AN APPROACH TO WIRELESS SCHEDULING CONSIDERING REVENUE AND USERS' SATISFACTION

L. Badia[1,*], M. Zorzi[2]

[1]Department of Engineering, University of Ferrara, Ferrara, Italy

[2]Department of Information Engineering, University of Padova, Padova, Italy

ABSTRACT

Recently, several proposals for wireless scheduling algorithms have been presented, using allocation models focused on users' subjective perception of the service. The goal of these investigations is to design scheduling policies aimed at satisfying more directly the users' preference. In this paper we extend this approach, by studying Radio Resource Management, and in particular the scheduler, considering an original model to represent the behavior of multimedia users. We include charging strategies and users' reaction to prices, so that qualitative and quantitative economic considerations are directly included. After a brief discussion on how to include both perceived quality and pricing, in order to achieve a user-centric evaluation of the QoS, we show how it is possible to schedule users by applying this model so as to obtain a more efficient resource usage, characterized by both larger users' appreciation and higher revenue for the service provider.

Keywords

Scheduling, Communication system economics, Resource management, Code division multiaccess.

1. INTRODUCTION

In current communication systems, many services are accessible on the wireless channel, thanks to the diffusion of packet transmission in Wideband Code Division Multiple Access (WCDMA) networks. However, in a WCDMA system it is necessary to consider the possible coexistence of several users interfering with each other, and the inherent unreliable nature of the radio channel. Moreover, in a wireless medium the errors are characterized by variability with time and location, as well as a considerably higher rate and burstiness. All these factors have a strong impact on the system performance and the achieved Quality of Service (QoS). In particular they usually prevent service requests from being satisfied with rigorous guarantees and/or in an equal manner for all the users.

*Corresponding author: *E-mail:* lbadia@ing.unife.it

 ISSN: 1744-2397

For these reasons, it becomes impossible to apply on the wireless medium classic scheduling strategies designed for the wireline case. To solve this problem, it is commonly exploited the fact that a scheduler can obtain an efficient usage of the system capacity by serving in a greedy way the users with the best channel conditions. This leads to the development of Channel-State Dependent Schedulers[1]. However, to have a wide service diffusion, it is necessary to provide a balanced resource supply, also serving users with bad channel conditions. Henceforth, several contributions have dealt with the problem of increasing fairness among users [2, 3]. We are interested in approaching the problem of increasing fairness not only from the theoretical point of view but also in terms of achieving a satisfactory service according to the price which the users are willing to pay, because users' satisfaction is connected with economic aspects, like the provider's task of achieving an adequate revenue. For a real operator, in fact, the network maintenance is possible only if the costs of service provisioning are compensated and overcome.

Users' appreciation of the supplied QoS can be investigated with an approach based on utility functions, which is a field of research widely explored in the recent literature, in particular for the Radio Resource Management (RRM) operations [4, 5]. Several scheduling strategies can be regarded as applications of a utility-based framework. Since it is possible to see the scheduling problem as a prioritization of users in a queue, this can be done by defining appropriate weights, as in the Weighted Fair Queuing (WFQ) [6] or WF^2Q [7] schedulers. A detailed mathematical formalization of the resource allocation problem, which involves utility functions to represent the QoS requirements, can be found in [8]. By combining these approaches, the scheduling can be seen as a strategy to maximize the total system utility. However, the system welfare is still seen only from the users' perspective, without considering the operator's economic counterpart.

In this paper we rather aim at considering also the provider's point of view, which means to include considerations about the earned revenue. To do so, it is necessary to extend the framework and consider the tariff of the service, by introducing a pricing function with properties similar to utility but with a negative impact on users' degree of satisfaction. Hence, we directly investigate the trade-off between utility and price by applying a micro-economic framework, developed in [9], to the scheduling problem. This allows us to refer to economics in two directions, since the allocation efficiency is eventually seen not only as the QoS provisioning but also through the money exchange between user and service provider.

In the preliminary investigation carried out in [9], we applied the model to dimension the network and to gain insight about the points of trade-off involved. The contribution of the present paper is to directly utilize the user-centric model of service appreciation as a concrete way to develop scheduling strategies in WCDMA system. In particular, we investigate how an acceptable QoS for the users and together a satisfactory income for the provider can be obtained. We evaluate existing techniques to enhance scheduling from the technical point-of-view (i.e., by maximizing throughput), and discuss their performance within the above framework. Moreover, we study and quantify possible margins of improvement for both users' satisfaction and provider's revenue, so to show that introducing service-perception awareness in the scheduler significantly increases the goodness of the allocation.

The remainder of this paper is organized as follows: in Section 2 we present the analytical model for the users' satisfaction, including both pricing and utilities. In Section 3 we discuss different scheduling strategies under the theoretical point-of-view and we outline how they can be modified to take into account the revenue improvement aspect. Section 4 presents simulation results and Section 5 concludes the work.

2. MODEL FOR USERS' SATISFACTION

In this section, we describe at first the properties of utility and pricing functions which will be exploited in the following analysis. Then, we consider a micro-economic model which quantifies the level of satisfaction for each of the users, determined by both the perceived QoS and the price paid, so that their trade-off can be explicitly taken into account. More details about the model are available in [9]. Here, we simply give a brief summary and then go directly into the application of the model to the scheduling for wireless CDMA networks.

Consider a set (network) of N users. In micro-economics [10], the concept of utility function is introduced to represent the evaluation among the customers (users) of an assigned resource, which we represent with a generic non-negative parameter r. For simplicity, we assume that r is a scalar, even though an extended analysis can be easily obtained for a multi-dimensional r by following the same approach presented here. The allocation of the resource r generates an assignment vector $\mathbf{r} = (r_1, \ldots, r_N)$, where r_i is the amount of resource allocated to user i. The utility of the ith user is related to the amount of resource received. An example of this can be obtained is shown in [11]. In other words, the element r_i is evaluated, from the perspective of the ith user, as mapped through the function $u_i(\cdot)$, determining a perceived utility $u_i(r_i)$. When network management is studied, it is reasonable to assume that r_i represents the share of network resources allocated to user i and the function $u_i(\cdot)$ is different for each user, as it depends on the technology of the terminals and the subjective perception of the users.

