Cellular network with continuum priority set

Jean-Marc Kelif
France Telecom Research and Development
92794 Issy Moulineaux, France
jeanmarc.kelif@orange-ftgroup.com

Eitan Alman
INRIA BP 93
06902 Sophia –Antipolis, France
Eitan.Altman@sophia.inria.fr

***Abstract*— We consider the following problem of spatial downlink prioritization. Mobiles arrive at a cell at locations that are determined according to some probability distribution. The further a mobile is from the base station, the weaker is its received power and thus the lower is the transmission rate to it. Beside this uncontrollable phenomenon that differentiates between mobiles according to their location, one can design other controlled mechanisms that differentiate between them. We analyse various priority policies where the assigned priority is given in terms of the distance of the mobiles from the base station. This gives rise to a whole continuum of priority levels. We study the influence that the combined location density and priority policy have on the quality of service of the mobiles and on the network overall performance. Applying our model to a HSDPA system, we calculate a quality of service indicator, the sojourn time, using a priority scheduling strategy, a processor sharing one and a first come first served one. Considering three types of arrival flow, a uniform one, a non uniform one and a flow which generates a constant load in the cell, we show the sojourn time depends on the adopted strategy, but also on the location of the mobile and on the arrival flow type. In particular, a numerical study based on our model shows that a maximum SIR priority does not provide in any case the minimum sojourn time.**

Keywords: priority, scheduling, continuum, HSDPA

INTRODUCTION

Radio communications systems manage more and more mobile data traffic. The available bandwidth being limited, different strategies are adopted by operators to share the limited resources, and to schedule the use of radio channels. We consider in this paper the scheduling of downlink traffic and we study the performance of the subclass of fixed priority policies, where priority rules are defined as a function of the distance of a mobile from the base station. This gives rise to a whole continuum of priority levels. We study the influence that the combined location density and priority policy have on the quality of service of the mobiles and on the network overall performance. This gives rise to a whole continuum of priority levels.

The first part of the paper is devoted to the study of the resulting priority queue model. We derive expressions for the expected sojourn times of connections as a function of the location of the mobile. We furthermore obtain conservation laws for the workload in the system for a whole class of work conserving scheduling policies. We then use these tools to study the influence that the combined location density and priority policy have on the quality of service of the **mobiles** and on the network overall performance. We study the quality of service (QoS) of a HSDPA system, considering three types of arrival flow, a uniform one, a non uniform one and a flow which generates a constant load density in the cell. We compare the sojourn time under various scheduling policies: a priority scheduling strategy (P), a processor sharing one (PS) and a first come first served one (FCFS). We analyze the expected sojourn time using a maximum SIR priority (SIR_{max}) and a minimum SIR priority (SIR_{min}). We particularly determine that a SIR_{max} priority does not provide *in any case* the minimum expected sojourn time.

Related work on scheduling policies. We briefly mention some other policies that have been proposed for downlink scheduling for HSDPA. The Fair algorithms are built on the share of the resources in a fairly way to the users present in the cell. In the Efficient algorithms (maximum SIR)[13], the resources are allocated to the mobile with the best instantaneous link quality, and so the throughput of the cell is maximized at each time. The Efficient-fair algorithms [13] do a compromise between efficiency and fairness, in order to overcome the drawbacks of the previous methods. The principle of the Proportional-Fair algorithm [14] [15] is based on the transmission to the user with the highest data rate relative to its current average data rate. The Score-Based algorithm [16] keeps track of the last n values of the feasible rate for each user and then selects the user with the best score, that is, with the best position. To decide which strategy will be adopted to share the transmission between a great number of consumers, a queueing priority analysis can be proposed. Different kind of priority analyses may be done: for example, the system can be considered as static or dynamic, the service discipline can be preemptive or non-preemptive. In data networks, a priority discipline has to improve the mean delay or to satisfy the stringent delay requirement of delay-sensitive traffics.

Related work on priority rules. Related works on the subject mainly consider a discrete limited set of priorities. The channel allocation mechanism proposed in [7] is based on two priority schemes, a high one for handoff calls and a low one for new calls. To reduce call blocking and failure in a mobile cellular network, a dynamical priority strategy is proposed in [8], for carrier allocation. In [9], the authors present an analytical framework for dynamic priority queueing of handover calls in wireless networks, with two classes of priority. In [10], the authors propose a combined preemptive or nonpreemptive priority discipline. In [11], the authors consider a priority queueing system with two different types of traffic, modelled by continuous fluid flows, high and low priority. They consider a simple priority queueing system.

I. PRIORITY MODEL

Consider a marked point process $\{T_n, \nu_n, X_n\}$ where T_n is the time when the n_{th} arrival occurs, $X_n \in [0,1]$ indicates the class to which belong the n_{th} arriving customer and $\nu_n \geq 0$ denotes its service requirement. A customer of class x has priority level $q(x)$.
Let $E[S_x]$ denote the expected service time required by an arrival of priority x. The arrival class of a customer is chosen according to some general distribution $F_x(c)=P(X\leq c)$. We define the workloads

$$\overline{\rho}_x = \lambda \int_{[0,x]}^{x} E[S_x] F_x(d\xi), \quad (1.1a)$$

and

$$\rho_x = \lambda \int_{[0,x)} E[S_x] F_x(d\xi), \quad (1.1b)$$

Define the priority distribution $F_q(c) = P(q(X) \leq c)$. (1.2)

II. MODEL: M/G/1 NON-PREEMPTIVE HOL QUEUE

Consider arrivals of customers according to a Poisson rate with parameter λ. The service requirement brought by a customer of class x is generally distributed (it may depend on x) with first moment $\overline{\sigma}_x$ and distribution Σ_x. The arrival priority sequence is *i.i.d*. We wish to compute the expected waiting $W(x)$ time of some tagged customer of priority level x. We proceed similar to [12] [Vol. 2, Chapter 3]. It is the sum of
(i) the expected residual service time W_{res} of the customer in service,
(ii) the expected service time $W_{present,c}$ of all the customers of priority larger than or equal to c that were in the system when the tagged customer arrived, and
(iii) the expected service time $W_{future,c}$ of all the arrivals of customers of priorities higher than the tagged customers that arrive during the waiting time of the tagged customer.

