Fair power and transmission rate control in wireless networks

C. Touati*, E. Altman*, J. Galtier*,†

March 28, 2002

Abstract

In third generation mobile networks, transmission rates can be assigned to both real time and non real time applications. We address in this paper the question of how to allocate transmission rates in a manner that is both optimal and fair. As optimality criterion we use the Pareto optimality notion, and as fairness criterion we use a general concept of which the max-min fairness (which is the standardized fairness concept in ATM networks) and the proportional fairness (which characterizes fairness obtained by transport protocols for the Internet) are special cases. We formulate the fair allocation problems as optimization problems and propose an approximating solution.

1 Introduction

Rate control of calls is an important network management issue in 3rd generation mobile networks. Indeed, not only data transfer but also real time audio and video applications can be transmitted at various rates by selecting an appropriate Codec. Indeed, in the case of voice applications, UMTS will use the Adaptive Multi-Rate (AMR) codec that offers eight different transmission rates of voice that vary between 4.75 kbps to 12.2 kbps, and that can be dynamically changed every 20 msec. Of course the transmission rate has an impact on the perceived quality. The reduction of the transmission rate is necessary for maintaining a call whose received energy per bit is too small, and it allows also to maintain a larger number of calls in the system.

A well studied problem is that of choosing transmission rates so as to maximize the system's throughput, see [10]. If two mobiles A and B transmitting at the same rate have the same received power at their base station, if A has larger attenuation than B, then it transmits with larger power than B, thus causing more interference than B causes in base stations of neighboring cells. Hence systemwise, it will be profitable to assign to mobile A lower throughput than to mobile B if there are not sufficient radio resources to assign the maximum throughput to both. This suggests large differences in throughputs assigned to mobiles according to their attenuation level according to the a system optimal viewpoint.

Fairness. Yet a second important consideration in assigning throughputs in networks is fairness. Several fairness concept have been suggested and implemented in various network architectures.

Max-min fairness In ATM networks, the standardized fairness concept in traffic whose rate is controlled (the ABR - Available Bit Rate class) is the so called "max-min fairness" [1]. An assignment vector of rates $\underline{r} = (r(1), ..., r(N))$ to N transmitting sources is said to be max-min fair if one cannot increase the assignment to a source i without decreasing the assignment of a source j that for which $r(j) \leq r(i)$ [1, 4]. The quantity that is assigned fairly in ATM is the excess of the throughput beyond a prenegociated minimum transmission rate. Another related standardized fairness concept in ATM is the "weighted max-min fairness" in which the quantity that is to be assigned fairly is the excess throughputs (beyond the minimum guaranteed) weighted by some multiplicative constant depending of the connections.

Proportional fairness In the Internet, the large majority of transfers use the TCP/IP transport control mechanism. The assignment of throughput to various connections using TCP/IP (and other related protocols) can be described using the concept of "proportional fairness" as was shown in [9]. An assignment r is said to be proportionally fair if it is feasible (satisfies the constraints) and if for any other feasible assignment $\underline{r}^*$, the aggregate of proportional changes is zero or negative:

$$\sum_{i=1}^{N} \frac{r^*(i) - r(i)}{r(i)} \leq 0. \tag{1}$$

*INRIA BP93, 06902 Sophia Antipolis Cedex, France
†France Telecom R&D, 06921 Sophia-Antipolis, France

The proportional fairness is known to maximize the quantity $\prod_i r(i)$. Equivalently, it is an assignment that maximizes $\sum_i \log r(i)$. A (weighted version of the) proportional fairness is also advocated for future developments of TCP, see e.g. [7]. The way TCP shares bandwidth between connection has become a reference for other real-time applications over the Internet that do not use TCP; such applications are called "TCP friendly".

Both proportional fairness and max-min fairness possess optimality properties: they are both Pareto optimal[1]. The proportional fairness is a good compromise between the system optimality notion and the welfare maximization approach of max-min fairness.

