A Hierarchical Proportional Fair Scheduler

Kinda Khawam
GET/ENST - Telecom Paris
khawam@enst.fr

Jean-Marc Kelif
France Telecom R&D
jeanmarc.kelif@francetelecom.com

***Abstract*— In 3G wireless networks, opportunistic schedulers take advantage from the delay-tolerance of data applications to ensure that transmission occurs when radio channel conditions are most favourable. In the well-known opportunistic scheduler "Proportional Fair" (PF), the Base Station (BS) serves only one user per time slot which is optimal when the Signal-to-Noise Ratio (SNR) scales linearly with the effective transmission rate. However, for a logarithmic relation between the channel quality and the transmission rate, scheduling one user at a time does not always result in maximum channel utilization. For that reason, we put forward in this paper a new scheduler that divides the cell into two categories, serving alternately the first one in a CDMA fashion and the second one according to the PF scheduler. The mathematical and simulation analysis provided proves that the proposed scheduler increases the overall throughput of the cell.**

I. Introduction

Data Traffic is increasingly popular in 3G mobile networks. New technologies like HDR (High Data Rate) [1] and its equivalent in 3GPP, HSDPA (High System Data Packet Access) [2], offer higher data rates than previous architectures notably through the opportunistic scheduling: time is divided into very short intervals (1.67 ms for HDR and 0.67ms for HSDPA) and the BS transmits at full power to a single user per time slot (intra-cell interference cancellation). However, because the SNR does not scale linearly with the feasible rate for all users, scheduling one user at a time may not result in maximum channel utilization. We suggest in this paper to divide the cell in two zones where slots will be distributed across the different zones in a Weighted Round Robin (WRR) fashion: users in the first zone will be served according to regular CDMA while users in the second zone will be served one at a time according to the PF algorithm. We will show how this new hierarchical scheduler, termed HPF, will increase the total throughput of the cell.

The rest of the paper is organized as follows. In Section II, we present the proposed HPF approach. In Section III, we analyse its performance; in particular we obtain analytical results for the mean rate in Standard PF (SPF) and in HPF for a fixed number of users. We also show how our hierarchical scheduler improves the performance of the cell in terms of average rate. In Section IV, we extend the model of our HPF algorithm to accommodate a dynamic user configuration and corroborate the results obtained by simulations. Finally, in Section V, we give a brief conclusion.

II. A Hierarchical Scheduling Approach

A. The Radio Resource

In this subsection, we present the model of the radio resource and the way it is shared among users. We then compute the peak rate of each user accordingly.

1) The Propagation Model: Let P be the transmission power emitted by the BS and γ_k the free space path loss. The power received by user k is then:

$$P_k = P \cdot \gamma_k \tag{1}$$

The adopted model for the free space path loss is:

$$\gamma_k = 1 \text{ if } r_k \leq \epsilon \text{ and } \gamma_k = (\tfrac{\epsilon}{r_k})^\beta \text{ otherwise}$$

where β is the path loss exponent (taking values between 2 and 5), ϵ is the maximum distance at which the full power P is received and r_k the distance separating user k from the BS.

2) The feasible rate: For user k, the signal-to-noise ratio and the peak rate C_k, which is assumed to follow a logarithmic relation according to Shannon, are respectively given by:

$$SNR_k = \frac{\phi_k \cdot P_k}{(\eta + I_{BS} + I_k)},\ C_k = W \cdot \log_2(1 + SNR_k) \tag{2}$$

where W is the cell bandwidth, P_k the power received by the user, η the background noise, ϕ_k the fraction of power transmitted to user k by its BS, I_{BS} the interference caused by other BSs to user k and I_k the interference caused to user k by its own BS and given by the following relation:

$$I_k = P_k \cdot (f_k(1 - \phi_k) + h_k\phi_k) \tag{3}$$

where f_k is the orthogonality factor and h_k is a self-noise coefficient. We assume that $h_k < 1$ which means that the interference experienced by a user from its own signal is not greater than the signal itself.

The authors in [3] showed that it is better for a BS to transmit to only one user at a time rather than to transmit to several such users simultaneously for $h_k < 1$, yet, we will show here that this is not necessarily true if the feasible rate does not scale linearly with the SNR.

For that, we consider a time interval during which the total power level P at the BS remains constant and we

0-7803-9455-0/06/$20.00 (c) 2006 IEEE

suppose that throughout the interval a proportion ϕ_k of the total power P is allocated to user k such that $\sum_k \phi_k = 1$. From (2) and (3), the feasible rate of user k is:

$$C_{k,CDMA} = W \log_2 \left(1 + \frac{\phi_k P_k}{\eta + I_{BS} + P_k(f_k(1-\phi_k) + h_k\phi_k)}\right)$$

Now we suppose that each user k is allocated the total power P but for a fraction ϕ_k of the interval. During the period that the BS transmits to user k, no power is allocated to other users in the cell, thus user k endures no interference from other users within the cell, and consequently the feasible rate of user k during the whole interval is:

$$C_{k,TDMA} = \phi_k W \log_2 \left(1 + \frac{P_k}{\eta + I_{BS} + P_k h_k}\right)$$

