when $\theta \to 0$ the EC is $\frac{K}{N}\frac{N_f}{K}\left(1-\frac{B}{S}\right)\log_2\left(1+\frac{1}{I_j^{scheme}}\right)$. With the assumption that users have same QoS requirements, let $f(K) = C_k(\theta)$ is the function of scheduled user number K at each time slot, with the fixed θ. Therefore, the derivative of this EC function with respect to K is:

$$\nabla f(K) = \frac{Ke^{\frac{\theta\nu}{K}} - K - \theta\nu}{(NK\theta)\left(Ne^{\frac{\theta\nu}{K}} - Ke^{\frac{\theta\nu}{K}} + K\right)}. \tag{14}$$

It is obviously, the value of Eq. (14) is greater than 0, at the point $K = 1$. And the denominator of the Eq. (14) is always greater than 0, the numerator of the Eq. (14) is a oscillatory function. The first inflection point of Eq. (14) is a suboptimal solution. By using binary search, it is easy to find the maximum value of formula (13).

3 Simulation Result

In our simulation, two hundred users that are served by one base station equipped one thousand antennas have the same QoS requirement. The pilot reuse factor is 1, and we set the coherence block length to 400, set the SNR to 5 dB, set the pathloss factor to 3.7. The QoS parameter θ gradually increase from e^{-10} to e^{10}, the low value of θ means non-strict demand for real-time, the high value of θ means that the service must satisfy high real-time request and high stability. The unit of EC is *bits/S/Hz*. In our simulation algorithm, we first compute the spectral efficiency based on the code in reference [12], and then the EC is obtained according to the formula (13). The simulation result is showed as Fig. 1. The Fig. 2 is a slice of the Fig. 1, as the $\theta = 0.15$. The EC is a smooth function of user number. In the Fig. 1, the QoS requirement is $\log \theta$. With the low value of θ, the point of optimal EC is near to the point of $K = 1$. As the value of θ increases, the optimal point moves to the point of $K = N$, and the achievable EC descends sharply. Because the higher θ requires higher stability.

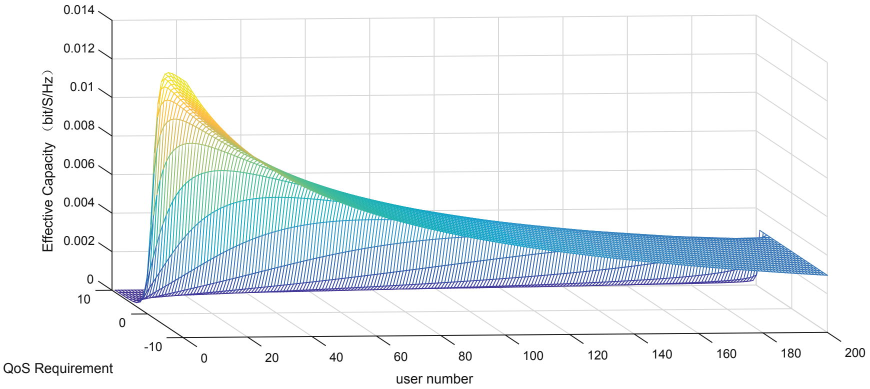

Fig. 1. Effective capacity under various QoS requirement.

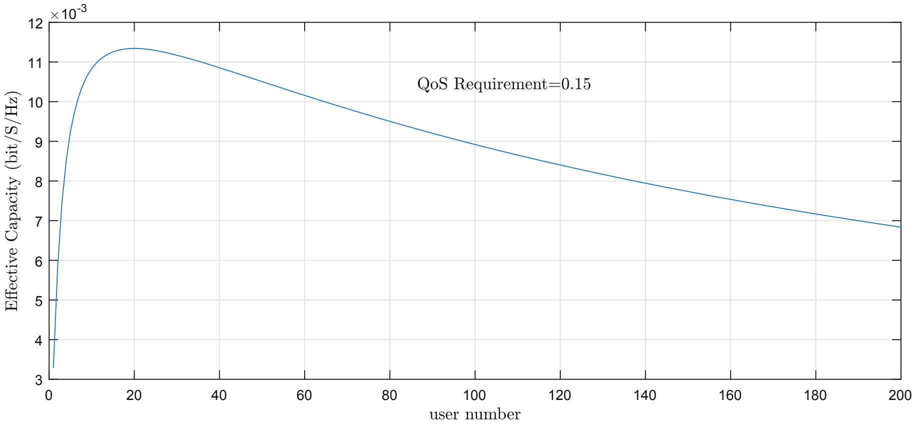

Fig. 2. Effective capacity with fixed QoS requirement.

4 Conclusion

We deduce the EC function with respective to the number of scheduling users and to the QoS requirement parameter in Massive MIMO system. The EC can be described by a smooth function of the number of scheduled users and the QoS requirement. A simulation is performed under various QoS requirement to survey the characteristics of the function. Designing the fast algorithm to find the optimal scheduling number of users under QoS requirement in Massive MIMO is a practical research work in the future.

References

1. Jiang, D., Wang, W., Shiand, L., Song, H.: A compressive sensing-based approach to end-to-end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. (2018). https://doi.org/10.1109/TNSE.2018.2877597
2. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **1**(2), 1–12 (2018)
3. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. Plos One **13**(5), 1–23 (2018)
4. Chen, L., et al.: A lightweight end-side user experience data collection system for quality evaluation of multimedia communications. IEEE Access **6**(1), 15408–15419 (2018)
5. Jiang, D., Zhang, P., Lv, Z., Song, H.: Energy-efficient multi-constraint routing algorithm with load balancing for smart city applications. IEEE Internet Things J. **3**(6), 1437–1447 (2018)
6. Jiang, D., Li, W., Lv, H.: An energy-efficient cooperative multicast routing in multi-hop wireless networks for smart medical applications. Neurocomputing **220**(2017), 160–169 (2017)
7. Jiang, D., Huo, L., Lv, Z., Song, H., Qin, W.: A joint multi-criteria utility-based network selection approach for vehicle-to-infrastructure networking. IEEE Trans. Intell. Transp. Syst. **19**(10), 3305–3319 (2018)
8. Chen, L., Jiang, D., Bao, R., Xiong, J., Liu, F., Bei, L.: MIMO scheduling effectiveness analysis for bursty data service from view of QoE. Chinese J. Electron. **26**(5), 1079–1085 (2017)
9. Abarghouyi, H., Razavizadeh, S.M., Bjornson, E.: QoE-aware beamforming design for massive MIMO heterogeneous networks. IEEE Trans. Veh. Technol. **67**(9), 8315–8323 (2018)
10. Chaudhari, S., Cabric, D.: QoS aware power allocation and user selection in massive MIMO underlay cognitive radio networks. IEEE Trans. Cogn. **4**(2), 220–231 (2018)
11. Gao, X., Edfors, O., Rusek, F., Tufvesson, F.: Massive MIMO performance evaluation based on measured propagation data. IEEE Trans. Wirel. Commun. **14**(7), 3899–3911 (2015)
12. Björnson, E., Larsson, E.G., Debbah, M.: Massive MIMO for maximal spectral efficiency: how many users and pilots should be allocated? IEEE Trans. Wirel. Commun. **15**(2), 1293–1308 (2016)
13. Wu, D., Negi, R.: Effective capacity: a wireless link model for support of quality of service. IEEE Trans. Wirel. Commun. **2**(4), 630–643 (2003)