Towards Opportunistic Fair Scheduling in Wireless Networks*

Daiqin Yang, Dongxu Shen, Wenjian Shao, and Victor O. K. Li
Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
Email: {dqyang,dxshen,wjshao,vli}@eee.hku.hk

***Abstract*— Opportunistic transmission scheduling schemes improve system capacity by taking advantage of independent time varying channels in wireless networks. In the design of such scheduling schemes, the fairness criterion plays an important role in the tradeoff of total system capacity and the achievable throughput of individual users. To meet different fairness demands with a unified opportunistic scheduling scheme, in this paper, we have extended the well known opportunistic scheduling scheme PFS into αPFS, which satisfies arbitrary fairness demands, varying from proportional fairness to max-min fairness, through adjusting the parameter α. To further improve the achievable diversity gains of αPFS, we extend the αPFS scheme into an αPFS-P scheme. Performances of αPFS and αPFS-P are studied and compared. As demonstrated in the simulation results, both αPFS and αPFS-P can achieve adjustable fairness criteria, varying from proportional fairness to max-min fairness. Compared with αPFS, αPFS-P achieves higher diversity gains with degraded short term performance, which is still better than the performance of PFS.**

I. Introduction

Recently, opportunistic transmission scheduling in wireless networks has drawn much research interests due to its attractive system capacity. With independent time-varying channels for different mobile users, opportunistic transmission scheduling increases system throughput for delay-tolerant traffics by always picking those mobile users with better channels to transmit or receive. However, if user channels are different not only in instantaneous quality but also in average quality, there is always a tradeoff between system capacity and fairness when exploiting such multi-user diversity gains. For example, if we always choose the users with the best instantaneous channel conditions, the system capacity can be maximized, but fairness may be unacceptable.

As fairness criterion plays an important role in the performance balancing of system capacity and the achievable throughput of individual users [1], many fair opportunistic scheduling schemes have been proposed to provide different tradeoffs between system capacity and fairness criteria. In [2], [3], [4], fairness is implemented in an outcome fair manner through max-min fair resource allocations. Mobile users receive equal throughput despite their average channel conditions. In [5], [6], [7], fairness is implemented in an effort fair manner, with which, mobile users receive equal share of "air time". The throughput achieved by each user is proportional to its average channel condition. In the proportional fair scheduling (PFS) scheme [8] of the Qualcomm HDR system, fairness is also implemented in such equal time share manner. In all the above fair opportunistic scheduling algorithms, fairness criteria are pre-assigned and fixed. There is a lack of unified opportunistic scheduling schemes which can provide adjustable fairness criteria and meet different fairness demands. In [9], [10], adjustable tradeoffs between system capacity and fairness criteria are provided from throughput maximization to proportional fairness. However, they are still lacking when higher fairness criteria are desired.

In this paper, we propose an opportunistic αPFS scheduling scheme that can both exploit multi-user diversity gains and provide flexible fairness adjustment from effort fairness to outcome fairness as α increases from 1 to ∞. With αPFS, as α increases, aside from the increase of system fairness, the achievable system diversity gain decreases. To maintain similar level of diversity gains for αPFS with different values of α, an αPFS-P scheme is further proposed. Performances of αPFS and αPFS-P are studied and compared. Simulation results show that, with the same value of α, αPFS and αPFS-P can achieve similar long term fairness, varying from proportional fairness to max-min fairness. Compared with αPFS, αPFS-P achieves higher diversity gains with relatively worse short term performance, which is still better than that of the PFS.

The rest of the paper is organized as follows. First, the system model and the PFS algorithm used in HDR system are introduced in Section II. In Section III, αPFS and αPFS-P opportunistic scheduling schemes are introduced and analyzed. Section IV gives the simulation results and discussions. Finally, conclusions are presented in Section V.

II. System Model and the PFS Scheme

In this paper we consider a centralized down link scheduling system with one base station and N mobile users. Each user have infinite backlogs at the base station. Time is divided into constant time slots and one transmission can be scheduled at each time slot. With independent path loss, shadowing, and fading of wireless channels, the received signal-to-noise ratio $SNR_i(t)$, $i \in \{1, 2, ..., N\}$, of mobile user i at time slot t are different and may vary independently across time slots due to fast fading. Through adaptive modulation and coding schemes, mobile user i can support transmissions of a maximum data

*This research is supported in part by the Research Grants Council, Hong Kong Special Administrative Region, China, under Grant No. 7152/05E.