For what concerns the wireless scheduling, it is possible to identify r with the assigned data rate. To correctly evaluate the utilities, r_i shall be averaged on the whole transmission, to obtain meaningful results, since subjective perception of the achieved data rate is possible only after a sufficiently large time window. A more complicated analysis which re-evaluates the situation every given time interval is possible, even though it is omitted here for the sake of simplicity.

A detailed investigation on how to derive proper utility functions for different specific communication systems is beyond the scope of this paper. Thus, we simply take standard assumptions performed in the literature for the utility-based RRM; if necessary, they can be specialized in more detail. Usually, utilities are assumed for every kind of assignment to be quasi-concave non-decreasing functions, i.e., $u''(r)<0$ for sufficiently large r. We further introduce the assumption of having sigmoid-shaped functions, as in [12]. In particular, we adopt the following expression:

$$\forall i = 1,2,\ldots,N \qquad u_i = \frac{(r/K_i)^{\zeta_i}}{1+(r/K_i)^{\zeta_i}} \tag{1}$$

where the parameters $K_i > 0$ and $\zeta_i \geq 2$ depend on the index i, so that different users may be characterized by different utility functions. In the simulations, K_i and ζ_i are randomly generated with uniform distribution within a given interval, which can be connected in particular with the specific type of service requested by the users.

These assumptions are quite common in the literature, and they are also suitable for our purpose: for example, sigmoid functions have also an upper-limit, which seems to be a reasonable description of the performance of multimedia services on wireless networks. In fact, according to the type of systems, in every case there is a maximum rate $r_{\max}$ allowed by the technological support and it is also realistic to assume that the highest perceived QoS is close to $u(r_{\max})$. Thus, it is possible to write:

$$\forall i = 1,2,\ldots,N \qquad \lim_{r\to\infty} u_i(r) = u_i(r_{\max}) \tag{2}$$

The above assumptions for the utilities are however very general, since the internal parameters K_i and ζ_i can be tuned so as to determine very different behaviors for the users. An appropriate way to choose them could be in connection with the user intrinsic parameters.

The above definition considers only non-negative values of the utilities, and the lowest utility, which is achieved when no service at all is supplied, is $u_i(0) = 0$. This condition can also be changed if run-time service degradation is considered, as in this case the utility could even go below 0; in fact, it is commonly assumed preferable not to be admitted at all than to be disconnected from the network while receiving the service. Thus, a lowest utility equal to 0 corresponds to a condition of *ideal Admission Control* [13].

The aggregate of the utilities can be regarded as the total network welfare $W(\mathbf{r})$, and a possible goal of the RRM can be considered to be the welfare maximization. If the utilities are additive, W is simply their sum. This leads to formulate the RRM task as an optimization problem [8]. A way to formalize it is:

$$\max\ W(\mathbf{r}),\ W(\mathbf{r}) = \sum_{i=1}^{N} u_i(r_i) \tag{3}$$

$$\text{s.t.}\ \mathrm{C}(\mathbf{r}) \leq C_{\max} \tag{4}$$

To have a properly defined optimization problem, one must also take into account a capacity constraint of the network, represented by (4) in the above formulation. Thus, $C(\mathbf{r})$ is a given function of the allocation vector r and $C_{\max}$ describes the upper limit allowed for C. Note that also this constraint can be easily extended to a multi-dimensional condition.

For a generic communication system, the simplest possibility to model the capacity constraint is to consider a *hard capacity* system, where a fixed maximum total rate $C_{\max}$ can be allocated on aggregate. For wireless systems, this holds for example for a Time or Frequency Division Multiple Access (TDMA, FDMA) where the capacity limit is related to the number of available time or frequency slots. Thus, the function $C(\cdot)$ is simply a sum, and the constraint becomes:

$$\sum_{i=1}^{N} r_i \le C_{\max}$$

However, note that, even though this kind of constraint is useful to understand what follows, it is not realistic for WCDMA networks. Code-division multiplexing has in fact a similar limitation in the maximum number of codewords, but usually this number is assumed to be very large for practical purposes. Another constraint, i.e., the interference limit, is usually more restrictive and will be considered in the following [14]. The interference limit can be modeled by considering the powers allocated to the users, imposing e.g., the Signal-to-Interference ratio of each user to be known and the total power allocated to the users in the same cell to be limited. According to the link gain conditions, Signal-to-Interference Ratio (SIR) requirements and allocated rates, an interference condition can be written for every user.

This condition can be formalized by introducing a vector of powers $\mathbf{P} = (P_1, \ldots, P_N)$, where P_i is the power allocated to user i. If for the sake of simplicity we consider a single-cell case, this condition is:

$$\Lambda_i(\mathbf{P}) r_i \le P_i \tag{5}$$

where $\Lambda_i(\cdot)$ is a first degree polynomial expression. This can be connected with the usual formulation of the interference condition for CDMA systems expressed as meeting an SIR-target for user i called g_i, as follows:

$$\gamma_i \le \frac{BW}{r_i} \frac{g_i P_i}{\sum_{j=1, j \ne i}^{N} \xi_{ij} g_i P_j + \eta_i} \tag{6}$$

where BW is the spreading bandwidth, g_k is the power link gain for the kth user, $0 \le \zeta_{ij} \le 1$ is the normalized cross-correlation between i and j at receiver i and the term h_i is due to background noise. This can be re-arranged so as to write, as in (5):

$$(BW \cdot \gamma_i)\left(\sum_{j=1, j \ne i}^{N} \xi_{ij} P_j + \frac{\eta_i}{g_i}\right) r_i \le P_i \tag{7}$$

In this case, a vector $\mathbf{r}$ is said to be *feasible* if these constraint are met. The feasibility condition can be complicated to investigate with analytical instruments, but is easy to check within a simulator. For this reason, in the present paper we adopt a simulation-based approach.

An extension of the problem described above can be given considering also the effect of pricing [15, 16]. Once the strategies to charge users for the offered service are given, it is possible to determine the generated revenue; it is also reasonable that, among allocations that are almost equally satisfactory for the users, the operator chooses the one that determines the highest revenue, since in this way its own satisfaction is increased. As long as this can be done without decreasing too much the users' welfare, this implies indeed a more efficient resource usage. On the other hand, pricing impacts negatively on users' choice. As an immediate consequence, by considering that the service does not come for free, it is possible to understand when an unnecessarily high amount of resource is provided to some users. Over-assignments are a trivial way to allocate a high amount of resource, but they are possible only if the price is neglected: in a more realistic case they are usually refused by users since they imply higher prices.