The SNR of a user depends on the radio channel and varies with time due to user mobility and fading effects. We will only consider the impact of fast fading in our model by replacing SNR_k of user k by:

$$SNR_k(t) = SNR_k \cdot x_k(t) \quad (4)$$

where $x_k(t)$ are i.i.d copies of some stationary process $x(t)$ with unit mean that represents the effect of fast fading. As we consider Rayleigh fading throughout the paper, $x_k(t)$ follows an exponential distribution.
It follows that the asymptotic rate for every user k is:

$$C_k = lim_{T\to\infty} \frac{1}{T} \int_0^T C_k(SNR_k(t))dt$$
$$= \int_0^\infty C_k(SNR_k \cdot x_k) \cdot g(x_k)dx_k$$

where $g(x_k) = e^{-x_k}$, $x_k \geq 0$

To compare the two schemes, we compute the following ratio $R_k = \frac{lim_{T\to\infty} \frac{1}{T}\int_0^T C_{k,CDMA}(SNR_k(t))dt}{lim_{T\to\infty} \frac{1}{T}\int_0^T C_{k,TDMA}(SNR_k(t))dt}$:

$$R_k = \frac{\int_0^\infty \log_2(1 + \frac{\phi_k snr_k x_k}{1+snr_k\cdot(f_k(1-\phi_k)+h_k\phi_k)})e^{-x_k}dx_k}{\phi_k \int_0^\infty \log_2(1 + \frac{snr_k x_k}{1+snr_k h_k})e^{-x_k}dx_k}$$

where

$$snr_k = \frac{\gamma_k P}{\eta + I_{BS}} \quad (5)$$

To evaluate R_k, we will plot it as a function of snr_k for different values of f_k, h_k and ϕ_k. In the special case where $f_k = 0$ and $\phi_k < 1$, R_k is strictly greater than 1 meaning that in the case of perfect orthogonality, it is always more profitable to apply CDMA rather than TDMA especially for high values of snr_k as we can see in Figure 1 where R_k is plotted for different values of ϕ_k. In Figure 2, R_k is depicted for $\phi_k = 1/30$, for $0.1 \leq f_k \leq 0.9$ and for $h_k = 0.03$ which is a reasonable value for h_k ([5], [13]). We see that for $f_k \leq 0.6$, it is better to apply CDMA rather than TDMA scheduling. In Figure 3, R_k is depicted as a function of snr_k for $\phi_k = 1/30$, for $0.01 \leq h_k \leq 0.09$ and for $f_k = 0.3$. And we see that for a moderate value of f_k, it is always better to apply CDMA rather than TDMA especially for high values of h_k.

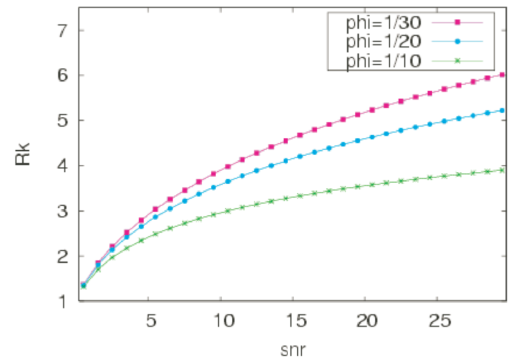

Fig. 1
RATIO FOR f=0

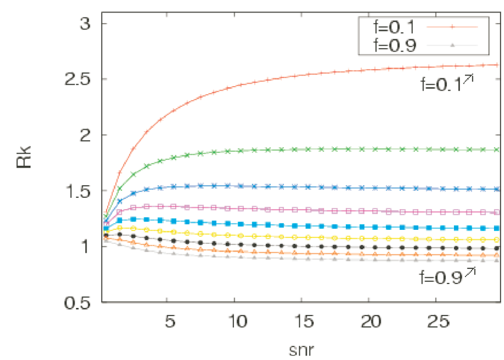

Fig. 2
RATIO FOR h=0.03

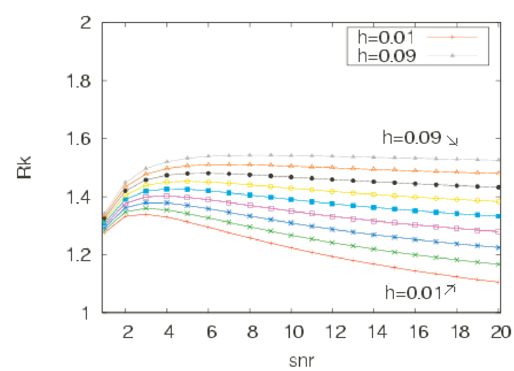

Fig. 3
RATIO FOR f=0.3

B. Delimiting the cell into zones

We saw in the preceding section that omitting to adopt a simplifying model where the transmission rate received by users varies linearly in the SNR has lead us to the conclusion that it is not always better to serve one user at a time. This conclusion has driven us to divide the cell into two zones: in the first zone, termed **Zone 1**, the peak rate is almost never proportional to the SNR and thus TDMA scheduling does not result in optimal performances, contrary to the second zone, termed **Zone 2**, where this linearity is satisfied with non-negligible probability and therefore TDMA scheduling is appropriate. Our target is to define the previously introduced zones in such a way that the two assumptions become more realistic inside each zone.

We know that $\log_2(1+x) \approx x/\ln(2)$ is valid for very small values of x. Thus, we will adopt that approximation for $x \leq 0.05$ leading to a maximum deviation of 2.4% from the actual value at $x = 0.05$. As a result, we assume that the following approximation:

$$C_{k,TDMA}(t) \approx \phi_k W \cdot SNR_k(t)/\ln(2) \tag{6}$$

is valid for

$$SNR_k(t) = SNR_k \cdot x_k(t) \leq 0.05 \tag{7}$$

where

$$SNR_k = \frac{P_k}{\eta + I_{BS} + P_k h_k} \tag{8}$$

denotes the SNR user k would get in the absence of fast fading.