1-4244-0355-3/06/$20.00 (c) 2006 IEEE

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

rate $R_i(t)$ with received signal-to-noise ratio $SNR_i(t)$. $R_i(t)$ is usually a bounded and nondecreasing function of $SNR_i(t)$. Both $SNR_i(t)$ and $R_i(t)$ are assumed to be constant within one time slot. Before scheduling each time slot, the base station can get full knowledge about the maximum supportable date rates of all mobile users within this time slot.

Let Π denote the set of scheduling strategies, and $S^{\pi}(t) \in \{1, 2, ..., N\}$ denote the index of the scheduled user at time slot t with scheduling strategy π, $\pi \in \Pi$. At the start of each time slot t, the base station schedules data transmissions to mobile user $S^{\pi}(t)$, which maximizes the scheduling metric $M_i^{\pi}(t)$:

$$S^{\pi}(t) = \arg\max_i M_i^{\pi}(t) \tag{1}$$

To exploit multi-user diversity gains and maintain "air time" fairness among the users, a well known proportional fair scheduling scheme, PFS [8], has been proposed and adopted in Qualcomm's HDR system. With PFS, at the beginning of time slot t, the scheduler picks the user $S^{\mathrm{PSF}}(t)$ to receive according to:

$$S^{\mathrm{PSF}}(t) = \arg\max_i \left(\frac{R_i(t)}{T_i(t)}\right) \tag{2}$$

Here $T_i(t)$ is the exponentially smoothed average of the service rate received by user i before the start of time slot t. At the end of each time slot t, $T_i(t)$ is updated as follows:

$$T_i(t+1) = \left(1 - \frac{1}{\omega}\right) \cdot T_i(t) + \frac{1}{\omega} \cdot \hat{R}_i(t) \tag{3}$$

where ω is the time constant for the exponential low pass filter, $\hat{R}_i(t)$ is the scheduled data rate of user i at time t.

$$\hat{R}_i(t) = \begin{cases} R_i(t) & : \quad i = S(t) \\ 0 & : \quad i \neq S(t) \end{cases} \tag{4}$$

It has been proved that, as $\omega \to \infty$, PFS provides equal share of "air time" to mobile users [11]. Besides, since $R_i(t)$ has been incorporated into the scheduling metric of PFS, the fluctuation of wireless channels can be tracked and multi-user diversity gains is achieved by PFS.

III. Opportunistic αPFS Scheduling

In this section, to meet different fairness demands through a unified scheduling scheme, we propose an opportunistic fair scheduling scheme, αPFS, which is extended from the PFS scheduling scheme. With αPFS, as we increase the value of α for better fairness, system diversity gain decreases as the price for a smoother short term performance. Since it may not be desirable to trade diversity gains for short term performance gains in systems with high α values, an αPFS-P opportunistic scheduling scheme is further proposed. With αPFS-P, both achievable diversity gains and short term performances are maintained at similar levels for systems with different α values.

A. Opportunistic αPFS scheme

As $\omega \to \infty$, the scheduling strategy of PFS is equivalent to a utility-based opportunistic scheduling scheme that maximizes the aggregate utility of the system [11]:

$$\max \sum_{i=1}^{N} U_i^{\mathrm{PFS}}(T_i) \tag{5}$$

$$\text{s. t.} \sum_{i=1}^{N} p_i \leq 1 \tag{6}$$

where, $T_i = \mathbf{E}\left[\hat{R}_i\right]$ is the average service rate received by mobile user i, and p_i is the probability that user i is scheduled to receive in a time slot. The utility function of each user in PFS is $U_i^{\mathrm{PFS}}(x) = \log x$. With this kind of utility-based scheduling scheme, the average service rates achieved by each user, T_i, satisfy *proportional fairness* [12], which means for any other sets of feasible average service rates T_i^*, the aggregate of the proportional change is negative:

$$\sum_{i=1}^{N} \frac{T_i^* - T_i}{T_i} \leq 0 \tag{7}$$

In [13], a generalized α *proportional fairness*, in which, proportional fairness is a special case of $\alpha = 1$, is defined. A set of average service rates, T_i, satisfies α proportional fairness when for any other sets of feasible average service rates T_i^*, the aggregate of α proportional change is negative:

$$\sum_{i=1}^{N} \frac{T_i^* - T_i}{(T_i)^{\alpha}} \leq 0 \tag{8}$$

As the value of α increases, the fairness criterion of α proportional fairness increases. Max-min fairness is approached as $\alpha \to \infty$. This α proportional fairness can be achieved through the same utility-based scheduling schemes as that of proportional fairness, except for a different utility function of $U_i^{\alpha\mathrm{PFS}}(x) = (1-\alpha)^{-1} \cdot x^{1-\alpha}$, $\alpha \neq 1$.

If we can modify the utility function of each user in PFS from $U_i^{\mathrm{PFS}}(x) = \log x$ to $U_i^{\alpha\mathrm{PFS}}(x) = (1-\alpha)^{-1} \cdot x^{1-\alpha}$, $\alpha \neq 1$, we can generalize PFS into αPFS and achieve α proportional fairness resource allocation. To implement αPFS, we follow the same form of scheduling metric defined in PFS, $M_i^{\mathrm{PFS}} = R_i(t) \cdot \left(U_i^{\mathrm{PFS}}(T_i(t))\right)'$, which can be interpreted as a greedy algorithm to the utility optimization problem. The scheduling metric of αPFS is therefore defined as:

$$M_i^{\alpha\mathrm{PFS}} = R_i(t) \cdot \left(U_i^{\alpha\mathrm{PFS}}(T_i(t))\right)' = \frac{R_i(t)}{(T_i(t))^{\alpha}} \tag{9}$$

where $T_i(t)$ is the same exponentially smoothed average of the service rate received by user i as defined in PFS (3). At the start of each time slot t, the base station chooses to transmit to the mobile user:

$$S^{\alpha\mathrm{PSF}}(t) = \arg\max_i \left(\frac{R_i(t)}{(T_i(t))^{\alpha}}\right) \tag{10}$$

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

With αPFS, as we increase the value of α, resource allocation with better fairness can be achieved. As $\alpha \rightarrow \infty$, max-min fairness can be approached.

B. Opportunistic αPFS-P scheme

In αPFS, as the value of α changes, aside from the change in system fairness, the achievable diversity gain and the short term performance also changes. This can be shown in the following analysis.

Let $D_i^{\pi}(t)$ denote the denominator of the scheduling metric M_i^{π}. Then

$$D_i^{\alpha\mathrm{PFS}}(t) = (T_i(t))^{\alpha} \tag{11}$$

For any un-scheduled time slot t, $D_i^{\alpha\mathrm{PFS}}(t)$ of user i is updated as:

$$\begin{aligned} D_i^{\alpha\mathrm{PFS}}(t+1) &= \left(\left(1-\frac{1}{\omega}\right)\cdot T_i(t)\right)^{\alpha} \\ &= \left(1-\frac{1}{\omega}\right)^{\alpha}\cdot D_i^{\alpha\mathrm{PFS}}(t) \end{aligned} \tag{12}$$

With $\omega \gg \alpha$, we have:

$$\begin{aligned} \left(1-\frac{1}{\omega}\right)^{\alpha} &= 1 + C_{\alpha}^{1}\cdot\left(-\frac{1}{\omega}\right) + \ldots + C_{\alpha}^{\alpha}\cdot\left(-\frac{1}{\omega}\right)^{\alpha} \\ &\approx 1-\frac{\alpha}{\omega} \end{aligned} \tag{13}$$

Then, the updating of $D_i^{\alpha\mathrm{PFS}}(t)$ during those un-scheduled time slots is equivalent to:

$$D_i^{\alpha\mathrm{PFS}}(t+1) = \left(1-\frac{\alpha}{\omega}\right)\cdot D_i^{\alpha\mathrm{PFS}}(t) \tag{14}$$

Compared with PFS, which uses an updating of $D_i^{\mathrm{PFS}}(t+1) = \left(1-\frac{1}{\omega}\right)\cdot D_i^{\mathrm{PFS}}(t)$ in those un-scheduled time slots, αPFS has an equivalent time constant of $\omega_e = \frac{\omega}{\alpha}$. The value of this ω_e affects the rate of decrease of $D_i(t)$ in non-scheduled time slots. The larger the ω_e, the slower the rate of decrease of $D_i(t)$, and thus the longer the allowable delay between two consecutive scheduled transmissions of a mobile user.