Generalized fairness criterion Recently, it has been shown in [15] that all three approaches: the system optimization, the max-min fairness and the proportional fairness are all special cases of a generalized fairness concept. Given a positive constant $\alpha \neq 1$, consider the optimization problem:

$$\text{Maximize} \sum_i \frac{r(i)^{1-\alpha}}{1-\alpha}$$

subject to the problem's constraints. Assume that the $r(i)$ are defined on a convex set. Then since the objective function is concave and the constraints are linear, this defines a unique allocation which is called the α-fair allocation. It turns out [15] that this allocation corresponds to the globally optimal allocation as $\alpha \to 0$, to the *proportional fairness* when $\alpha \to 1$, to the *harmonic mean fairness* (another well known fairness concept) when $\alpha \to 2$, and to the *max-min* allocation when $\alpha \to \infty$.

We should mention that other aspects of fair assignment have been studied previously. These were aspects related to scheduling back logged packets [5, 14, 11], so as to achieve already given average transmission rates of different sources. Our study aims, in contrast, to fairly assign the transmission rates.

The structure of the paper is as follows. We introduce the model in the next section, and show how the fair assignment problem can be formulated as one of two possible optimization problems: one in which transmission rate can be assigned any real value within some given interval, and one in which finitely many transmission rates are available to each mobile. In Section 3 we then address the solution of these problem focusing mostly on the continuous model. A concluding section ends this paper with with extensions to utility-based fairness concepts.

[1] An assignment $\underline{r}$ is Pareto optimal if one cannot increase the assignment to one source i without strictly decreasing an assignement to another source j

2 The model

We use the model and notations of [10]. Consider a cellular radio system with N mobile terminals. Mobile i can transmit with power p_i within the interval $[0, \overline{p}_i]$. Let g_{ij} be the link gain between mobile j and base station i. We assume that time intervals are sufficiently short for g_{ij} to be constant within the interval. Let $a_i = j$ if mobile i is assigned to base station j. Given a power vector $P = (p_1, ..., p_N)$, the received signal to interference ratio of mobile i is given by

$$SIR_i(P) = \frac{g_{a_i,i}p_i}{\nu_{a_i} + \sum_{j\neq i} g_{a_i,j}p_j}, \quad 1 \leq i \leq N, \quad (2)$$

where $\nu_i > 0$ is the background noise power at base station i.

Let us denote $(E_b/I_o)_i$ the ratio of bit energy to interference power spectral density ratio of mobile i, and W_i the spreading bandwidth at chip rate for mobile i.

We next describe two possible settings for the power and transmission rate control.

2.1 The continuous model

In the first model, mobile i can use any value of throughput between a minimum guaranteed value MR_i and a maximum value PR_i. This can be achieved if a packet mode is used with an appropriate scheduling (see e.g. [3, 8] and references therein). Denote $r(i)$ the transmission rate assigned to mobile i, We assume that for each mobile, there is a minimum required value of SIR_i per transmitted bit which we denote by δ_i:

Then we have

$$\delta_i \leq \frac{1}{W_i}\left(\frac{E_b}{I_o}\right)_i = \frac{SIR_i(P)}{r(i)}, \quad (3)$$

implying $\delta_i r(i) \leq SIR_i(P)$.

Thus the solution of our joint power and transmission rate assignment problem is constrained to belong to the set $\Pi^c = (P, r)$, such that:

$$\begin{cases} MR_i \leq r(i) \leq PR_i, \\ 0 \leq p_i \leq \overline{p}_i, \\ \delta_i r(i) \leq \dfrac{g_{a_i,i}p_i}{\nu_{a_i} + \sum_{j\neq i} g_{a_i,j}p_j}, \end{cases} \quad 1 \leq i \leq N.$$

A fair allocation can now be obtained using the following optimization problem:

$$\text{Find } (P,r) \in \Pi^c \text{ that Maximize } \sum_{i=1}^{N} \frac{r(i)^{1-\alpha}}{1-\alpha}. \quad (4)$$

2.2 A discrete model

There are a finite number of available transmission rates to each mobile. Let $r_i^1 < r_i^2 < ... < r_i^{K(i)}$ be the available transmission rates for mobile i.

One way to formulate the discrete model is to repete the formulation of the continuous model, and then add a constraint on the discrete values that the throughputs can have:

$$\text{Find } (P,r) \in \Pi^c \text{ that Maximize } \sum_{i=1}^{N} \frac{r(i)^{1-\alpha}}{1-\alpha}, \quad (5)$$

$$r(i) \in \{r_i^1, r_i^2, ..., r_i^{K(i)}\}. \quad (6)$$

We present below an alternative formulation of the problem following [10].