Accordingly, we will define the geographical region in the cell, termed **Zone 1**, where users do not profit fully from TDMA scheduling because they almost never (we can say with probability 99%) satisfy formula (7), i.e.:

$$\mathbb{P}(SNR_k \cdot x_k(t) \geq 0.05) > 0.99 \tag{9}$$

To compute (9), we need to evaluate (8) beginning with I_{BS}. For that we adopt the approach in [4] where the inter-cell interference is caused by all BSs (working at maximum power P) in the network taken to be homogeneous with a density of stations ρ_{BS} per unit of surface and thus we get:

$$I_{BS} = P \cdot \gamma(r) \frac{2\pi\rho_{BS}}{\beta - 2} r^{\beta}[(2R-r)^{2-\beta} - (R_S - r)^{2-\beta}]$$

with R being the cell radius and R_S the network radius.

By defining $\eta = \alpha P$ and taking $\rho_{BS} = \frac{1}{\pi R^2}$, we get from the path loss model and (5):

$$\frac{1}{snr_k} = \alpha max\big((\frac{r_k}{\epsilon})^{\beta}, 1\big) + \nu(r_k) \tag{10}$$

with $\nu(r_k) = \frac{2}{\beta-2} \cdot (\frac{r_k}{R})^2[(2\frac{R}{r_k} - 1)^{2-\beta} - (\frac{R_S}{r_k} - 1)^{2-\beta}]$

We obtain from (8) and (10):

$$\frac{1}{SNR_k} = \alpha max\big((\frac{r_k}{\epsilon})^{\beta}, 1\big) + \nu(r_k) + h_k \tag{11}$$

Finally, from (9) and (11), and taking $R = 2.0 \cdot \epsilon$ (higher values of R induce very low rates at the border of the cell), we have the following:

$$e^{\frac{-0.05}{SNR_k}} > 0.99 \Rightarrow r < R_1 \approx 1.4 \cdot \epsilon \tag{12}$$

obtained for $\beta = 4.0$ (urban environment), $R_S = 10R$ (quasi-infinite network [4]), $\alpha = 10^{-10}$, $h_k = 0.03$ which are realistic parameters ([5], [6]) and $f_k = 0.1$. Assumed values for the orthogonality factor [6] are 0.4 for urban macro-cells and 0.06 for urban micro-cells. The size of **Zone 1** being close to a micro-cell (a micro-cell is approximately equal in size to half of a macro-cell), it is reasonable to assume that $f_k = 0.1$.

Therefore, users in **Zone 2** are those who are situated at a distance greater than R_1 and for whom formula (6) is valid with non-negligible probability.

Because TDMA scheduling is not appropriate to users in **Zone 1**, we suggest to serve them in a CDMA fashion (we prove in section III that this will result in a performance enhancement) while continuing on serving users in **Zone 2** in a TDMA fashion, more specifically according to the well-known PF algorithm. To carry out this hybrid scheduling, we suggest in this paper an alternative scheduling approach to plain PF, termed HPF. At its first hierarchical level, HPF distributes the slots between the two zones in a WRR fashion (**Zone 1** will be served with probability $\mathbb{P}(A_1)$ and **Zone 2** with probability $\mathbb{P}(A_2)$). And at its second level, HPF serves users inside **Zone 1** in a CDMA fashion and users inside **Zone 2** according to the PF algorithm. We show next that HPF augments the overall mean capacity of the cell.

III. Analytical study of HPF Vs. Standard PF

The total number of users present in **Zone 1** is n_1 and in **Zone 2** is n_2. Traffic demand is uniformly distributed in the cell.

A. The Average Peak rate in HPF

Proposition 1:
The Average Peak Rate obtained in **Zone 1** is:

$$C_{HPF,1} = W\mathbb{P}(A_1) \int_0^{\infty} e^{-x_k} dx_k \big[\log_2(1 + SNR(\epsilon) \cdot x_k)\frac{\epsilon^2}{R_1^2} + \int_{\epsilon}^{R_1} \log_2(1 + SNR(r_k) \cdot x_k)\frac{2r_k dr_k}{R_1^2}\big]$$

with $SNR(r_k) = \frac{\phi_k \cdot snr(r_k)}{1 + snr(r_k)\cdot(f_k(1-\phi_k) + h_k\phi_k)}$, $\phi_k = \frac{1}{n_1}$ and event A_1={**Zone 1** is being served}.

Proof:
The average peak rate of a user k belonging to **Zone 1** served according to HPF (CDMA scheduling) is:

$$C_{HPF,k,1} = W\mathbb{P}(A_1) \int_0^{\infty} \log_2(1 + SNR(r_k) \cdot x_k) e^{-x_k} dx_k \tag{13}$$

The probability that user k is located between r_k and r_k+dr_k in **Zone 1** being $\frac{2\pi r_k dr_k}{\pi R_1^2}$, we obtain $C_{HPF,1}$.