We have as in [12]:

$$W_{res}=\lambda/2\int_{[0,1]} E[S_\xi^2]\, F_x(d\xi) \quad (2.1)$$

Note that W_{res} does not depend on the priorities. With M_q denoting the number of customers with priority q, we have by PASTA and by Little's law

$$W_{present,x} = \int 1_{\{q(\xi)\geq q(x)\}} \overline{\sigma}_\xi dE[N_\xi] = \int 1_{\{q(\xi)\geq q(x)\}} W(\xi) d\rho_\xi \quad (2.2)$$

and finally

$$W_{future,x} = W(x)\int 1_{\{q(\xi)\geq q(x)\}} d\rho_\xi \quad (2.3)$$

We conclude that

$$W(x) = W_{res} + \int 1_{\{q(\xi)\geq q(x)\}} W(\xi) d\rho_\xi + W(x)\int 1_{\{q(\xi)\geq q(x)\}} d\rho_\xi$$

We conclude that

$$W(x) = \frac{W_{res} + \int 1_{\{q(\xi)\geq q(x)\}} W(\xi) d\rho_\xi}{1 - \int 1_{\{q(\xi)\geq q(x)\}} d\rho_\xi} \quad (2.4)$$

whose solution is

$$W(x) = \frac{W_{res}}{\left(1 - \int 1_{\{q(\xi)\geq q(x)\}} d\rho_\xi\right)^2} \quad x\in[0,1] \quad (2.5)$$

III. APPLICATION: SCHEDULING IN HSDPA

The HSDPA (High Speed Downlink Packet Access) has been introduced in UMTS (3GPP specifications for Release 5) in order to adapt more dynamically the radio resource to the nature of packet switched traffic, and to offer a throughput transmission higher than in WCDMA [2][3][4]. The HSDPA is, by nature, well adapted to Non-Real Time (NRT) services (no delay constraint). The scheduling strategies represent one of the key factors of the QoS of that system.
We first consider a single cell with a unit radius and a single base station that transmits at its maximum power P. Mobiles arrive according to a Poisson process and their distance to the base station is determined according to the distribution F_x. The power gain between the base station and a mobile at a distance x is given by $h_x = \min(1, x^{-\eta})$ where η, the powerloss exponent, is typically between 2 and 3. We assume that mobiles are served according to the HOL non-preemptive priority scheme.

III.1 Transmission rate and service time

The transmission rate B_x to a mobile at a distance x depends on the signal to noise ratio. Specifically, we can consider the following models,

$$B_x = aPh_x/N_{th}, \quad (3.1)$$

where N_{th} is the thermal noise. *i.e.* it is linear in the signal to noise ratio. a is a constant

As a special case, we may assume that the sessions' sizes do not depend on the arrival location: Denoting S_x the service time (second) at the position x, and v_x the service requirement (bits), we can write: $v_x = v$, and then the service time required for a file transfer is given as

$$S_x = \frac{v}{B_x} \quad (3.2)$$

III.2 Priority queues

We focus here on non-real time (NRT) data transfers. Whereas all calls use CDMA, we assume that NRT calls are time-multiplexed (which diminishes the amount of interference, thus increasing the available average throughputs). This combination of time multiplexing over CDMA is typically for high speed downlink data channels, such as the High Speed Downlink Packet Access (HSDPA) [4] and the High Data Rate (HDR) in CDMA-2000 systems [1].

We apply the model developed in section II. Considering independent mobiles arrivals in a cell, each mobile arrives at a given position *x* from its serving BS. And each mobile has to be served with a given priority depending on its distance from the BS. Let's consider that mobiles close to the BS have a higher priority than the ones far from it: the priority is a decreasing function of the distance.
The service requirement is distributed according to a distribution *g* with a first moment denoted $E[v]$ and a second moment denoted $E[v^2]$. We can write the first and second moment of the service time as:

$$E[S_x] = \int_0^\infty S_x g(u)du = \frac{1}{B_x}\int_0^\infty v(u)g(u)du = E[v]B_x^{-1} \qquad (3.3)$$

and

$$E[S_x^2] = \int_0^\infty S_x^2 g(u)du = \frac{1}{B_x^2}\int_0^\infty v^2(u)g(u)du = E[v^2]B_x^{-2} \qquad (3.4)$$

We consider a network and denote P_i the total transmitting power of a BS *i* of the network. P_{CC} represents the power dedicated to the common channels, α is the orthogonal factor, N_{th} is the thermal noise, h_x^i the pathloss between the mobile located at *x* in a given cell 0 of the network and the BS *i*. The transmitting power dedicated to a mobile belonging to the cell 0 is the total available power $P_0 - P_{CC}$ of the serving base station BS_0. And the noise *N* is due to the common channels of the BS_0, the power transmitted by the N_{BS} other base stations of the network and the thermal noise. The expression (3.1) can be written as:

$$B_x = \frac{aPh_x^0}{N} = a\frac{(P_0 - P_{CC})h_x^0}{\alpha P_{CC}h_x^0 + \sum_{i=1}^{N_{BS}} P_i h_x^i + N_{th}} \qquad (3.5)$$

Introducing the interference factor *f(x)*, for a homogeneous network (all the base stations transmit at the same power), we can write:

$$f(x) = \frac{1}{P_0 h_x^0}\sum_{i=1}^{N_{BS}} P_i h_x^i = \frac{1}{h_x^0}\sum_{i=1}^{N_{BS}} h_x^i \qquad (3.6)$$

and $\varphi = \frac{P_{cc}}{P_0}$ we have

$$B_x = a\frac{1-\varphi}{\alpha\varphi + f(x) + \frac{N_{th}}{Ph_x^0}} \qquad (3.7)$$

The term with thermal noise is very low compared to the other interferences so we can write, denoting $B(x) = B_x$:

$$B(x) = a\frac{1-\varphi}{\alpha\varphi + f(x)} \qquad (3.8)$$

III.3 Load of the system
We denote *F(x)* the location arrival distribution of the mobiles (at the position *x*). From (1.1b) and (3.3), we can express the load density as:

$$d\rho(x) = \lambda E[v]\frac{F(x)}{B(x)}dx \qquad (3.9)$$

Using (3.8) it can be written:

$$d\rho(x) = \lambda E[v](f(x) + \alpha\varphi)\frac{1}{a(1-\varphi)}F(x)dx \qquad (3.10)$$

IV. PRIORITY SCHEDULING STRATEGY

We propose hereafter a priority scheduling strategy based on the SIR received by the mobile. Denoting λ the mobiles arrival rate in the cell, we consider a circular symmetry in the cell: *F(x)* can be written *F(r)* where *r* is the distance from the BS.