To properly receive messages at transmission rate r_i^k with tolerable error probability, mobile i is expected to attain an $SIR_i(P)$ not less than a target γ_i^k.

Let $Y = (y_i^k)$ be a 0-1 matrix such that for every mobile i and rate r_i^k,

$$y_i^k = \begin{cases} 1, & \text{if mobile } i \text{ is transmitting with rate } r_i^k, \\ 0, & \text{otherwise.} \end{cases}$$

As in [10], we introduce M_i^k which is an arbitrarily large number satisfying

$$M_i^k \geq \max_P \frac{p_i \gamma_i^k}{SIR_i(P)}.$$

It can be interpreted as the amount of transmission power that mobile i needs in order to attain γ_i^k, regardless of the interference power. (Recall that the power of mobile j is bounded by $\overline{p}_j$, so the right hand side of the above equation is finite. M_i^k is introduced for mathematical needs and is allowed to exceed $\overline{p}_i$.) Thus the solution of our joint power and transmission rate assignment problem is constrained to belong to the set $\Pi^d = (P,y)$ such that:

$$\left\{ \begin{array}{ll} y_i^k \in \{0,1\}, \quad \sum_{k=1}^{K_i} y_i^k \leq 1, & \\ \quad 0 \leq p_i \leq \overline{p}_i, & i = 1, ..., N \\ p_i + (1-y_i^k) M_i^k \geq \dfrac{p_i \gamma_i^k}{SIR_i(P)}. & \end{array} \right.$$

A fair allocation can now be obtained using the following optimization problem:

$$\text{Find } (P,y) \in \Pi^d \text{ that Maximize } \sum_{i=1}^{N} \frac{r_i^{1-\alpha}}{1-\alpha}. \quad (7)$$

3 Solution of the fair assignment problem

In order to solve continuous model, i.e. the fair assignment problem (4), we note that for fixed transmission rates $r(i)$, the minimal power $P^{\min}$ that satisfies the constraints (2) is given by the solution of

$$r(i)\delta_i = \frac{g_{a_i,i} p_i}{\nu_{a_i} + \sum_{j \neq i} g_{a_i,j} p_j}, \quad 1 \leq i \leq N,$$

which can be written as a system of linear equations. Note that for the single cell case, $P^{\min}$ is given by [2]

$$p_i^{\min}(r) = \frac{1}{g_{a_i,i}} \frac{\nu \Delta(r,i)}{1 - \sum_{j=1}^{N} \Delta(r,j)} \quad (8)$$

$$\text{where } \Delta(r,j) = \frac{\delta_j r(j)}{1 + \delta_j r(j)}.$$

Assuming that the system has a single solution $P^{\min}(r)$, we can now rewrite the problem (4) as:

$$\text{Find } r \text{ that Maximize } \sum_{i=1}^{N} \frac{r(i)^{1-\alpha}}{1-\alpha} \quad (9)$$

$$\text{s.t. } (r, P^{\min}(r)) \in \Pi^c.$$

(9) is a non convex nonlinear program and thus does not have in general a polynomial solution algorithm. As in [10], the alternative for an exact solution would be to propose instead polynomial algorithms to obtain good sub-optimal solutions.

In that aim, we consider below the special case in which the following assumption holds:

A1: There is a single isolated cell. There is a feasible joint assignment of power and transmission rate. Moreover, we assume that an optimal assignment r^* exists such that $\delta_i r^*(i)$ is much smaller than one for all $1, ..., N$.

We note that the last part of assumption A1 is quite a natural one. Indeed, assume for example that the number of mobiles is large, that δ_j are all equal, that for all $i = 1, ..., N$ we have $PR_i \leq PR$, $MR_i \geq MR$ and $PR = 2MR$ for some constants MR

and PR. Then in order for a solution $P^{\min}(r)$ for (8) to exist (to be finite) then $\delta \sum_j r(j) < 1$, so that

$$N\delta r(i) < N\delta PR \leq 2\delta \sum_{j=1}^{N} r(j) < 2$$

so that $\delta r(i)$ is bounded by $2N^{-1}$ which is indeed much smaller than 1 for N large.