Therefore, as we increase α in αPFS for better fairness, the equivalent time constant of ω_e gets smaller, which in turn shortens the allowable delay between consecutive transmissions. In opportunistic scheduling schemes, this shortened allowable delay affects both the achievable diversity gains and the short term performances of the system. Smaller allowable delay imposes stricter constraint for the scheduler to choose high data rate users to transmit to. Thus less diversity gains can be exploited, while smoother short term performances can be achieved as the delay between consecutive transmissions is restricted to small values.

However, for the sake of system capacity, it may not be desirable to trade diversity gains for smoother short term performance as α increases. To maintain similar level of achievable diversity gains as we increase α for fairer resource allocation, an αPFS-P scheme is proposed which increases the time constant ω proportionally as α increases. In αPFS-P, the scheduling metric is the same as that of αPFS, while the updating of $T_i(t)$ is changed to keep the equivalent time constant ω_e the same for different values of α:

$$T_i(t+1) = \left(1-\frac{1}{\omega\cdot\alpha}\right)\cdot T_i(t) + \frac{1}{\omega\cdot\alpha}\cdot\hat{R}_i(t) \tag{15}$$

With this modified updating of $T_i(t)$, α becomes a parameter that is only related to the desired fairness criteria. The achievable diversity gains together with the short term performances can be easily adjusted through only one parameter, namely, the time constant ω.

IV. Numerical Examples

To evaluate and study the performances of αPFS and αPFS-P, we use computer simulations. In the simulated system, the duration of a time slot is 1.67 ms, and the maximal Doppler Shift fm is 7.5Hz. Each user encounters an independent Rayleigh fading channel with different long term average signal-to-noise ratio SNR_i, which is evenly distributed from 1 to 97 (0dB to about 20dB) as the user index i increases from 1 to N. The maximum achievable data rate $R_i(t)$ is a logarithmic function of the instantaneous $SNR_i(t)$: $R_i(t) = W\cdot\log(1+SNR_i(t))$, where W is the system bandwidth. The final service rates received by each user with different scheduling schemes are normalized to the long term system average service rate achieved by a Round Robin (RR) scheme, with average channel conditions.

In the rest of this section, we will first show the achievable service rate by each individual mobile user under different scheduling schemes. Then short term fairness performance and the relationship between system capacity and the value of the time constant ω will be studied.

A. Achievable service rates of individual users

To evaluate the performance of αPFS and αPFS-P, we have simulated a system with $N = 25$ and $\omega = 1000$. Achievable service rates at individual users under αPFS and αPFS-P are compared with a round robin (RR) scheme and an absolute max-min fair scheduler (MFS), which does not vary with time varying channel conditions and always picks the user with the least average service rate to transmit to.

Fig. 1 and Fig. 2 show the long term average service rate received by each user under αPFS and αPFS-P. We see that system fairness achieved with both αPFS and αPFS-P improves as α increases. As α increases, the system average service rate achieved by αPFS decreases more rapidly than that of the αPFS-P system. The decrease in system average service rate of αPFS-P is mainly due to the increase of fairness. As fairness increases, more transmission time is spent on users with worse channel conditions, degrading the whole system capacity. With αPFS, system capacity decrease with increasing α comes not only from increased fairness, but also from decreased diversity gain due to shortened equivalent time constant ω_e. As shown in Fig. 1 and Fig. 2, when $\alpha = 16$, the system average service rate of 16PFS-P is about 60%