IV.1 Maximum SIR Priority

The priority is given to mobiles with the highest received SIR. From (3.8) since *f(.)* is an increasing function of the distance *r* (as established in [5] and in section V.4), the maximum SIR is received by the base station's closest mobiles. Consequently, the expected waiting time given by (2.5) can be written as $W_{max}(r)$ and is expressed as:

$$W_{\max}(r) = \frac{\frac{\lambda}{2}\int_0^{R_c} E\left[S_x^2\right]F(x)dx}{\left(1 - \int_0^r \lambda E[S_u]F(u)du\right)^2} \qquad (4.1)$$

We denote ω^2 the variance of the sessions sizes arrivals, $\omega^2 = E[v^2] - E[v]^2$, and R_c the cell radius. Using the expressions (3.3), (3.4) and (3.8) we finally can write

$$W_{\max}(r) = \frac{1}{2}\left(1 + \frac{\omega^2}{E[v]^2}\right)\frac{\int_0^{R_c}\frac{1}{a^2}(f(x) + \alpha\varphi)^2 F(x)du}{\lambda\left(\frac{1-\varphi}{\lambda E[v]} - \int_0^r \frac{1}{a}\left(f(u) + \alpha\varphi\right)F(u)du\right)^2} \qquad (4.2)$$

The expected waiting time is an increasing quadratic function of the sessions' size variability expressed by the standard deviation ω

IV.2 Minimum SIR priority

The highest priority is given to mobiles with the lowest received SIR. From (3.8) since *f(.)* is an increasing function of the distance *r*, the minimum SIR is received by the base station's furthest mobiles. Consequently, we can write the expected waiting time as a function of the distance r, using its expression given by (2.5), and denoting $W(r)$ as $W_{\min}(r)$ in this case:

$$W_{\min}(r)=\frac{\frac{\lambda}{2}\int_0^{R_c}E\left[S_x^2\right]F(x)dx}{(1-\int_r^{R_c}\lambda E[S_u]F(u)du)^2} \qquad (4.3)$$

IV.3 Conservation law

The conservation law is expressed as in [12]:

$$\sum_{i=1}^{P}\rho_i W_i=\frac{\rho}{1-\rho}W_{res} \qquad (4.4)$$

where W_{res} is the residual service time, W_i is the expected waiting time of the service i ρ_i is the load of the service i and ρ is the total load of the system. In our analysis we consider a continuum of priorities in the cell so the expression (4.4) can be written:

$$\int_0^{R_c}\rho(r)W(r)dr=\frac{\rho}{1-\rho}W_{res} \qquad (4.5)$$

where the load density (3.9) can be expressed:

$$d\rho(r)=\lambda E[v]\frac{F(r)}{B(r)}dr \qquad (4.6)$$

The expression (2.5), and consequently (4.1) and (4.2), was established for a conservative system: no work (*i.e.* service requirement) is created or destroyed within the system [12]. It appears important to verify that the system is conservative (see Annex).

IV.4 Stability condition

Considering the expression of the expected waiting time (2.5) and (4.1), the necessary stability condition of the queue can be written as

$$1-\int_0^{r}\lambda E[v]\left(f(u)+\alpha\varphi\right)\frac{1}{a(1-\varphi)}F(u)du>0$$

That condition means that the expected waiting time of any mobile has to be finite. It can be rewritten:

$$\frac{1-\varphi}{\lambda E[v]}>\int_0^{r}\frac{1}{a}\left(f(u)+\alpha\varphi\right)F(u)du$$

This stability condition expresses that when the amount of data arriving in the system is higher than the available transmitting amount of data, the system can no more answer the demand: the amount of data in the system increases indefinitely, there is no balance between the data arrival and their departure. In the case of a priority scheduling strategy based on a minimum SIR, using the expression (4.3), we have the following condition:

$$1-\int_{R_c}^{r}\lambda E[v]\left(f(u)+\alpha\varphi\right)\frac{1}{a(1-\varphi)}F(u)du>0 \qquad (4.7)$$

V. OTHER SCHEDULING STRATEGIES

It is interesting to compare the scheduling strategy based on priorities between mobiles, to other ones. We analyze the cases First Come First Served (FCFS), and Processor Sharing (PS).

V.1 FCFS case

Hereafter we develop the case "without priority" FCFS [12, chap.1]. We can write the mean expected waiting time as:

$$W_{FCFS}=\frac{\frac{\lambda}{2}\int_0^{R_c}E\left[S_x^2\right]F(u)du}{1-\int_0^{R_c}\lambda E\left[S_x\right]F(u)du}$$

Using (3.3) and (3.4) it becomes:

$$W_{FCFS}=\frac{\lambda}{2}\frac{E[v]^2}{a^2(1-\varphi)^2}\left(1+\frac{\theta^2}{E[v]^2}\right)\frac{\int_0^{R_c}(f(x)+\alpha\varphi)^2F(x)du}{\left(1-\int_0^{R_c}\frac{\lambda E[v]}{a(1-\varphi)}\left(f(u)+\alpha\varphi\right)F(u)du\right)} \qquad (5.1)$$