This economic aspect enforces even more the intrinsic property of interference-limited system, like the CDMA environment, of being unable of providing strict QoS guarantees, unless resource are severely over-provisioned. Henceforth, the best effort allocation that can be commonly provided by such systems is not always considered acceptable by the end users. Our claim is that this decision is also strongly impacted on by the tariff paid for the service and the willingness to pay of the users.

More in general, without considering the pricing impact it is impossible to understand whether the allocation is not only technically efficient, but also sensible from a micro-economic perspective, where users that do not get both adequate QoS and affordable price are unsatisfied customers. It can be assumed that these users only pay a certain fraction of the tariff due, or alternatively they leave the service with a certain probability. Hence, the model presented in [9] proposes to define the satisfaction A_i of the ith user as a value into the range [0,1]. This satisfaction value can be seen as a weight, or a probability of service acceptance, to evaluate all metrics related to resource assignment, so that only resource coming from satisfactory allocations is efficiently used. For example, the average revenue R earned by the provider is evaluated as

$$R = \sum_{i=1}^{N} A_i p_i \tag{8}$$

where p_i is the price paid by the ith user. Similar expressions can be used to evaluate other metrics [13], for example the sum of the A_i's represents the fraction of satisfied customers, i.e., the ones who keep paying for the service without abandoning it or being driven to other operators. However, the interesting point is that this framework is not only useful for evaluations, but also indicates a possible approach to scan the solutions of the allocation problem to improve the scheduling strategy, as will be outlined in next Section.

To map the trade-off between the offered QoS and the price paid, we associate A_i with u_i and p_i. Hence, we write $A_i = A(u_i, p_i)$ for each i, by assuming that every user in the network adopts the same criterion to decide whether the service conditions are satisfactory or not. A possible expression [9] for $A(\cdot)$ is as follows:

$$A(u_i, p_i) = 1 - e^{-k(u_i/\psi)^{\mu}(p_i/\phi)^{-\varepsilon}} \tag{9}$$

where k, μ, ε, ψ, ϕ are positive constants. The exponents μ and ε tune the sensitivity to utility and price, respectively, whereas ψ, ϕ and k are simply normalization constants (a reference utility, a reference price, and the opposite of the logarithm of a reference value for A, respectively). Note that the above equation is written in this form only to emphasize the trade-off, but actually both u_i and p_i depend on the allocated resource r_i. In fact, in this work we always consider that u_i's follow (1), whereas $p_i = p(r_i)$, with $p(\cdot)$ being a non decreasing function, which is the same for all users (note that this assumption is made for fairness reasons, since in this paper we consider only one service class for all users). If there are different service classes, then a set of pricing functions should be considered. However, it is realistic to assume that this set is small; thus, an extension to this case is straightforward). The definition reported in (9) satisfies several properties which are expected to characterize A_i, like monotonicity or boundary conditions (for details see [9]). However, it is adopted here only for the sake of simplicity, but the conclusions drawn in the following are still valid for other choices of $A(u,p)$.

The total revenue of the provider, evaluated in (8) can be rewritten as $R(\mathbf{r})$. This opens up the formulation of a different optimization problem, in which the goal function is no longer the users' welfare but the revenue. Note that this goal, besides being an alternative which might be interesting for the provider, is not disjoint from welfare maximization, since generally the larger the utility, the larger A_i and henceforth the revenue. This follows directly from (8): however, A_i decreases with increasing p_i. With a similar formulation of the problem (3)-(4) we write then:

$$\max\ R(\mathbf{r}),\quad R(\mathbf{r}) = \sum_{i=1}^{N} A(u_i(r_i), p_i(r_i))\, p_i(r_i) \qquad (10)$$

$$\text{s.t. } C(\mathbf{r}) \le C_{\max} \qquad (11)$$

where (10) is still related to the WCDMA interference management. The scheduling strategies derived in the following will be identified as approximate solutions to this problem.

3. ANALYSIS OF WIRELESS SCHEDULING WITHIN THE MODEL

This section is dedicated to the application of the above framework to scheduling strategies for WCDMA systems. In wireless systems, scheduling heavily impacts on the performance and in particular the scheduler must be aware of the radio conditions, because of the location-dependent and bursty errors. For example, a user in a fading dip may experience a bad channel and may be unable to transmit. In this respect, also time variability, which is determined by fading and users' mobility, has to be considered. For these reasons, we assume in the following to have a wireless system with mechanisms to predict channel conditions. An example is the High Speed Downlink Packet Access (HSDPA) release of UMTS [17]. In HSDPA, the channel conditions might be rapidly tracked to improve the system throughput, thanks to the Medium Access Control features located in the node-B, to evaluate the rapid variations of the wireless channel [18].

The knowledge of the channel state makes it possible to give priority to users that perceive a clean channel, whereas users with a poor SIR will be delayed. In this way the scheduler tends to maximize the total throughput because it minimizes packet retransmissions, while on the other hand fairness is decreased. Conversely, fairness is obtained, e.g., with a Round Robin (RR) approach, where resources are allocated to the communication link on a sequential basis. However, this results also in the potential risk of having high number of corrupted packets and consequent retransmissions. Due to the poor performance of the pure RR scheduler, in the following we will not analyze this strategy. To better guarantee the QoS requirements it is necessary to find a trade-off between a pure SIR-based heuristic and a round robin scheduling, i.e., between maximizing the throughput and allowing the highest possible number of users to achieve a satisfactory QoS.

In our analysis we consider first of all a traditional SIR-based scheduler, called C/I, exploiting a greedy assignment of the available resources [19]; such a strategy obtains the maximum sector throughput, but with a high degree of unfairness. A possible solution to cope with this problem was proposed in [3], where a similar scheme called in the following WCA C/I (short for *Weighted Code Assignment* C/I) scheduler, was introduced. The WCA C/I scheduler adopts a utility-function-based assignment, with the goal of increasing fairness. In other words, the allocation procedure considers a weighted combination of exponential functions related to possible objectives of the scheduler and tries to heuristically increase it with respect to the pure channel-state dependent case.