Proposition 2:
The Average Peak Rate obtained in **Zone 2** is:

$$C_{HPF,2} = W\mathbb{P}(A_2)\int_0^\infty f_{n_2}(x_k)dx_k\Big[\int_{R_1}^{R}\log_2(1+SNR(r_k)\cdot x_k)\frac{2r_k dr_k}{R^2-R_1^2}\Big]$$

with $SNR(r_k) = \frac{snr(r_k)}{1+snr(r_k)\cdot h_k}$, $f_{n_2}(x_k) = e^{-x_k}(1-e^{-x_k})^{n_2-1}$ and event A_2={**Zone 2** is being served}.

Proof:
In **Zone 2**, users are served according to the PF algorithm and we know that the latter schedules, at time slot t, the user with the highest SNR relative to its current average SNR, i.e.,

$$\text{user } k^* = arg_k max \frac{SNR_k(t)}{T_k(t)}$$

with $SNR_k(t) = SNR(r_k)\cdot x_k(t)$ and T_k being the exponentially smoothed SNR, given by:

$$T_k(t+1) = (1-\frac{1}{\tau})\cdot T_k(t) + \frac{1}{\tau}\cdot SNR_k(t)\cdot \mathbb{1}_{user(t)=k}$$

with $\mathbb{1}_{user(t)=k}$ being the indicator function which equals 1 if user k was chosen at time slot t and 0 otherwise. τ is a time constant that captures the time-scales of the PF scheduler.

The random variables representing the fading being i.i.d, we have that $T_k = SNR(r_k)\cdot U_k$, where U_k are identically distributed random variables (but not independent).
If $\frac{1}{\tau}\to 0$, then:

$$T_k \to SNR(r_k)\cdot V \tag{14}$$

where V is a constant. In practice, τ has large values because this offers the opportunity of waiting a long time before scheduling a user when its channel quality is maximal: the scheduler is then expected to better exploit multi-user diversity. As a result, we will adopt formula (14) in our analysis. We refer to [7] for rigorous justifications of the above claims.
Therefore, the average rate of a user k belonging to **Zone 2** served according to the standard PF is:

$$C_{HPF,k,2} = W\mathbb{E}[\log_2(1+SNR_k(t))\cdot \mathbb{1}\{A_2\}\cdot \mathbb{1}\{\frac{SNR_k(t)}{V\cdot SNR(r_k)} = max_{l=1..n_2}\frac{SNR_l(t)}{V\cdot SNR(r_l)}\}]$$

$$= W\mathbb{P}(A_2)\mathbb{E}[\log_2(1+x_k SNR(r_k))\cdot \mathbb{1}\{x_k = max_{l=1..n_2}x_l\}]$$
$$= W\mathbb{P}(A_2)\mathbb{E}[\log_2(1+SNR(r_k)x_k)|B_{n_2}]\mathbb{P}(B_{n_2})$$

with $B_{n_2} = \{x_k = max(x_1,..,x_{n_2})\}$.

Knowing that $\mathbb{P}(B_{n_2}) = \frac{1}{n_2}$ because the random variables $x_k(t)$ are i.i.d, we obtain:

$$C_{HPF,k,2} = W\mathbb{P}(A_2)\int_0^\infty \log_2(1+SNR(r_k)\cdot x_k)f_{n_2}(x_k)dx_k \tag{15}$$

with $f_{n_2}(x_k) = e^{-x_k}(1-e^{-x_k})^{n_2-1}$.

The probability that user k is located between r_k and r_k+dr_k in **Zone 2** being $\frac{2\pi r_k dr_k}{\pi(R^2-R_1^2)}$, we get $C_{HPF,2}$.

B. The Average Peak rate in Standard PF

The analysis is identical to the one done in **Zone 2** except that $\mathbb{P}(A_2) = 1$ because the cell is considered as a whole. Thus, we get from (15) the average peak rate of a user k:

$$C_{PF,k} = W\int_0^\infty \log_2(1+SNR(r_k)\cdot x_k)f_n(x_k)dx_k \tag{16}$$

where $SNR(r_k) = \frac{snr(r_k)}{1+snr(r_k)\cdot h_k}$ and $f_n(x_k) = e^{-x_k}(1-e^{-x_k})^{n-1}$.

Hence, the Average Peak Rate of the cell in Standard PF is:

$$C_{PF} = W\int_0^\infty f_n(x_k)dx_k\Big[\log_2(1+SNR(\epsilon)\cdot x_k)\frac{\epsilon^2}{R^2} + \int_\epsilon^R \log_2(1+SNR(r_k)\cdot x_k)\frac{2r_k dr_k}{R^2}\Big]$$

We define the following two mean peak rates to evaluate later the impact of HPF on both categories of users:

$$C_{PF,1} = W\int_0^\infty f_n(x_k)dx_k\Big[\log_2(1+SNR(\epsilon)\cdot x_k)\frac{\epsilon^2}{R_1^2} + \int_\epsilon^{R_1} \log_2(1+SNR(r_k)\cdot x_k)\frac{2r_k dr_k}{R_1^2}\Big]$$

$$C_{PF,2} = W\int_0^\infty f_n(x_k)dx_k\Big[\int_{R_1}^{R} \log_2(1+SNR(r_k)\cdot x_k)\frac{2r_k dr_k}{R^2-R_1^2}\Big]$$

C. Average Gain

1) Numerical Results: We begin by computing the ratio of the mean peak rate obtained by applying CDMA divided by the mean peak rate obtained by applying the PF algorithm:

$$\Re_k = \frac{\int_0^\infty \log_2(1+\frac{snr_k x_k}{n+snr_k\cdot(f_k(n-1)+h_k)})e^{-x_k}dx_k}{\int_0^\infty \log_2(1+\frac{snr_k x_k}{1+snr_k h_k})(1-e^{-x_k})^{n-1}e^{-x_k}dx_k}$$

We plot, in Figure 4, $\Re_k$ as a function of snr_k, for $h_k = 0.03$, $f_k = 0.1$ and $n = 30$. We can see that it is better to apply CDMA when serving users until snr_k goes approximately below 5.0, which means for $r_k \geq 1.5\cdot\epsilon$ where it is preferable to switch to the PF scheduling. This highlights the validity and sensibility of the idea behind our hierarchical scheduler.