V.2 Processor Sharing case

In a Processor Sharing case, a mobile is served as soon as it enters the cell: the expected waiting time has no significance. The total capacity is equally distributed (in time) between the mobiles present in the cell: all the mobiles have the same access channel duration. In HSDPA, each mobile is scheduled alone. So we follow, for the mobiles, a Round Robin proportional scheduling allowing to each mobile the same channel duration [6]. In this case, we consider the mean sojourn time T of a mobile in the system. That one only depends on the mean service time and the load of the system, and can be written $T=\frac{E(S_x)}{1-\rho}$. Using the expressions (3.3), (3.9) and (3.10) we have

$$T=\frac{\int_0^{R_c}E\left[S_u\right]F(u)du}{1-\int_0^{R_c}\lambda E\left[S_u\right]F(u)du}$$

and finally

$$T=\frac{\int_0^{R_c}\frac{1}{a}(f(u)+\alpha\varphi)F(u)du}{\left(\frac{(1-\varphi)}{E[v]}-\lambda\int_0^{R_c}\frac{1}{a}\left(f(u)+\alpha\varphi\right)F(u)du\right)} \qquad (5.2)$$

V.3 Expressions of the three scheduling strategies

We denote

$$I_N = \int_0^{R_c} (f(x)+\alpha\varphi)^2 F(x)dx \qquad (5.3)$$

and

$$I_D(r) = \int_0^{r} (f(x)+\alpha\varphi) F(x)dx \qquad (5.4)$$

We can rewrite the expressions of the expected waiting times and the sojourn times as follows.

V.3.1 Priority case

Expected waiting time:

$$W(r) = \frac{\lambda}{2}\frac{E[v]^2}{a^2(1-\varphi)^2}\left(1+\frac{\omega^2}{E[v]^2}\right)\frac{I_N}{\left(1-\frac{\lambda E[v]}{a(1-\varphi)}I_D(r)\right)^2} \qquad (5.5a)$$

Sojourn time:

$$T(r) = W_{\max}(r) + \frac{E[v]}{B(r)} \qquad (5.5b)$$

V.3.2 FCFS case

Expected waiting time:

$$W_{FCFS} = \frac{\lambda}{2}\frac{E[v]^2}{a^2(1-\varphi)^2}\left(1+\frac{\omega^2}{E[v]^2}\right)\frac{I_N}{1-\frac{\lambda E[v]}{a(1-\varphi)}I_D(R_c)} \qquad (5.6a)$$

Sojourn time

$$T(r) = W_{FCFS} + \frac{E[v]}{B(r)} \qquad (5.6b)$$

V.3.3 Processor Sharing case

Sojourn time:

$$T = \frac{\frac{E[v]}{a(1-\varphi)}I_D(R_c)}{1-\frac{\lambda E[v]}{a(1-\varphi)}I_D(R_c)} \qquad (5.7)$$

Considering the whole cell, $r=R_c$, we observe that the stability condition (denominator >0) is the same for the three cases. In the priority case however, there is a singularity degree of order 2 and for the other cases the singularity is of order 1.

V.4 Analytical Model

Using the model developed in [5] the expression (5.4) of the factor *f(x)* can be written only depending on the distance *r*: $f(x) = f(r)$ as:

$$f(r) = \frac{\rho_{BS}.2\pi r^{\eta}}{(2-\eta)}\left[(kR_c - r)^{2-\eta} - (2R_c - r)^{2-\eta}\right] \qquad (5.8)$$

Where ρ_{BS} is the BS density, R_c is the cell radius and $k\,R_c$ characterizes the size of the network.

We consider an infinite network $k >> 2$ and $\rho_{BS} = \frac{1}{\pi R_c^2}$

$$f(r) \approx \frac{8}{\eta-2}\left(\frac{r}{R_c}\right)^{\eta}\left(1-\frac{r}{2R_c}\right)^{2-\eta} \qquad (5.9)$$

Considering η=3, and $u = r/R_c$ we have:

$$f(r) \approx 8\frac{u^3}{(1-u/2)} \qquad (5.10)$$

VI. Scheduling strategies comparison

We will compare the expected waiting times and the sojourn times using three scheduling strategies, maximum SIR (SIR_{max}) priority, FCFS, and Processor Sharing (PS). For each strategy we will assume three kinds of arrival distributions in the cell: a uniform one (UA), a non-uniform one (denoted NA), and an arrival for which the cell's load is a constant (denoted CL). Considering the expressions (5.5a) (5.6a) and (5.7), to analyze the different cases, it is sufficient to calculate the parameters

$$I_N = \int_0^{R_c} (f(x)+\alpha\varphi)^2 F(x)dx \qquad (6.1)$$

and
$$I_D(r) = \int_0^{r} (f(x)+\alpha\varphi) F(x)dx \qquad (6.2)$$

VI.1 Uniform location arrival probability

We consider an arrival probability equivalent at any location of the cell, so we have $F(r,\theta) = \frac{\lambda}{\pi R_c^2}$

In this case we can write

$$I_N = \int_0^1 2u\left(8\frac{u^3}{1-\frac{u}{2}}+\alpha\varphi\right)^2 du \qquad (6.3)$$

and
$$I_D(r) = \int_0^{r/R_c} 2u\left(8\frac{u^3}{1-\frac{u}{2}}+\alpha\varphi\right)du \qquad (6.4)$$

VI.2 Non uniform arrival probability

The traffic can be non uniform. In this case we can consider a non uniform arrival probability $\lambda(r) = \lambda F(r)$, where for example $F(r) = \kappa \exp(-br^2)$, where κ and b are constant and the normalisation condition is $\int_0^{2\pi}\int_0^{R_c} F(r) r dr d\theta = 1$. So $F(r,\theta) = \frac{b}{\pi}\frac{1}{1-\exp(-bR_c^2)}\exp(-br^2)$ represents the probability density to enter the cell at the point $M(r,\theta)$ and $F(r) = b\dfrac{2r}{1-\exp(-bR_c^2)}\exp(-br^2)$ is the probability density to enter the cell at the distance *r*.