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

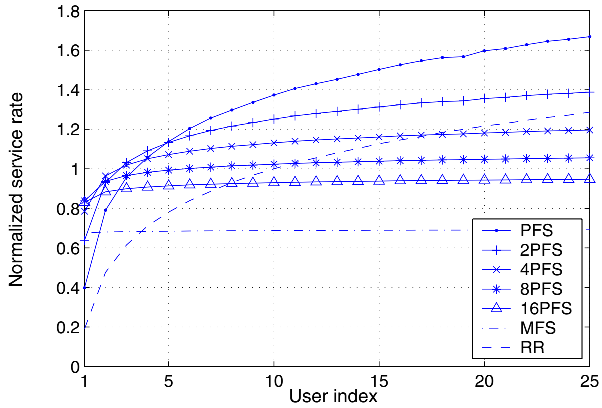

Fig. 1. Service rate of different users under αPFS with ω=1000.

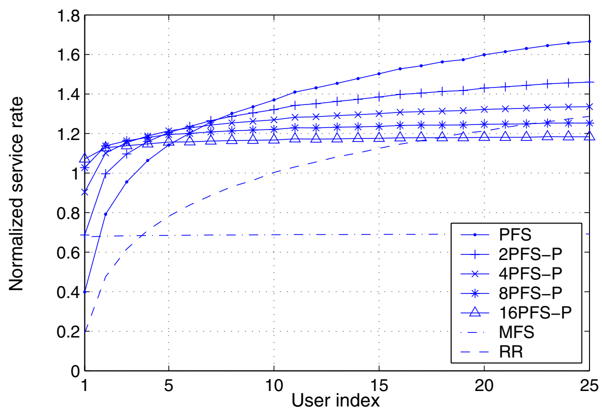

Fig. 2. Service rate of different users under αPFS-P with ω=1000.

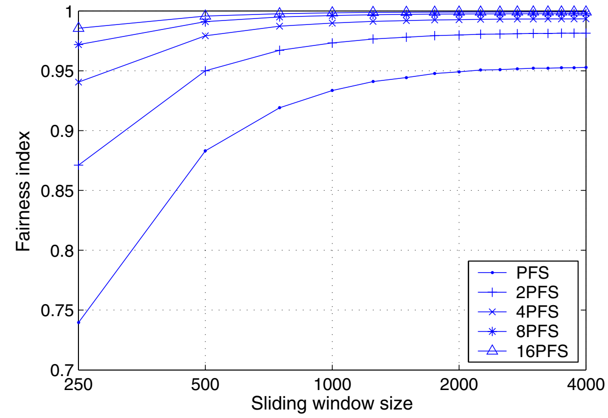

Fig. 3. Fairness index under αPFS (N=25, ω=1000).

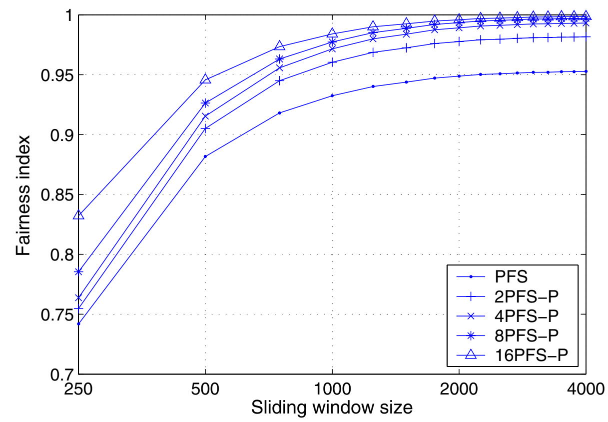

Fig. 4. Fairness index under αPFS-P (N=25, ω=1000).

higher than that of the MFS scheme, while with 16PFS, the improvement is less than 40%.

B. Short term fairness performances

To compare the short term performances of αPFS and αPFS-P, which are highly related to their equivalent time constant ω_e, we use a sliding window method [14] to quantize their short term fairness performances.