Consider the following power assignments:

$$p_i^* = \frac{\nu \tilde{\Delta}(r,i)}{g_{a_i,i}(1 - \sum_{j=1}^{N} \tilde{\Delta}(r,j))}$$

$$\text{where } \tilde{\Delta}(r,j) = \delta_j r(j), \quad i = 1, ..., N.$$

It is easily seen that if finite, then this solution satisfies the constraint

$$\delta_i r(i) \leq \frac{g_{a_i,i} p_i}{\nu_{a_i} + \sum_{j \neq i} g_{a_i,j} p_j}, \qquad i = 1, ..., N$$

so it yields a feasible power assignment as long as $p_i \leq \overline{p}_i$.

Under assumption **A1**, this power assignment is close to the minimal one, so we can approximate problem (9) by the following one.

$$\text{Find } r \text{ that Maximize } \sum_{i=1}^{N} \frac{r(i)^{1-\alpha}}{1-\alpha} \text{ s.t.}$$

$$\left\{ \begin{array}{c} MR_i \leq r(i) \leq PR_i, \\ \\ 0 \leq \nu \delta_i r(i) \leq \overline{p_i} g_{a_i,i}(1 - \sum_{j=1}^{N} \delta_j r(j)), \end{array} \right. \tag{10}$$

$$1 \leq i \leq N.$$

This is a standard minimization problem with linear constraints and concave objective function that can be easily solved by either decentralized Lagrangian algorithms or with efficient centralized methods based on SDP (Semi Definite Programming), see e.g. [16]. Going back to the original problem (10) we should note that the fact that the constraints in (10) are close to those of (9) does not in general mean that the solution is also a good approximation. In particular, we could have the extreme situation where problem (10) has no solution (there are not sufficient resources to guarantee to all mobiles their required minimum rates) whereas problem (9) does have a solution. Yet (10) can be used as a simple heuristic to get an approximation for (9). We show that this approximation gives in fact a lower bound on the optimal transmission rates and powers.

Define the set $\Pi^{app} = (P, r)$ s.t.:

$$\left\{ \begin{array}{l} MR_i \leq r(i) \leq PR_i, \\ \\ p_i \leq \overline{p}_i, \\ \\ \delta_i r(i) \leq \dfrac{g_{a_i,i} p_i}{\nu_{a_i} + \sum_{j=1}^{N} g_{a_i,j} p_j}, \end{array} \right. \qquad 1 \leq i \leq N.$$

The only difference between Π^{app} and Π^c is that the summation in the third constraint is over all j. In other words, we approximate the constraint on the ratio between the signal and interference with a constraint on the ratio between the signal and total received energy. It is easily seen that the solution of (10) is the same as the solution of the problem

$$\text{Find } (P, r) \in \Pi^{app} \text{ that Maximize } \sum_{i=1}^{N} \frac{r(i)^{1-\alpha}}{1-\alpha}. \tag{11}$$

Clearly, $\Pi^{app} \subset \Pi^c$ as every pair $(P, c) \in \Pi^c$ clearly satisfies the constraints defining Π^{app}. Hence the value of the approximating problem is in fact a lower bound for the value of the original problem (since we maximize over a smaller set).

We finally comment briefly on the discrete model. The solution of the model (7) can be done in the same way as in [10], using a distributed algorithm based on Lagrangian relaxation. Alternatively, one can use the formulation (5). Its solution can follow a similar path as we had for the continuous case: first express for given transmission rates the corresponding minimum power that satisfies the constraint (3). For the case of a single cell we can again use (8). Then we reduce the problem to (9) with the extra constraint (6). For the case of a single case, we can again use our approximation technique to reduce the problem to one with linear constraints and obtain (10) with the extra integrity constraint (6).

4 Concluding remarks and extensions

In this paper we addressed the problem of joint transmission rate and power control in wireless networks so as to be both fair and optimal.

The paper is in line with many references [9, 12, 15, 17] that considered the throughput as the object to be fairly assigned, in other networking contexts. Another question not addressed here is how to achieve these throughputs in practice if packet mode is used, or in other words, how to schedule packets in order to achieve the throughputs that were fairly assigned. This second question has been well studied see e.g. [5, 14, 11].