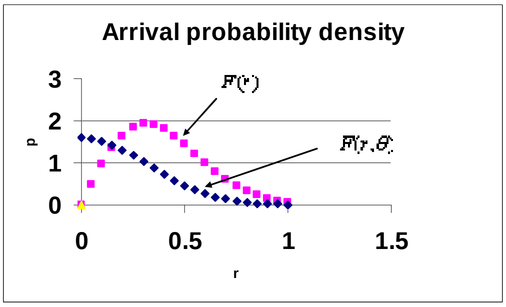

Figure 1: Arrival probability density in the cell

We thus can write (5.3) and (5.4) as

$$I_N = \frac{2b}{1-\exp(-bR_c^2)}\int_0^{R_c}\left(8\frac{\left(\frac{r}{R_c}\right)^3}{1-\frac{r}{2R_c}}+\alpha\varphi\right)^2 r\exp(-br^2)dr \qquad (6.5)$$

$$I_D(r) = \frac{2b}{1-\exp(-bR_c^2)}\int_0^{r}\left(8\frac{\left(\frac{r}{R_c}\right)^3}{1-\frac{r}{2R_c}}+\alpha\varphi\right) r\exp(-br^2)dr \qquad (6.6)$$

VI.3 Constant load in the cell

The provider may adopt a scheduling strategy based on a constant load density in the cell. From the load expression (3.10) we write: $\lambda E[v](f(x)+\alpha\varphi)\dfrac{1}{a(1-\varphi)}F(x)dx = Kdx$

where *K* is a constant. So we have the condition

$$F(x) = K\left(\lambda E[v](f(x)+\alpha\varphi)\frac{1}{a(1-\varphi)}\right)^{-1}$$

And the normalisation condition is $\int_0^{R_c} F(x)dx = 1$. So we can write:

$$F(x) = \frac{1}{I_B}\left(f(x)+\alpha\varphi\right)^{-1}$$

where $$I_B = \int_0^{R_c}\frac{du}{\alpha\varphi + f(u)}$$

We finally can write

$$I_N = \frac{1}{I_B}\int_0^{R_c}(f(u)+\alpha\varphi)du \qquad (6.7)$$

and $$I_D(r) = \frac{r}{I_B} \qquad (6.8)$$

Remark: The stability condition (4.7) can be written as

$$I_D(R_c) < \frac{a(1-\varphi)}{\lambda E[v]} \qquad (6.9)$$

VII. NUMERICAL APPLICATION

We consider a unit cell radius R_c = 1, an orthogonality factor α=0.8 and φ=0.2, and $\dfrac{E[v]}{a} = 0.1s$. For the non uniform case we consider a value *b*=5.

We notice that considering the three strategies, we have, from the expressions (6.4), (6.6) and (6.8):

Uniform case	$I_D(R_c) = 6.20$
Non-uniform case	$I_D(R_c) = 0.19$
Constant load case	$I_D(R_c) = 1.45$

The less restrictive stability condition (6.9) is obtained with the non-uniform case.

We first consider a variance equal to zero: $E[v]^2 = E[v^2]$. To generalize the results, it will be sufficient to calculate the influence of the variance, using the expression (4.2).

VII.1 Expected Waiting Times comparison

The figures 2a, 2b and 2c show the expected waiting time variations with the distance. They allow to compare two scheduling approaches, the maximum SIR priority one and the FCFS one, in the case of uniform (fig. 2a), non uniform (fig. 2b) mobiles arrivals in the cell, and also mobiles arrival such as the cell's load density is constant (fig. 2c). We first observe that the expected waiting time increases with the distance. For uniform arrivals, this increase becomes very high close to the edge of the cell (fig. 2a). In the case of non uniform and "constant load" arrivals, this increase is relatively slow and linear when the distance increases (fig. 2b and fig.2c). We observe that the priority scheduling strategy is better than the FCFS one, until a given distance d_{th} depending on the mobiles arrival distribution: d_{th} =0.9 for uniform arrivals, d_{th} =0.6 for non uniform ones, and d_{th} =0.5 for "constant load arrival". Moreover, the non uniform and constant load cases (fig. 2b and 2c) give expected waiting times very low compared to the uniform one (fig.2a)

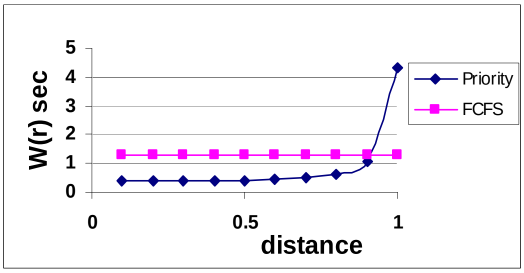

Figure 2a: Expected waiting time with Uniform arrival distribution λ=1

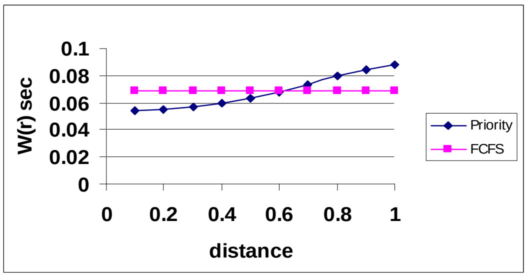

Figure 2b: Expected waiting time with
Non uniform arrival distribution λ=1

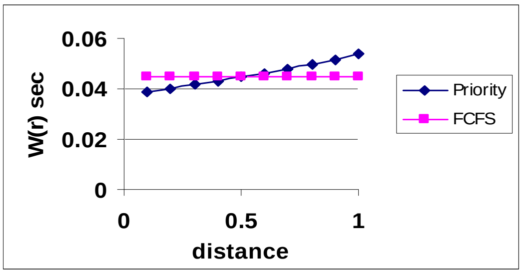

Figure 2c: Expected waiting time with
Constant load arrival distribution λ=1

VII.2 Sojourn Times Comparisons

Fig 3a, 3b and 3c show that the sojourn time is mainly better (i.e. lower except at the edge of the cell) in the priority case and the FCFS' one than in the processor sharing (PS) case, for uniform, non uniform and constant load mobiles arrivals.