In this sliding window method, Jain's fairness index [15] is used as the fairness indicator. In a sliding window from time slot j to time slot $j+k-1$, where j is the starting slot, and k is the sliding window size, the fairness index is calculated as:

$$F_j^k = \frac{\left(\sum_{i=1}^{N} T_i^{j,k}\right)^2}{N \cdot \sum_{i=1}^{N} \left(T_i^{j,k}\right)^2} \tag{16}$$

where $T_i^{j,k} = \frac{1}{k} \cdot \sum_{n=j}^{j+k-1} \hat{R}_i(n)$ is the average service rate received by user i in this sliding window. A higher fairness index F corresponds to a fairer resource allocation. Max-min fairness is achieved when $F = 1$ and absolute unfairness will result in $F = \frac{1}{N}$.

After this fixed-size sliding window has slid across the whole simulated time series, the fairness indices of each sliding windows are averaged and plotted as the short term fairness performance under this sliding window size k. A smaller difference in the averaged fairness between small and large sliding window sizes represents a smoother short term performance.

Fig. 3 and Fig. 4 show the short term fairness performances of αPFS and αPFS-P under different sliding window sizes for the simulated system of $N = 25$, $\omega = 1000$. We find that, with both αPFS and αPFS-P, average fairness increases with α for all sliding window sizes, and the long term fairness performances of both αPFS and αPFS-P are equal when their α values are the same. These results represent the basic property of α proportional fairness. We also find that, with αPFS, the difference in the average fairness between small and large sliding window sizes decreases as α increases. However, with αPFS-P, the differences are invariant with different values of α. This is because, compared with αPFS-P, αPFS has smaller equivalent time constant ω_e with $\alpha > 1$. This shortened ω_e decreases the achievable diversity gains while increasing the smoothness of the short term performances in αPFS. In αPFS-

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.

P, the equivalent time constant ω_e is kept the same for all α values. Thus, both the achievable diversity gains and the short term performances are kept the same in αPFS-P with different α values, including the case of PFS with $\alpha = 1$.

C. Impact of time constant ω on system capacity

In this subsection, we will study the impact of the time constant ω on the diversity gains of both αPFS and αPFS-P. In our simulated systems, N, the number of active users, is fixed at 25, while the time constant ω varies from 100 to 4000. Fig. 5 and Fig. 6 depict the long term system service rate achieved by these two mechanisms.

As shown in Fig. 5 and Fig. 6, the system capacity of both αPFS and αPFS-P increases logarithmically with the increase of the time constant ω. This is because, with a larger time constant ω, the allowable delay between consecutive scheduled transmissions is increased. So, as ω increases, it allows opportunistic schedulers more freedom to choose the user with the highest currently supportable data rate to transmit to. Thus, diversity gains can be increased. Comparing Fig. 5 and Fig. 6, we also find that, with any $\alpha > 1$, αPFS-P achieves higher system capacity than αPFS. The reason is just as we have explained in previous subsections. Besides, as shown in Fig. 6, after the time constant ω has increased to more than 1000, the speed of increase of system capacity with high α valued αPFS-P decreases due to the saturation of diversity gains. It is also easier for higher fairness demanding αPFS-Ps (with larger α values) to saturate, when the equivalent time constant ω_e is kept the same.

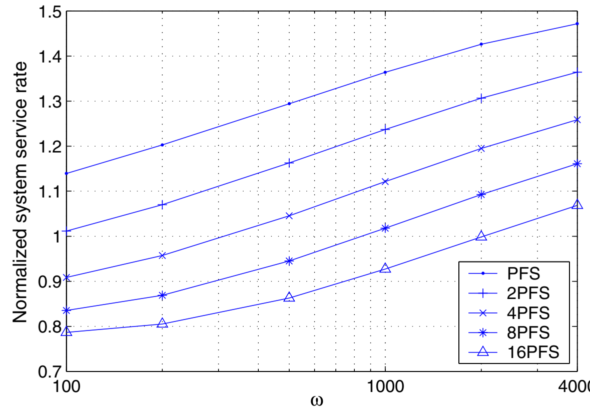

Fig. 5. System service rate in systems with different ω under αPFS.