In particular, the allocation criterion of the WCA C/I scheduler considers the following function as a sorting value to prioritize the allocation of a packet to the requesting users:

$$\beta_S e^{C_i S_i} + \beta_Q e^{C_i Q_i} + \beta_D e^{-C_i D_i} \qquad (12)$$

where S_i is the normalized SIR of user i in the considered HSDPA Transmission Time-Interval (TTI), Q_i is its buffer occupancy and D_i is the normalized head packet deadline. The term C_i simply refers to the possibility of managing different classes of traffic (we did not investigate this here, so C_i can be equal to 1 for all the users). Finally, the weights β_k, with $k \in \{Q, S, D\}$ are tunable weights which can be adjusted to change the behavior of the scheduler (only their relative ratios are relevant). Note for example that letting $\beta_s \neq 0$ and $\beta_Q = \beta_D = 0$ corresponds to adopting a pure C/I scheduler. In the following we set the weights so that $\beta_S = \beta_Q = \beta_D$, which corresponds to a scheduler trying to combine several objectives at the same time and in a similar manner. Within this approach, users' utility can account for more parameters like SIR, buffer state, deadline of the packets, which can be mixed in a more efficient manner than in the C/I case. Henceforth, the dependence on the utility permits to obtain not only a better degree of fairness, but also a generally better allocation. An optimization of the scheduling weights, which would be surely possible, is beyond the goal of the present paper. Note, however, that the results shown in the following hold in a qualitatively similar manner for other choices of β_s, β_Q, and β_D.

With the micro-economic framework is possible to compare for example the pure C/I strategy and the WCA C/I one. The merit of the framework outlined in Section 2 is in this case to show that it is possible to compare the policies of scheduling algorithms in terms of economic metrics. For example, it is immediate to think that a better matching of users' utilities can also determine a higher revenue.

Table 1. Local search algorithm

```
local_search ( )
{
    rev1=R(r);
    rev0=0;
    while (rev1>rev0)
    {
        rev0=rev1;
        for (i=0; i<N; i++)  x_i = r_i + Δ ;
        k = max_{i=1...N} u'_i(x_i) ;
        for (i=0; i<N; i++)  y_i = r_i − Δ ;
        h = min_{i=1...N} u'_i(y_i) ;
        if (k==h) break;
        else
        {
            for (i=0; i<N; i++)  z_i = r_i ;
            z_k = x_k ;
            z_h = y_h ;
        }
        Rev1 = R(z);
        if (rev1>rev0)
        for (i=0; i<N; i++)  r_i = z_i ;
    }
}
```

However, we can also think to address the problem of the users' service appreciation and revenue generation more directly. Thus, we propose an original contribution in which the solution presented in [3] is taken as an initial condition, but the assignment is modified iteratively by means of a local-search algorithm, obtaining a locally optimum solution. A possibility is to use the algorithm reported in Table 1. Here, the scheduling process starts from a solution obtained with a greedy heuristic and modifies this assignment by giving more resources to the user with the highest marginal utility, in order to improve the total sector utility. Resources are subtracted to users with the lowest marginal utility, to obtain a variation of the total utility as small as possible. The algorithm ends when the goal function reaches a local maximum.

The revenue will also depend on the pricing strategy. Thus, the choice of the function $p(r)$ should be indicated to clarify the above definition of revenue, as given in (8). In the literature [16], different pricing strategies have been proposed, and obviously the pricing strategy choice heavily affects the value of the total revenue. In this work we will consider two kinds of pricing policies, mainly for their conceptual simplicity. The first one is a *flat price* strategy, i.e., the price is fixed for any value of the assigned rate. The second policy represents a simple usage-based pricing with linear price. This means that $p(r)=kr$ is linearly related to r through a given constant k. It is interesting to observe that in (8) there is expressed a double dependence of the revenue on the pricing, as also A_i is a function of the price. These two metrics are also representative of other values of interest from the technical point of view. Since a flat price policy assigns the same price to all users, the revenue is directly proportional to the number of users accepting the service. Hence, the revenue for the flat price policy can be seen also as a measure of the number of admitted users. Instead, the revenue in the linear pricing case equals the unit price times the throughput; thus, in this case a weighted version of the throughput is considered. In general, it is at the same time true that a real pricing policy is likely to be something hybrid between these two policies [15], but also the interest for the provider in having a satisfactory revenue is connected to having both high throughput and a large fraction of satisfied users. For these reasons, we claim that this choice of pricing policies is interesting, as it allows to separate these two opposite goals, among which a trade-off can be cut.

In the next we will consider the behavior of the C/I scheduling policies in the classic [1, 19] and modified [3] version against our proposal introduced to improve the revenue. The policies will be compared by means of simulation in terms of generated revenue in order to highlight the consequences on the provider side.

Table 2. List of Parameters of Simulation Scenario

Parameter (symbol)	value
cell radius (d)	250 m
gain at 1 m (A)	-30 dB
path loss exponent (α)	3.5
shadowing parameter (σ)	4 dB
Doppler frequency (f_d)	2 Hz
mean SNR at cell border	40 dB
max assignable rate (r_{max})	96
utility parameter (ζ)	5.0÷8.0
utility parameter (K)	0.2÷6.0
acceptance prob. parameter (k)	-log 0.9
acceptance prob. parameter (ϕ)	1.0
acceptance prob. parameter (ψ)	1.0
acceptance prob. parameter (μ)	2.0
acceptance prob. parameter (ε)	4.0

4. RESULTS

In this Section we will present the results obtained with a HSDPA UMTS simulator developed at the University of Ferrara. A cellular cluster is simulated with a 3x3 hexagonal cell structure and wrapped onto itself in order to avoid border effects. In radio channel propagation, path loss, fast fading and shadowing have been included. To consider the environment mobility, a non-zero Doppler frequency is assigned, even though stationary users are considered. All these effects are included in the Power Control module, so that the spreading gain (which determines the rate) and the transmitted power are tuned to allocate a vector **r** which is feasible with the interference constraints. This applies (11) to our case. Table 2 reports the parameters for the simulation scenario and the Acceptance-probability model.