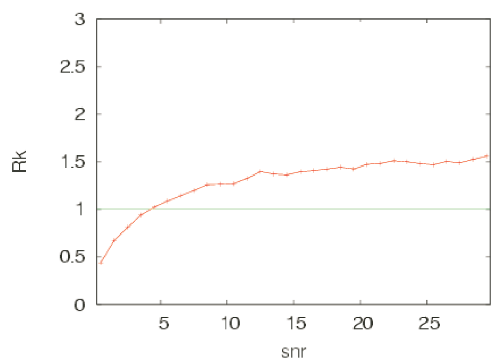

Fig. 4
RATIO FOR f=0.1, h=0.03

To evaluate the average gain in the total average peak rate of the cell, we compute the following ratios:

$$\Re = \frac{C_{HPF,1}+C_{HPF,2}}{C_{PF}},\ \Re_1 = \frac{C_{HPF,1}}{C_{PF,1}} \text{ and } \Re_2 = \frac{C_{PF,2}}{C_{HPF,2}}$$

Following the assumption that users are uniformly distributed in the cell, it is reasonable to suppose that $n_1 = n \cdot \frac{R_1^2}{R^2}$ and $n_2 = n \cdot \frac{R^2-R_1^2}{R^2}$. To compute $\Re$, we still have to set the values taken by $\mathbb{P}(A_1)$ and $\mathbb{P}(A_2)$. We begin by setting the value of $\mathbb{P}(A_2)$ such that users in **Zone 2**, in our HPF, have the same chance to access the channel as users in Standard PF situated in the same geographical area, for that reason $\mathbb{P}(A_2) = \frac{R^2-R_1^2}{R^2}$ and thus $\mathbb{P}(A_1) = 1 - \mathbb{P}(A_2) = \frac{R_1^2}{R^2}$.

We obtain for $n = 30$: $\Re_1 \approx 1.31$, $\Re_2 \approx 1.1$ and $\Re \approx 1.16$ which means that our HPF will lead to a total gain of 16% in the average peak capacity of the cell without really affecting users with low snr situated in **Zone 2**.
Nevertheless, if we want to realize more gain in the total peak rate of the cell, we increase $\mathbb{P}(A_1)$ at the expense of lowering the rates realized by users in **Zone 2**. For instance, if we serve **Zone 1** twice as often as **Zone 2**, in other words if we set $\mathbb{P}(A_1) = 2/3$ and $\mathbb{P}(A_2) = 1/3$, we obtain $\Re \approx 1.39$ leading to a gain of 39%.

2) Simulation Results: We present in this section our numerical experiments performed to illustrate the previous results. Users are served according to the Standard PF (SPF) and according to our HPF where slots are distributed such that $\mathbb{P}(A_1) = \frac{R_1^2}{R^2}$ and $\mathbb{P}(A_2) = \frac{R^2-R_1^2}{R^2}$. The number of users in each zone is fixed and equal to 15. We take $W = 1.0$, h_k=0.03, $f_k = 0.1$ and $\epsilon = 1.0$. In all experiments, we determine the average rate per user and display the results for users belonging to the same zone. We plot the average rate realized by users occupying a ring of internal radius r and external radius $r + dr$ such that $dr = 0.01$ and compare results with what we obtain using formulae (13) and (15) respectively for **Zone 1** and **Zone 2** in HPF, and formula (16) in SPF at r.

Zone 1	SPF	HPF	$\Re_1$
NUM	0.187	0.246	1.31
SIM	0.189	0.251	1.33
Zone 2	SPF	HPF	$\Re_2$
NUM	0.112	0.102	1.10
SIM	0.12	0.110	1.04

TABLE I
TOTAL AVERAGE PEAK RATE

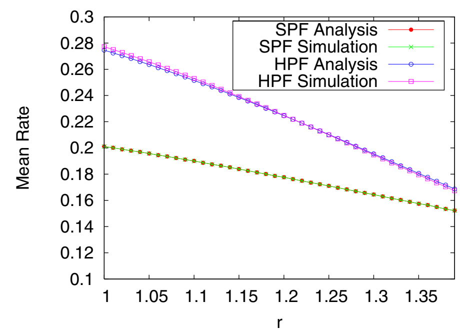

Fig. 5
AVERAGE RATE FOR ZONE 1

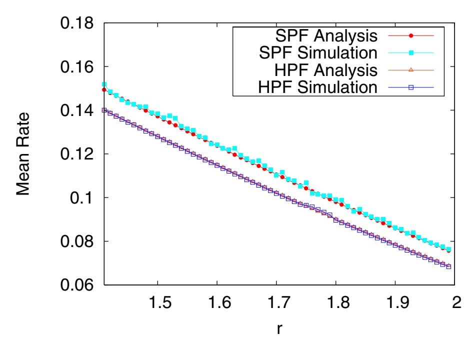

Fig. 6
AVERAGE RATE FOR ZONE 2

In Figure 6, we notice a slight degradation in the average rate in our model in comparison with SPF as predicted. Whereas in Figure 5, we can see how close users benefit from being served in a CDMA fashion rather than in a TDMA fashion as it is done in the PF scheduler. Furthermore, we can see how the analytical results give very precise estimations of the simulation results as the two curves are indistinguishable.