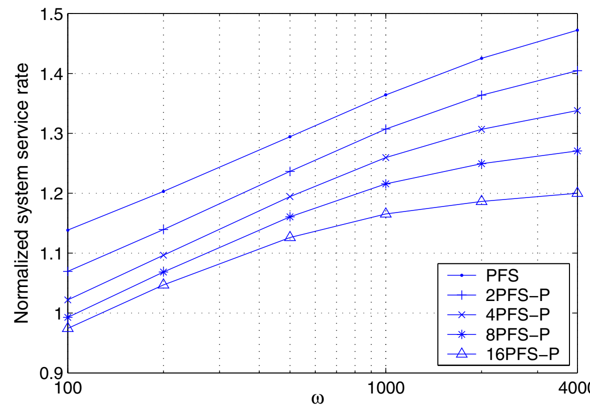

Fig. 6. System service rate in systems with different ω under αPFS-P.

V. Conclusions

In this paper, we have proposed opportunistic fair scheduling schemes, αPFS and αPFS-P, which can provide flexible tradeoffs between fairness and system capacity when exploiting multi-user diversity gains. Fairness criteria from proportional fairness to max-min fairness can be achieved by both αPFS and αPFS-P as α increases from 1 to ∞. With αPFS, as α increases to meet higher fairness demands, the achievable diversity gain decreases as the price to pay for a smoother short term performance. With αPFS-P, the achievable diversity gains together with the short term performances are kept the same for different α values, and α becomes a parameter that is only related to the desired fairness criteria. Compared with αPFS, αPFS-P achieves higher diversity gains with degraded short term performance, which is still better than the performance of PFS.

References

[1] Y. Cao and V.O.K. Li, "Scheduling algorithms in broadband wireless networks," *Proc. IEEE*, vol. 89, no. 1, pp. 76–87, Jan. 2001.

[2] S. Borst and P. Whiting, "Dynamic rate control algorithms for HDR throughput optimization," in *Proc. IEEE INFOCOM '01*, 2001.

[3] Y. Cao and V.O.K. Li, "Wireless packet scheduling for two-state link models," in *Proc. IEEE GLOBECOM '02*, 2002.

[4] Y. Liu and E. Knightly, "Opportunistic Fair Scheduling over Multiple Wireless Channels," in *Proc. IEEE INFOCOM '03*, 2003.

[5] X. Liu, E. K. P. Chong, and N. B. Shroff, "Opportunistic transmission scheduling with resource-sharing constraints in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 10, pp. 2053–2064, 2001.

[6] Y. Liu, S. Gruhl, and E. W. Knightly, "WCFQ: an opportunistic wireless scheduler with statistical fairness bounds," *IEEE Trans. Wireless Commun.*, vol. 2, no. 5, pp. 1017–1028, Sept. 2003.

[7] Y. Cao, V.O.K. Li, and Z. Cao, "Scheduling delay-sensitive and best-effort traffics in wireless networks," in *Proc. IEEE ICC '03*, 2003.

[8] A. Jalali, R. Padovani, and R. Pankaj,, "Data throughput of CDMA-HDR a high efficiency-high data rate personal communication wireless system," in *Proc. IEEE VTC 2000*, 2000.

[9] D. Panigrahi and F. Khaleghi, "Enabling trade-offs between system throughput and fairness in wireless data scheduling techniques," in *World Wireless Congress (3G Wireless)*, May 2003.

[10] L. Yang, M. Kang, and M. S. Alouini, "On the capacity-fairness tradeoff in multiuser diversity systems," in *International Symposium on Wireless Systems and Networks (ISWSN05)*, 2005.

[11] P. Viswanath, D.N.C. Tse, and R. Laroia, "Opportunistic beamforming using dumb antennas," *IEEE Trans. Inform. Theory*, vol. 48, pp. 1277–1294, 2002.

[12] F. Kelly, "Charging and rate control for elastic traffic," *Eur. Trans. Telecommun.*, vol. 8, pp. 33–37, Jan./Feb. 1997.

[13] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *IEEE Trans. Networking*, vol. 8, no. 5, pp. 556–567, 2000.

[14] T. Nandagopal, T. Kim, X. Gao, and V. Bharghavan, "Achieving MAC layer fairness in wireless packet networks," in *Proc. of ACM Mobicom '00*, 2000.

[15] R. Jain, D. Chiu, and W. Hawe, "A Quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *DEC Research Report TR-301*, Sept. 1984.

This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2006 proceedings.