In Fig. 1-6 we compare different scheduling strategies, by evaluating the earned revenue for the cases of flat and linear price. In particular we compare the classic C/I "C/I" and "WCA C/I" respectively. Also, we consider improved versions for both strategies in which the Local Search (LS) procedure introduced in the previous Section is implemented. These strategies will be referred to as "LS with C/I" and "LS with WCA C/I" respectively. Note that the strategy "LS with C/I" is not reported in all graphs, namely where it obtains similar results to the basic "C/I". Hence, sometimes we choose not to plot it to avoid unnecessary confusion in the graphs.

From all the results reported, it is clear that the classic C/I scheduling obtains worse revenue performance with respect to the weighted version. This is reasonable, since the C/I method has the problem of assigning the resource to the best users, without considering their utility. In fact, a given user might be already satisfied and not require a larger assignment, even though his channel conditions are good and a C/I scheduler would assign him more resource. In other words, the C/I scheduler might introduce over assignments, which are avoided with the WCA version.

Note that we are introducing here two kinds of contributions: first of all, we evaluate the performance of the scheduler in terms of economic quantities, which reflect also efficiency metrics, as discussed before. Moreover, we are able to go further by introducing the LS algorithm, where the micro-economic concepts of users' satisfaction are employed also to drive the scheduler. This strategy is more revenue-aware, and therefore is able to improve the performance, with respect not only to the pure "C/I" scheduler, which is clearly outperformed, but also to the weighted assignment.

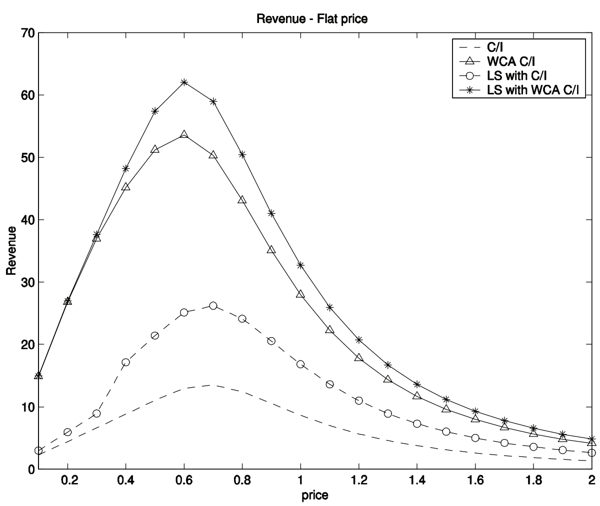

Figure 1. Provider revenue for flat price, 120 users, as a function of the price

There are two types of improvement that the LS solutions present: the first one is a general increase in the curves, which means that the revenue that the provider can achieve is increased. Secondly, the LS curves present a wider range of prices for which a given reference revenue is overtaken, which means an increased robustness. In fact, one might expect that the price can not always be set up optimally by the provider. Thus, the LS strategy is also able to increase even more significantly the revenues given by suboptimal prices, which implies a higher robustness of the system performance. Finally, an important comment which distinguishes between the flat and linear pricing is as follows: for the flat pricing a revenue improvement means that the resource is better allocated since satisfaction is achieved by a higher number of users. For the linear pricing, things are more complicated. If the number of users is sufficiently large, a revenue increase can not be achieved only by increasing the number of satisfied users, since if two assignment vectors allocate the same global amount of resource C, the revenue is identical (it is equal to the unit price times C). Thus, to improve the revenue in the linear case, the scheduler must also increase the total throughput. This is more difficult to do when the system is saturated, i.e., the rate request is higher than the resources available for allocation, since the total rate allocated is supposed to be high anyway. However, in WCDMA systems it becomes possible thanks to a better interference management, since the total allocated rate is not fixed, but depends on the interference constraints, as discussed in Section 2. A Local Search aimed at improving revenue is therefore able to increase the assignment where it is more efficient, i.e., for users which cause less interference to the system.