In table I, we computed the mean rate obtained in each zone by simulation and we compared it to the mean rate obtained numerically from $C_{HPF,1}$ and $C_{HPF,2}$ for HPF and from $C_{PF,1}$ and $C_{PF,2}$ for SPF respectively for **Zone 1** and **Zone 2**. We can see how the two sets of values are very close.

Finally, we obtain for $\Re$ the value 1.17 by simulation and 1.16 numerically.

IV. Dynamic Model of HPF

In order for **Zone x** to behave like a Generalized Processor Sharing (GPS) [8] system, the service rate of the latter must only depend on the total number of users occupying it.
In **Zone 1**, the SNR of user k is:

$$SNR_{k,1} = \frac{\phi_k}{\alpha/\gamma_k + \nu_k + f\cdot(1-\phi_k) + h\cdot\phi_k} \quad (17)$$

In order for **Zone 1** to behave like a GPS system, we need to adopt the following two assumptions in our analytical model:

Hypothesis I: We compute the SNR at an "average location" within **Zone 1** by replacing in (17) γ_k and ν_k by their sample averages γ_I and ν_I for r_k ranging from ϵ to R_1. In practical radio network dimensioning, it is not possible to use the parameters of each user but average values of the individual parameters among users ([9], [10]) which makes this first assumption fairly plausible. Besides, considering two ranges of r by considering two zones will give us much more precise values for these parameters than those obtained by considering the cell as a whole. The average parameters are obtained according to the following:

$$\alpha/\gamma_I + \nu_I = \int_{\epsilon}^{R_1} \frac{\alpha(\frac{r_k}{\epsilon})^\beta + \nu(r_k)}{R_1^2} 2r_k dr_k + (\frac{\epsilon}{R_1})^2\cdot(\alpha+\nu(\epsilon)) \quad (18)$$

Hypothesis II: Due to the fact that we are considering an average value of the SNR, the impact of fading will not be taken into consideration in the analysis. That is again a reasonable assumption in a CDMA system where the fading of the channel will be combated through the various available techniques [11] (Space Time coding, Adaptive Modulation, Dynamic Frequency Selection, etc.) and where the high rates available in **Zone 1** will further minimize the effect of fading.

In section IV-B, we will prove the validity of these two hypothesizes through simulation by showing that the discrepancy between the values obtained by making these two assumptions in the analytical model, from those obtained by simulation, where these assumptions are omitted, is negligible. Finally, we have the following, for $\phi_k = \frac{1}{n_1}$:

$$SNR_I(n_1) = (n_1\cdot(\alpha/\gamma_I + \nu_I + f) + h - f)^{-1}$$

As a consequence, each user will get the following rate when the BS divides its transmission power P equally between active users:

$$C_1 = W\log_2(1 + SNR_I(n_1))\mathbb{P}(A_1)$$

In **Zone 2**, we know from (15) that the rate of user k is:

$$C_{HPF,k,2} = W\mathbb{P}(A_2)\int_0^\infty e^{-x_k}(1-e^{-x_k})^{n_2-1}\cdot \log_2(1 + x_k\cdot SNR_{k,2})dx_k$$

with $SNR_{k,2} = (\alpha/\gamma_k + \nu_k + h)^{-1}$.

We define the following:

$$F_{k,2}(n_2) = \frac{C_{HPF,k,2}}{W\mathbb{P}(A_2)\cdot\log_2(1 + SNR_{k,2})} \quad (19)$$

By plotting $F_{k,2}(n_2)$ as a function of $R_1 \le r \le R$ for different values of n_2, we can see in Figure 7 that it varies slightly and can be well approximated by a constant for every value of n_2.

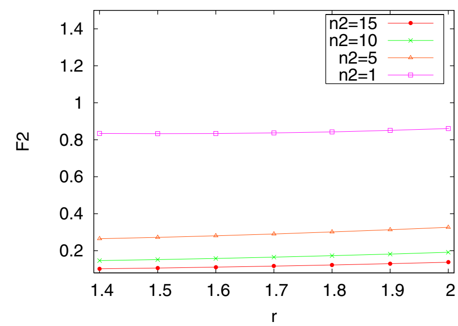

Fig. 7
$F_{k,2}(n_2)$ AS FUNCTION OF r

Thus, we define $F_2(n_2)$ that consists in replacing γ_k and ν_k by their average values γ_{II} and ν_{II} in (19). The average parameters are obtained according to:

$$\alpha/\gamma_{II} + \nu_{II} = \int_{R_1}^{R} \frac{\alpha(\frac{r_k}{\epsilon})^\beta + \nu(r_k)}{R^2 - R_1^2} 2r_k dr_k \quad (20)$$

We adopt the subsequent approximation $F_{k,2}(n_2) \approx F_2(n_2)$ necessary for **Zone 2** to behave like a GPS system. We compute the following standard deviation $\frac{|F_{k,2}(n_2) - F_2(n_2)|}{F_{k,2}(n_2)}$ for $R_1 \le r \le R$ and $1 \le n_2 \le 15$ realizing that the maximum deviation from the real value is approximately 0.15 at $n_2 = 15$ and $r = R$. We will further validate this approximation through simulation.