In the fairness analysis, one may consider the case in which the utility corresponding to the transmission rates should be fairly assigned, rather than the directly the throughputs. Indeed, since utility represents the degree of satisfaction as a function of the assigned throughput, which may be application dependent, assigning the same throughput to two applications might be highly unfair. In fact, mathematical frameworks for defining fairness indeed exist, within the area of cooperative game theory, and they always relate to utilities. The central concept of this type that has been applied to fair resource allocation problems is the so called Nash Bargaining solution [16, 13, 17]; it turns out to agree with the proportional fairness concept when utilities are linear. If $f_j(r(j))$ is the utility for mobile j to have a transmission rate of $r(j)$, then the Nash Bargaining Solution is given by the solution of the optimization problem:

$$(P_J) \quad \max \prod_{i=1}^{N} (f_i(r(i)) - f_i(MR_i)), \qquad (12)$$

where the maximization is over the appropriate constrained set (Π^c in the continuous model, Π^{app} for the corresponding approximating problem, and Π^d in the discrete problem), see [16]. For real-time voice application f_i are typically concave functions over the interval $[MR_i, PR_i]$, which implies that the objective function in the above problem is concave. In particular, if we consider use the approximating approach to compute bounds for the continuous model and solve P_j over the constrained set Π^{app}, this is then again a standard concave optimization problem with linear constraint for which many efficient (polynomial) methods exist. In particular, several solution methods are proposed in [16] for such problems if we express the utilities using quadratic functions.

References

[1] The ATM forum Technical Committee, *Traffic Management Specification,* Version 4.0, April 1996.

[2] E. Altman, "Capacity of Multi-service CDMA Cellular Networks with Best-Effort Applications", INRIA Research Report, March 2002.

[3] A. Baiocchi, F. Cuomo and C. Martello, "Optimizing the radio resource utilization of multiaccess systems with a traffic-transmission quality adaptive packet scheduling", *Computer Networks*, **38**, 225-246, 2002.

[4] D. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, 1987.

[5] V. Bharghavan, S. Lu and T. Nandagopal, "Fair queuing in wireless networks, issues and applications", *IEEE Personal Communications* **6**(1), Feb. 1999.

[6] T. Bonald and L. Massouillé, "Impact of fairness on internet performance", *Sigmetrics*, 2001.

[7] J. Crowcroft and P. Oechslin, "Differentiated end-to-end Internet services using a weighted proportional fair sharing TCP", *Computer Communications Review*, 28(3):53–67, July 1998.

[8] M. Ferracioli, V. Tralli, R. Veerdone, "Channel adaptive scheduling for a wideband TDD/TCDMA wireless system under heterogeneous traffic conditions", *Computer Networks* **38**, 207-223, 2002.

[9] F. P. Kelly, A. Maulloo and D. Tan, "Rate control for communication networks: shadow prices, proportional fairness and stability", *Journal of the Operational Research Society* **49**, 1998.

[10] S. L. Kim, Z. Rosberg and J. Zander, "Combined power control and transmission selection in cellular networks", Proceedings of *IEEE Vehicular Technology Conference*, Fall, 1999.

[11] S. Lu, V. Bharghavan and R. Srikant, "Fair scheduling in wireless packet networks", *IEEE/ACM Trans. on Networking*, **7**(4), 473-489, 1999.

[12] L. Massouillé and J. W. Roberts, "Bandwidth sharing and admission control for elastic traffic", *Telecommunication Systems*, 2000.

[13] R. Mazumdar, L. G. Mason, and C. Douligeris. Fairness in network optimal flow control: optimality of product forms. *IEEE Trans. on Comm.*, 39:775–782, 1991.

[14] G. Miklós and S. Molnár, "Fair allocation of elastic traffic for a wireless base station" Proceedings of *IEEE Globecom*, Rio de Janeiro, 1673-1678, Dec. 1999.

[15] J. Mo and J. Walrand, "Fair End-to-End Window-based Congestion Control". SPIE '98, International Symposium on Voice, Video and Data Communications, 1998.

[16] C. Touati and E. Altman and J. Galtier, "On fairness in bandwidth allocation", *INRIA Research Report*, September 2001.

[17] H. Yaiche, R. R. Mazumdar and C. Rosenberg, "A game theoretic framework for bandwidth allocation and pricing in broadband networks", *IEEE/ACM Transactions on Networking* **8**, No. 5, pp. 667-677, 2000.