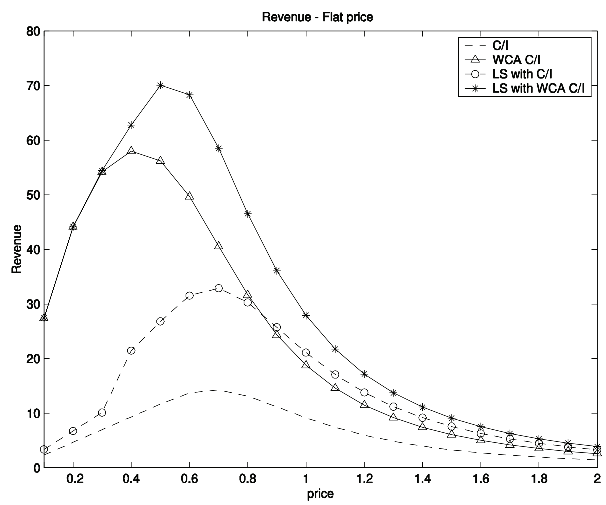

Figure 2. Provider revenue for flat price, 180 users, as a function of the price

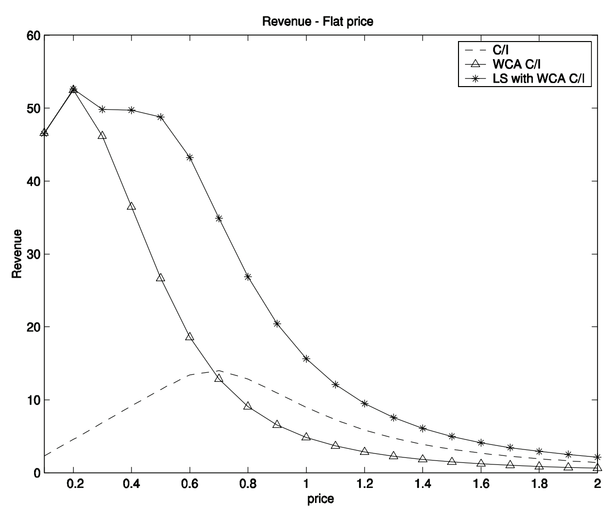

Figure 3. Provider revenue for flat price, 300 users, as a function of the price

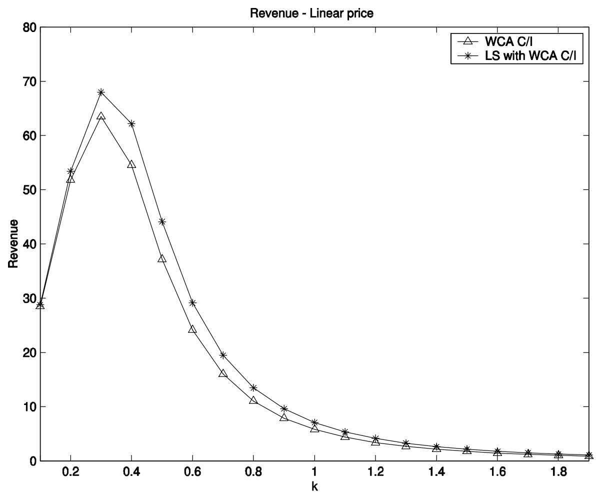

Figure 4. Provider revenue for linear price *p*(*r*)=*kr*, 120 users, as a function of *k*

We now discuss the results in more detail. Fig. 1 to 3 refer to the case of flat pricing policy. As discussed before, here the revenue is directly related to the number of admitted users, as the tariff paid by each user is the same, regardless of the quality, as long as the service is satisfactory. In this case it can be shown that there is a margin for increasing the number of satisfied users with respect to the C/I strategies. This happens because the LS allocation scheme avoids unnecessary over assignments, which waste resources without improving users' satisfaction. For the flat pricing policy, we plotted results also for the original C/I strategy since it is comparable with WCA (even though the general performance is poorer). In Fig. 1 and 2 we considered also the LS with C/I, without weighted code assignment. It is emphasized here that the LS strategy, being only directed toward a local optimum, heavily relies on the first solution. However, there is anyway a significant gain in using the Local Search: the revenue increase achieved by the improved allocation scheme is between 15 and 20 percent for the cases with low load (120–180 users). For higher loads, this gain decreases if we consider the peaks of the curves only. Two remarks should be made, though. First of all, the gain in applying the LS to heavily loaded networks is lower since the larger the number of users, the higher the revenues in general, since it is likely that the capacity is fully allocated even by an inefficient scheme. On the other hand, even though the peaks of the curves are approximately the same, which means that the maximum achievable revenue is only slightly increased by the LS strategy, as remarked before the price setup is more robust because of the general increase by the whole curve. In particular, the range of prices achieving a generally high revenue is wider.

Similar conclusions can be drawn also for the case of linear pricing policy, analyzed in Fig.s 4-6. For this pricing strategy, the revenue can be equivalently seen as a measure of the total assigned throughput, i.e., how much the provider succeeds in assigning as much resource as possible to the users. We have a gain of the same order of magnitude as in the previous case, and this shows that even with a completely different pricing strategy our proposal is still able to significantly improve the assignment. However, note the following: first of all, the simple C/I strategy is not reported here (neither in the original nor in the LS version) due to the poor performance exhibited. The reasons for the failure of the standard C/I policy when we adopt a linear pricing are in the higher inefficiency of over provisioning. In this case in fact, not only over assignments are a waste, but they also decrease the users' appreciation of the service, since they imply a higher cost. Another phenomenon that should be observed is that the gain of the LS strategy is slightly lower than with flat pricing (its average is around 10%), even though it is not always decreasing with the number of users. However, as previously discussed, this is indeed the result of a strong improvement in the assignment efficiency, since it corresponds to a throughput increase.

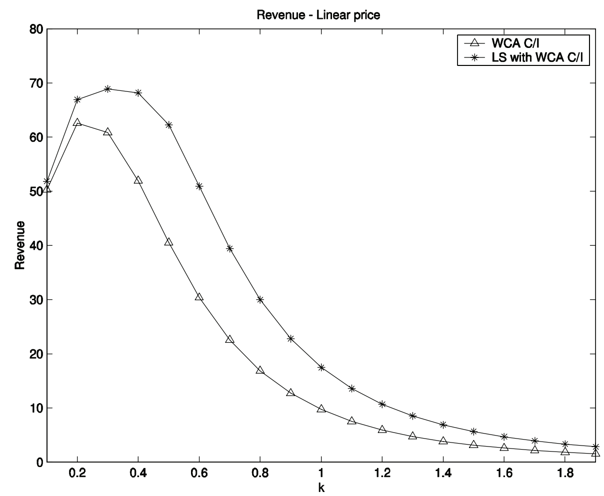

Figure 5. Provider revenue for linear price *p*(*r*)=*kr*, 180 users, as a function of *k*

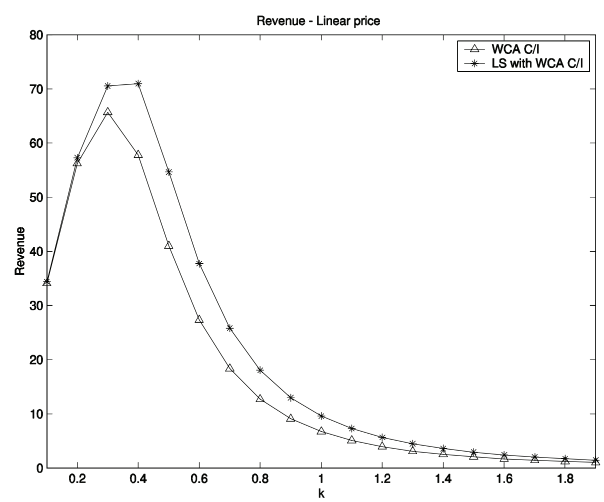

Figure 6. Provider revenue for linear price *p*(*r*)=*kr*, 300 users, as a function of *k*

To sum up, for all the cases reported in Fig. 1-6 the Local Search strategy outperforms significantly the C/I (or WCA C/I) algorithm taken as the initial solution. The price for such improvements is in the increased computational complexity required by the LS procedure. Fortunately, for the examined cases this is not very high, since the number of iterations is usually low (10 iterations at most, but usually 4-5); thus, in this case a simple variation from the initial solution allows to greatly improve the earned revenue. This remark can be extended to make the scheme even more tunable, since further research can be done in the direction of applying different local search heuristics so to explore the solutions space in more depth.