A. The Processor Sharing Model

Users arrive as a Poisson process of intensity $\lambda\cdot ds$ in any area of surface ds. Flow sizes are i.i.d and σ is the corresponding random variable. We denote by $\rho = \lambda\cdot\mathbb{E}[\sigma]$ the traffic density and by $d\rho(r) = \rho\cdot 2\pi r dr$ the traffic intensity generated by users whose distance to the BS ranges between r and $r + dr$. The maximum number of simultaneously admitted users in **Zone x** will be limited to $n_{Max,x}$ in order to guarantee a minimum rate which is a QoS notion appropriate for non-real time users. New transfers generated in a zone where they are already $n_{Max,x}$ transfers in progress are blocked and lost. Based on the above analysis, we see that, in every zone, each user k is served at a fraction $F_x(n_x)$ of some constant $c_{k,x}$ whenever there are n_x active users in its zone:
In **Zone 1**

$$F_1(n_1) = \log_2(1 + SNR_I(n_1)) \quad (21)$$

$$c_{k,1} = W \cdot \mathbb{P}(A_1)$$

In **Zone 2**

$$F_2(n_2) = \frac{\int_0^\infty e^{-x_k}(1-e^{-x_k})^{n_2-1}\log_2(1+x_k \cdot SNR_{II})dx_k}{\log_2(1+SNR_{II})} \tag{22}$$

$$c_{k,2} = W \cdot \mathbb{P}(A_2)\log_2(1+SNR_k)$$

with $SNR_{II} = (\alpha/\gamma_{II} + \nu_{II} + h)^{-1}$

We see from (21) and (21) that $F_x(.)$ is an arbitrary positive function satisfying the subsequent constraints for $n_x \leq n_{Max,x}$:

$$0 \leq F_x(n_x) \leq \infty \text{ (I) and } n_x \cdot F_x(n_x) \leq \infty \text{ (II)}$$

Thus, every zone behaves like a GPS system with equal but time varying service allocation as users randomly enter and leave the system. As a consequence, we can directly apply the formulae in [8] regarding the stationary state distributions of our system with the assumption that users arrive according to a Poisson process. Such a system has the well-known insensitivity property which means that performance depends mainly on the load factor (and the maximum number of users in presence of an admission control policy), and not on the distribution of the flow size which is continually changing given the ever varying nature of data applications. Hence, our scheduling approach, in addition to increasing the overall throughput of the cell, relieves it from dimensioning issues.
The stationary distribution of the number of active users:

In **Zone x**:

$$\pi_x(n) = \frac{\frac{\rho_x^n}{\psi_x(n)\cdot n!}}{\sum_{k=0}^{n_{Max,x}} \frac{\rho_x^k}{\psi_x(k)\cdot k!}}$$

where $\psi_x(n) = \prod_{i=1}^{n} F_x(i)$ and ρ_x is the load in **Zone x**:

with $\rho_1 = \int_\epsilon^{R_1} \frac{d\rho(r_k)}{c_{k,1}} = \frac{\rho\pi(R_1^2-\epsilon^2)}{W\mathbb{P}(A_1)}$ and $\rho_2 = \int_{R_1}^{R} \frac{d\rho(r_k)}{c_{k,2}} = \frac{\rho\pi}{W\mathbb{P}(A_2)}\int_{R_1}^{R} \frac{2r_k dr_k}{\log_2(1+SNR_{k,2})}$

Using Little's law, we find that the flow throughput $Th_{k,x}$ of user k in **Zone x**, defined as the ratio of the mean flow size $\mathbb{E}[\sigma]$ to the mean flow duration, is given by:

$$Th_{x,k} = c_{x,k} \cdot \frac{\rho_x\,(1-B_x)}{\mathbb{E}[n_{MAX,x}]}$$

where, in **Zone x**, $B_x = \pi_x\,(n = n_{Max,x})$ is the blocking probability and $\mathbb{E}[n_{Max,x}] = \sum_{i=1}^{n_{Max,x}} i \cdot \pi_x(i)$ is the mean number of active users.