Remark: the figure 3d shows the influence of the variability of the session' sizes, for different distances *d* from the BS. For almost the whole cell, when the distances are lower than 0.9, we observe that for $\frac{\omega}{E[v]} < 1$, the priority scheduling strategy gives lower sojourn times than the PS one. This result is moreover observed when the standard deviation reaches the value 2 for distances lower than 0.5.

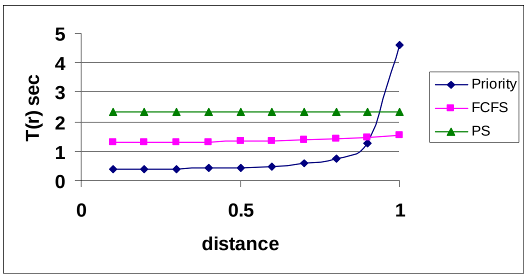

Figure 3a: Sojourn time with
Uniform arrival distribution λ=1

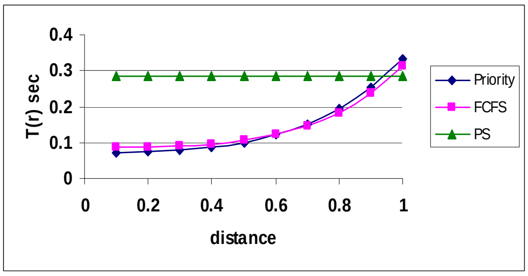

Figure 3b Sojourn time with
Non uniform arrival distribution λ=1

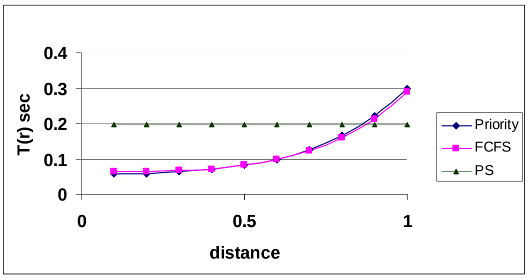

Figure 3c Sojourn time with
Constant load arrival distribution λ=1

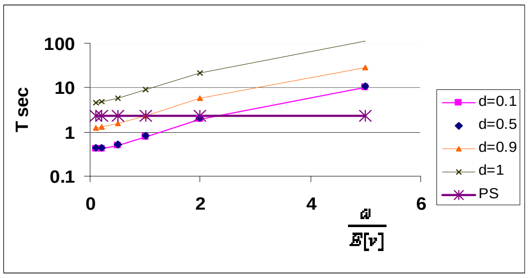

Figure 3d: Sojourn time vs standard deviation of the session' size with uniform arrival λ=1

VII.3 Arrival rate influence

VII.3.1 Expected waiting time

Fig 4a and Fig 4b show the expected waiting time increases with the arrival rate. The priority case remains better than FCFS' one.

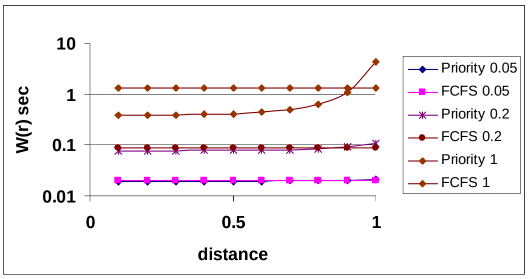

Figure 4a: Expected waiting time with
Uniform arrival distribution λ=0.05, 0.2, 1

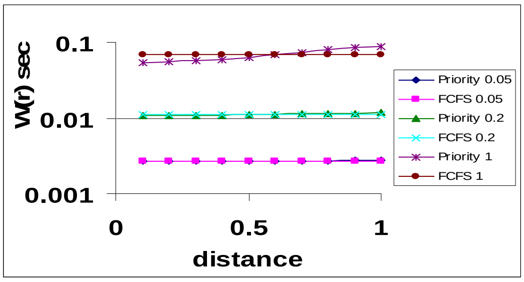

Figure 4b Expected waiting time with
Non uniform arrival distribution λ=0.05, 0.2, 1

VII.3.2 Sojourn time

For non uniform arrivals, Fig 5a and 5b show the sojourn time is better in the priority case than in PS and FCFS ones, even for high arrival rates, until a distance of 0.8.

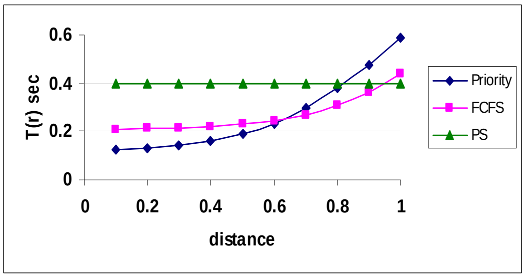

Figure 5a: Sojourn time with
Non uniform arrival distribution λ=2

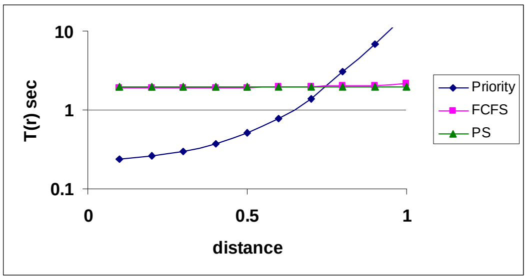

Figure 5b: Sojourn time with
Non uniform arrival distribution λ=4

VII.4 Hybrid scheduling strategy

In the case of non uniform arrivals, we observe (Fig 6a and 6b) the expected waiting time is lower using a scheduling based on a maximum SIR priority (SIR_{max}) than a scheduling based on minimum SIR priority (SIR_{min}) for distances lower than 0.6. This is a well-known result. We however also observe that this last strategy is better than the first one in term of expected waiting time when the distance is higher than 0.6: the expected waiting time becomes lower. It could be interesting to adopt a SIR_{min} priority scheduling strategy, in some given cases.

This conclusion could drive us to adopt a "hybrid" strategy, dividing the cell into two zones: a first one, close to the BS until a given distance Z_{th}, and a second one far from the BS, beginning at Z_{th} and ending at the edge of the cell. In the first one, a SIR_{max} scheduling strategy would be adopted and in the second one a SIR_{min} strategy would be adopted.