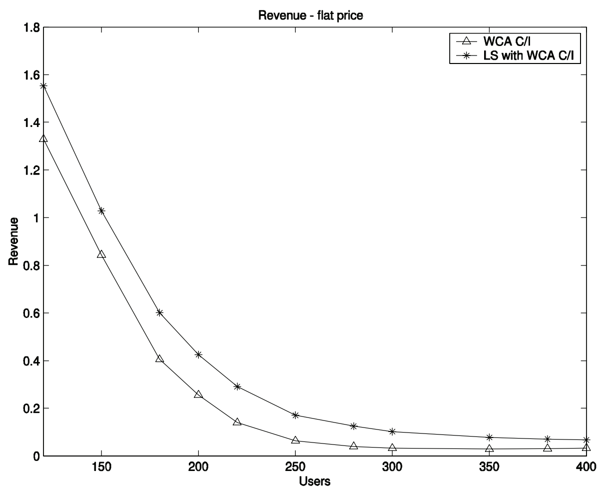

Figure 7. Normalized provider revenue for flat price equal to 0.8, as a function of the load

We also reported a linear interpolation of the collected prices. These results can serve as a practical guideline in price setting. In particular, the most interesting conclusion which can be drawn from these results is that flat pricing is very sensitive to the number of users (and hence the optimal price setting is more critical), whereas linear pricing is almost insensitive on the number of users, which confirms the previous discussion about its suitability for the purpose of tariff collection.

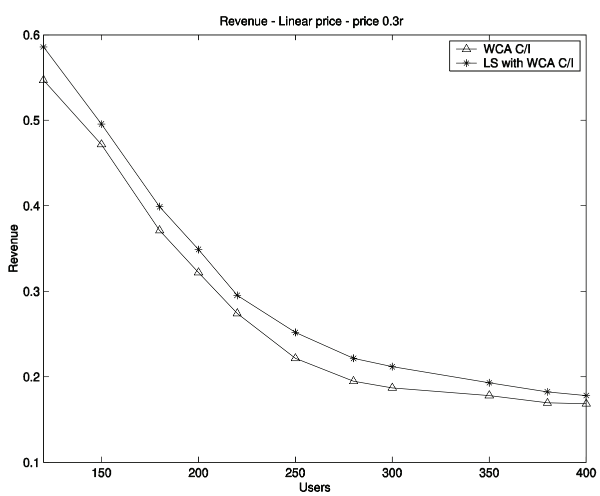

Figure 8. Normalized provider revenue for flat price equal to 0.8, as a function of the load

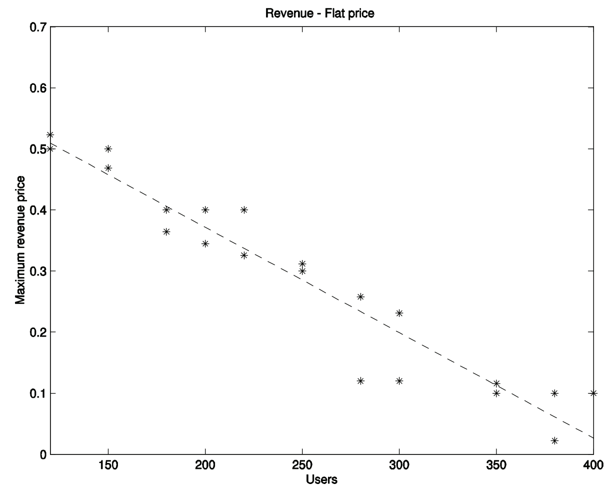

Figure 9. Price maximizing revenue for flat price, as a function of the load

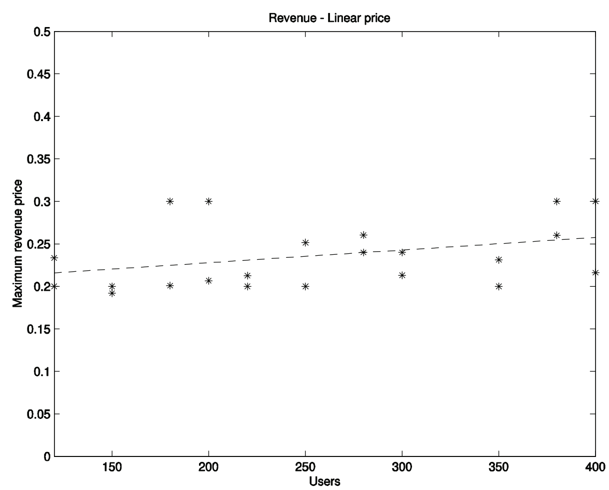

Figure 10. Price maximizing revenue for linear price, as a function of the load

To better study the dependence on the network load (i.e., the number of users), and to gain insight on the price setting issue, we might consider the collection of the revenues achieved by the same scheduler for different load conditions for a given pricing policy setup. For example, in Fig. 7-8 this is reported for flat (with price $p(r) = 0.8$) and linear ($p(r) = 0.3r$) pricing, respectively. Other choices of flat or linear prices, respectively, exhibit completely similar results, so that these choice are more or less representative of the entire class.

From these Figures, one can see that the gain in using the LS strategy is present for different load conditions. However, this holds if the price is kept constant and not adaptively set. Hence, it could be also interesting to collect all the price *values maximizing the provider revenue*, to understand how the price setting is sensitive to the load. The result of this operation is represented in Fig. 9 and 10, for flat and linear pricing, respectively. Noise is present in these Figures, because the granularity of the results is not always accurate, since only certain price values are simulated.

Finally, we remark that we also tested rather different choices for the local search policy, and the results obtained are more or less equivalent to those shown in this section. In general, it is emphasized that the real gain comes from the scheduling strategy being driven by a utility- and price-aware model, which makes it capable of better allocating the constrained resource.

5. CONCLUSIONS

We analyzed the scheduling from the provider's point-of-view, by including also revenue maximization among the goal of the scheduler. This is made possible by the introduction of the *Acceptance-probability* model, which accounts for the joint effect of user utility and price, and includes economic considerations into the analysis.

The results show interesting possibilities of improving the network management. In particular, the application of a classical efficient strategy, like the C/I scheduler, by neglecting the economic counterpart of the allocation, can lead to unsatisfactory results for the operator, since the maximized throughput provided by the C/I strategy might not be what users want. On the other hand, a simple strategy that locally searches for higher values of the revenue is able to greatly improve the profit and the economic efficiency of the resource management, by keeping the users' satisfaction level almost constant, if not increased. Thus, the usefulness of the economic considerations is highlighted.