B. Simulation Results

We present here our numerical experiments performed to illustrate the above results. We consider a system where users initiate file transfer requests as a Poisson process of intensity $\lambda_T = \lambda\pi R^2$ and traffic demand is uniformly distributed in the cell. At most 15 users are admitted simultaneously in each zone ($n_{Max,x} = 15$). The system operates in a time-slotted fashion with a slot duration equal to 2ms. We take $W = 1.0$, $h = 0.03$, $f = 0.1$ and $\epsilon = 1.0$. Flow sizes are independent and exponentially distributed with normalized mean equal to 20. Users are served according to our hierarchical approach of PF where each zone is served at a time ($\mathbb{P}(A_1) = \mathbb{P}(A_2) = 1/2$). We obtain respectively from (18) and (20), $\alpha/\gamma_I + \nu_I \approx 0.05$ and $\alpha/\gamma_{II} + \nu_{II} \approx 0.5$. The divergence in the preceding two values show the difficulty to give fitting mean values for the parameters γ_k and ν_k when we consider the entire cell. This explains why we failed to find an appropriate analytical model for the Standard PF as it covers the whole cell.
We determine the Normalized Throughput per user $T_{x,k}/c_{x,k}$ and display the Average Normalized Throughput for users belonging to the same zone as a function of $\frac{1}{\lambda_T}$.
The simulation results obtained are compared to the analytical results of subsection IV-A. For **Zone 1**, we run two sets of simulations to show that the proposed GPS model is still valid in spite of the two restrictive hypothesizes considered in section IV-A: in the first set of simulations, fading is omitted, while in the second set, we consider a Rayleigh fading channel. However, in two sets of experiments the parameters γ_I and ν_I are not taken as constants as in the analytical model but vary with distance. In the simulations we run for **Zone 2**, the parameters γ_{II} and ν_{II} are not taken as constants as in the analytical model but vary with distance. Figures 8 and 9 depict the Average Normalized Throughput as a function of $\frac{1}{\lambda_T}$.

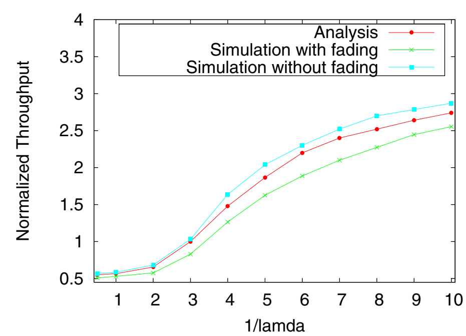

Fig. 8
THROUGHPUT AS A FUNCTION OF $\frac{1}{\lambda_T}$ IN **ZONE 1**

In Figure 8, we can see how the analytical results are very close from the simulation results when fading is omitted. Nevertheless, when fading is taken into consideration the simulation results follow the same trend as the analytical results and are reasonably close. In Figure 9, we can see how the analytical formulae provide highly accurate estimates of the simulation results especially at moderate to high loads. Therefore, we can fairly consider that the proposed analytical model approximates very well the behaviour of our HPF.

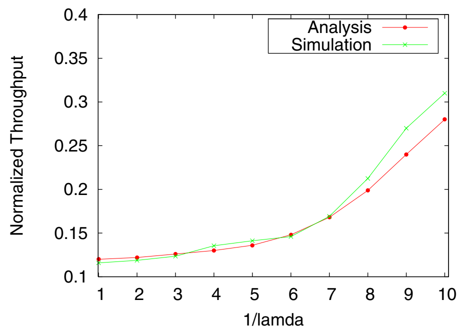

Fig. 9

THROUGHPUT AS A FUNCTION OF $\frac{1}{\lambda_T}$ IN **ZONE 2**

V. CONCLUSION

This work suggested decoupling the cell into two zones and serving each zone in a WRR fashion using an appropriate scheduler. We chose to apply the PF algorithm for Zone 2 because the SNR is almost proportional to the feasible rate of users belonging to this zone and because of the moderate to bad conditions of its channel which requires serving one user at a time to cancel intra-cell interference. Whereas, for Zone 1 we chose to serve its users in a CDMA fashion because of the logarithmic relation between their SNR and their feasible rate and because of the good channel condition they experience. We proved that our scheduler globally increased the performance in terms of average rates besides the fact that it is completely scalable which frees it from dimensioning issues.

REFERENCES

[1] Bender P., Black P., Grob M., Padovani R., Viterbi A., CDMA/HDR: A Bandwidth-Efficient High-Speed Wireless Data Service for Nomadic users, IEEE Communications Magazine, pp. 70-77, 2000.

[2] Love R., Gosh A., Nikides R., Jalloul L., Cudac M., High-Speed Downlink Packet Access Performance, Proceedings of IEEE VTC Spring, 2001.

[3] Bedekar A., Borst S., Ramanan K, Downlink Scheduling in CDMA data Networks, Proceedings of IEEE Globecom, 1999.

[4] Kelif J., Altman E., Downlink fluid model of CDMA networks, VTC 2005 Spring.

[5] Viterbi A.J., CDMA: Principles of Spread Spectrum Communication, Addison-Wesley, 1995.

[6] Technical Specification Group (TSG) RAN WG4; RF System Scenarios, 3G TR 25.942, V2.2.1.

[7] Kushner H.J., Convergence of Proportional-Fair Sharing Algorithms Under General Conditions, IEEE Transactions on Wireless Communications, Vol.3, No.4, 2004.

[8] Cohen J.W., The multiple phase service network with generalized processor sharing. Acta Informatica 12, 245-284.

[9] Holma H., Toskala A., WCDMA for UMTS: Radio access for third generation mobile communications, John Wiley & Sons, 2002.

[10] Sipilä K., Honkasalo Z., Estimation of capacity and required transmission power of WCDMA downlink based on a downlink pole equation, VTC Spring 2000.

[11] OS WP Space Time Coding US 22-Aug-05.

[12] Zhao Y., Standardization of mobile phone positioning for 3G systems, IEEE Comm. Mag., vol. 40, no. 7, pp. 108-116, July 2002.

[13] Crohn I., Bonek E., Modeling og Intersymbol-Interference in a Rayleigh Fast Fading Channel with Typical Delay Power Profiles, IEEE Transaction on Vehicular Technology, vol. 41, no. 4, pp. 438-447, November 1992.