We can compare these strategies by using the expected sojourn time *T* of a mobile in the cell, whatever its position, as an indicator (Table 1). Considering the first cell's zone until a distance Z_{th}=0.6, and the second one until the edge of the cell, we observe (Table 1) that the "hybrid" strategy can be better in terms of expected sojourn time than a SIR_{max} strategy.

	T (sec)	
	λ=2	λ=4
SIR_{max}	0.18	0.75
SIR_{min}	0.28	3.92
Hybrid	0.16	0.39

Table 1: Expected Sojourn Time

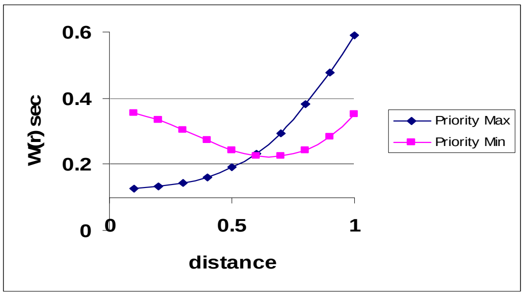

Figure 6a: Expected waiting time with
Non uniform arrival distribution λ=2

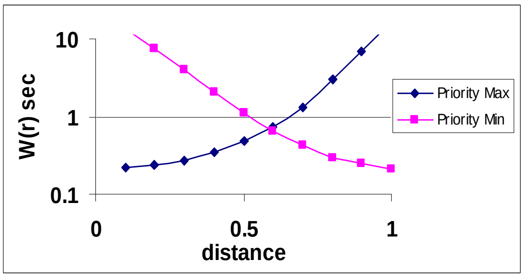

Figure 6b: Expected waiting time with
Non uniform arrival distribution λ=4

The figure 7a, 7b and 7c show the sojourn time probability densities, for different arrival rates, when the mobiles arrivals are uniform. For a low arrival rate, λ=0.1, the priority scheduling based on SIR_{max} and SIR_{min} are equivalent. When the arrival rate increases, the behaviour of the curve describing the priority SIR_{min} is modified. For a high value of λ we observe that the sojourn times probability density comprised between 0.8 and 1 is higher for a scheduling based on a SIR_{min} priority: it can be interesting, in that last case, to adopt a SIR_{min} priority scheduling strategy than a SIR_{max} one.

VIII. CONCLUSION

We considered a continuum of priority levels to analyse various priority policies where the priority is given in terms of the distance between mobiles and their serving base station. We established the influence that the combined location density and priority policy have on the expected waiting time and on the sojourn time. Applying our analysis to a HSDPA system, we compared a priority scheduling strategy to a processor sharing one and a FCFS one. Considering three types of arrival flow, a uniform one, a non uniform one and a flow which generates a constant load density in the cell, we showed the sojourn time depends on the adopted strategy, on the location of the mobile and on the arrival flow type. We moreover showed that a SIR_{max} priority scheduling strategy does not provide in any case the lowest expected sojourn time: dividing the cell into two zones and considering a non-uniform arrival of mobiles, we showed a "hybrid" strategy combining a SIR_{max} scheduling strategy and a SIR_{min} one allows to obtain a lower expected sojourn time.

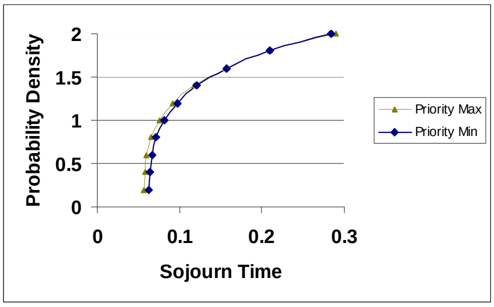

Figure 7a: Sojourn time Probability density with Uniform arrival distribution λ=0.1

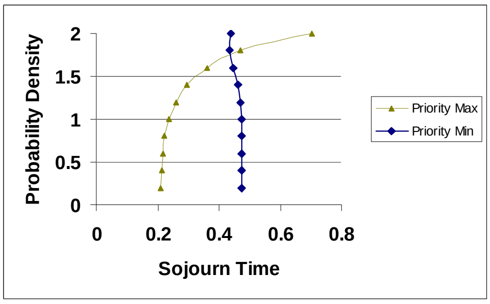

Figure 7b: Sojourn time Probability density with Uniform arrival distribution λ=0.5

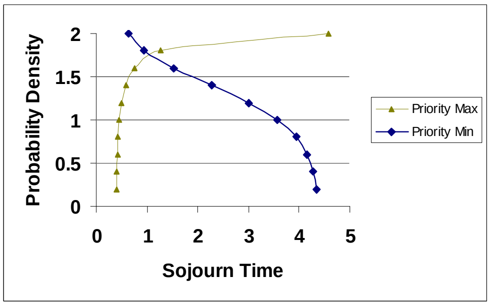

Figure 7c: Sojourn time Probability density with Uniform arrival distribution λ=1

ANNEX: CONSERVATION LAW

Maximum SIR Priority

Using (2.1) (3.3) and (3.4), the expected waiting time (4.1) can be written:

$$W_{\max}(r) = \frac{W_{res}}{(1-\lambda E[v]\int_0^r \frac{F(u)}{B(u)}du)^2} \qquad \text{(A-1)}$$

So we have

$$\int_0^{R_c} \rho(r)W_{\max}(r)dr = \int_0^{R_c} \frac{\lambda E[v]\frac{F(r)}{B(r)}}{(1-\lambda E[v]\int_0^r \frac{F(u)}{B(u)}du)^2} W_{res}dr \qquad \text{(A-2)}$$

Remark: We can denote the following key change of variables (A-3) and (A-4)
:

$$D(r) = \lambda E[v]\int_0^r \frac{F(u)}{B(u)}du \qquad \text{(A-3)}$$

and

$$\frac{\lambda E[v]\frac{F(r)}{B(r)}}{(1-\lambda E[v]\int_0^r \frac{F(u)}{B(u)}du)^2} = \frac{D'(r)}{(1-D(r))^2} \qquad \text{(A-4)}$$

Consequently we can write

$$\int_0^{R_c} \rho(r)W_{\max}(r)dr = \int_0^{R_c} \frac{D'(r)}{(1-D(r))^2}W_{res}dr,$$

$$= W_{res}\int_0^R \left(\frac{1}{1-D(r)}\right)' dr \qquad \text{(A-5)}$$

So we have

$$\int_0^{R_c} \rho(r)W_{\max}(r)dr = \left(\frac{1}{(1-D(R_c))} - \frac{1}{(1-D(0))}\right)W_{res} \qquad \text{(A-6)}$$

We notice that

$$D(R_c) = \lambda E[v]\int_0^{R_c} \frac{F(u)}{B(u)}du = \rho \qquad \text{(A-7)}$$

and

$$D(0) = \lambda E[v]\int_0^0 \frac{F(u)}{B(u)}du = 0 \qquad \text{(A-8)}$$

Finally we can write the conservation law

$$\int_0^{R_c} \rho(r)W_{\max}(r)dr = \left(\frac{1}{(1-\rho)} - 1\right)W_{res} = \frac{\rho}{1-\rho}W_{res} \qquad \text{(A-9)}$$

Minimum SIR priority

In a similar approach, when the highest priority is given to the base station's furthest mobiles (with the minimum SIR), we can write:

$$W_{\min}(r) = \frac{W_{res}}{(1-\lambda E[v]\int_r^{R_c} \frac{F(u)}{B(u)}du)^2} \qquad \text{(A-10)}$$

We can write

$$\lambda E[v]\int_r^{R_c} \frac{F(u)}{B(u)}du = \lambda E[v]\int_0^{R_c} \frac{F(u)}{B(u)}du - \lambda E[v]\int_0^{r} \frac{F(u)}{B(u)}du$$

So we have

$$\lambda E[v]\int_r^{R_c} \frac{F(u)}{B(u)}du = \rho - \lambda E[v]\int_0^{r} \frac{F(u)}{B(u)}du$$

The expression (A-10) can be written

$$W_{\min}(r) = \frac{W_{res}}{(1-\rho+\lambda E[v]\int_0^{r} \frac{F(u)}{B(u)}du)^2} \qquad \text{(A-11)}$$

So we can write

$$\int_0^{R_c} \rho(r) W_{\min}(r)dr = \int_0^{R_c} \frac{\lambda E[v]\frac{F(r)}{B(r)}dr}{(1-\rho+\lambda E[v]\int_0^{r} \frac{F(u)}{B(u)}du)^2} W_{res}$$

So we have

$$\int_0^{R_c} \rho(r) W_{\min}(r)dr = \int_0^{R_c} \frac{D'(r)}{(1-\rho+D(r))^2} W_{res}$$

In an analogue way as the one used for the maximum SIR priority, we write:

$$\int_0^{R_c} \rho(r) W_{\min}(r)dr = -(\frac{1}{(1-\rho+D(R_c))} - \frac{1}{(1-\rho+D(0))})W_{res}$$

$$= (\frac{1}{(1-\rho)} - \frac{1}{(1-\rho+\rho)})W_{res} = \frac{\rho}{1-\rho} W_{ers}$$

Finally we can write the conservation law

$$\int_0^{R_c} \rho(r) W_{\min}(r)dr = \frac{\rho}{1-\rho} W_{res} \qquad \text{(A-12)}$$

REFERENCES

[1] P. Bender, P. Black, M. Grob, R. Padovani, N. Sindhushayana and A. Viterbi, CDMA/HDR: A bandwidth-e_cient high-speed wireless data service for nomadic users", IEEE Commun. Magazine, 70{77, July 2000.

[2] 3GPP TS 25.308, UTRA High Speed Downlink Packet Access (HSDPA); Overall description, Release 5, 2003.

[3] 3GPP TS 25.214, Physical Layer Procedures(FDD), Release 5, 2003.

[4] S. Parkvall, E. Dahlman, P. Frenger, P. Beming,M. Persson, The high speed packet data evolutionof WCDMA, in: Proc. of the 12th IEEE PIMRC,2001.

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

[6] A proportionally fair scheduling algorithm with QoS and priority in 1XEV-DO, Kuenyoung Kim, Hoon Kim, Youngnam Han, Proceedings PIMRC 2002

[7] Tsang-Ling Sheu and Hsin-Yuan Lin, A Priority-based Channel Allocation Scheme for Cellular Networks, Proceedings of WCNC 2003

[8] Xuefeng Dong and Ten H. Lai, An Efficient Priority-Based Dynamic Channel Allocation Strategy for Mobile Cellular Networks, Infocom 1997

[9] Ariton E. Xhafa and Ozan K. Tonguz, Dynamic Priority Queueing of Handover Calls in Wireless Networks: An Analytical Framework, IEEE Journal of Selected Areas in Communications, Vol. 22, n° 5, June 2004

[10] You Ze Cho and Chong Kwan Un, Analysis of the M/G/l Queue under a Combined Preemptive/Nonpreemptive Priority Discipline, IEEE Transactions on communications, Vol. 41, N° 1, January 1993

[11] Charles Knessl and Charles Tier, A Simple Fluid Model for Servicing Priority Traffic, IEEE Transactions on automatic control, Vol. 46, N° 6, June 2001

[12] L. Kleinrock, Queueing Systems, Vol 2, Wiley, 1976.

[13] Yoshiaki Ofuji, Akihito Morimoto, Sadayuki Abeta, and Mamoru Sawahashi,, Comparison of Packet Scheduling Algorithms focusing on user throughput in HSDPA, Proceedings PIMRC 2002

[14] J.M. Holtzman, “CDMA forward link water filling power control”. Proc. of IEEE VTC Spring, 2000.

[15] J.M. Holtzman, “Asymptotic analysis of Proportional Fair algorithm”. Proc. of 12th IEEE PIMRC, 2001.

[16] T. Bonald, A score-based opportunistic scheduler for fading radio channels, Proc. of European Wireless 2004.