Several further observations can be made on the strategy used for optimization. The heuristic strategies discussed here offer the advantages of simplicity and fast evaluation; however, more detailed procedures can improve even further the performance and/or the convergence rate. It could be also possible to develop, within the given framework, a theoretical analysis of the scheduling, in which the optimization problem (10)–(11) for revenue maximization is explored analytically. This study, that can allow better understanding of the RRM issues, is left for future research.

REFERENCES

[1] C. Fragouli, V. Sivaraman, & M. Srivastava, "Controlling multimedia wireless link via enhanced class-based queueing with channel state dependent packet scheduling," Proc. INFOCOM, San Francisco, CA, USA, 1998, Vol. 2, pp. 572–580.

[2] X. Liu, E. K.-P. Chong, & N. B. Shroff, "Transmission Scheduling for Efficient Wireless Utilization," Proc. INFOCOM, Anchorage, AK, USA, 2001, Vol. 2, pp. 776–785.

[3] M. Boaretto, D. Angelini, & M. Zorzi, "Fast scheduling policies using WCDMA high-speed downlink packet access," Proc. VTC Fall, 2003, Vol. 4, pp. 2650–2654.

[4] V. Siris, "Resource control for elastic traffic in CDMA networks," in Proc. ACM MobiCom, Atlanta, GA, 2002, pp. 193–204.

[5] M. Xiao, N.B. Shroff, & E.K.-P. Chong, "A Utility-based Power-control scheme in Wireless Cellular Systems," IEEE/ACM Transactions on Networking, Vol. 11, No. 2, 2003, pp. 210–221.

[6] A. Demers, S. Keshav, & S. Shenker, "Analysis and simulation of a fair queuing algorithm," Proc. ACM SIGCOMM, Vancouver, BC, Canada, 1998, pp. 1–12.

[7] J.C.R. Benner & H. Zhang, "WF2Q: worst case fair weighted fair queuing," Proc. INFOCOM, San Francisco, CA, USA, 1996, pp. 120–128.

[8] F.P. Kelly, A. Maulloo, & D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability", Journal of the Operational Research Society, Vol. 49, 1998, pp. 237–252.

[9] L. Badia, M. Lindstr, J. Zander, & M. Zorzi, "An Economic Model for the Radio Resource Management in Multimedia Wireless Systems," Elsevier's Computer Communications, Vol. 27, 2004, pp. 1056–1064.

[10] H. R. Varian, "Intermediate Microeconomics: A Modern Approach." Norton, New York, 1999.

[11] G. Bianchi, A. T. Campbell, & R. R.-F. Liao, "On utility-fair adaptive services in wireless networks," Proceedings of IWQoS, Napa, CA, USA, 1998, pp. 256–267.

[12] D. Famolari, N. Mandayam, D. Goodman, & V. Shah, "A New Framework for Power Control in Wireless Data Networks: Games, Utility and Pricing", in R. Ganesh, K. Pahlavan, & Z. Zvonar, Wireless Multimedia Network Technologies, Kluwer Academic Publishers, 1999, chapter 1, pp. 289–310.

[13] L. Badia & M. Zorzi, "On Utility-based Radio Resource Management with and without Service Guarantees," Proc. ACM/IEEE MSWiM'04, Venice, Italy, 2004, pp. 244-251.

[14] F. Berggren, S.-L. Kim, R. Jäntti & J. Zander, "Joint power control and intra-cell scheduling of DS-CDMA non-real time data," IEEE JSAC, Vol. 19, No. 10, 2001, pp. 1860–1870.

[15] H. Jiang & S. Jordan, "A Pricing Model for High Speed Networks with guaranteed quality of service," Proc. INFOCOM, San Francisco, CA, USA, 1996, Vol. 2, pp. 888–895.

[16] C. Courcoubetis, F.P. Kelly, V.A. Siris, & R. Weber, "A study of simple usage-based charging schemes for broadband networks," Telecommunications Systems, Vol. 15, 2000, pp. 323–343.

[17] 3rd Generation Partnership Project, TR 25.950 V 4.0.0, "UTRA High Speed Downlink Packet Access", 2001.

[18] H-Y. Wei & R. Izmailov, "Channel-aware soft bandwidth guarantee scheduling for wireless packet access," IEEE WCNC, Atlanta, GA, USA, 2004, Vol. 2, pp. 1276-1281.

[19] T.J. Moulsley, "Throughput of HSDPA for UMTS", Proc. 2nd Int. Conference on 3G Mobile Communication Technologies, London, UK, 2001, pp. 363-367.

Biographies

Leonardo Badia was born in Ferrara, Italy, in 1977. He received the Laurea Degree (MS) in Electrical Engineering and the Ph.D. in Information Engineering from the University of Ferrara, Italy, in 2000 and 2004, respectively. In 2001 he joined the Department of Engineering of the University of Ferrara, where he is a currently a Post-Doc Researcher. During 2002 and 2003 he was on leave at the Royal Institute of Technology of Stockholm, Sweden. His research interests include energy efficient Ad Hoc Networks, transmission protocol modelling, Admission Control and economic modelling of Radio Resource Management for Wireless Networks.received his Ph.D. degree in 1989 from Oregon State University, USA. Currently he is a faculty member with the department of Information and Computer Science at KFUPM, Dhahran, SA. He has varied research interests and publications in the fields of distributed computing and web publishing.

Michele Zorzi was born in Venice, Italy, in 1966. He received the Laurea Degree and the Ph.D. in Electrical Engineering from the University of Padova, Italy, in 1990 and 1994, respectively. During the Academic Year 1992/93, he was on leave at the University of California, San Diego (UCSD), attending graduate courses and doing research on multiple access in mobile radio networks. In 1993, he joined the faculty of the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Italy. After spending three years with the Center for Wireless Communications at UCSD, in 1998 he joined the School of Engineering of the Università di Ferrara, Italy, where he became a Professor in 2000. Since November 2003, he has been on the faculty at the Information Engineering Department of the University of Padova. His present research interests include performance evaluation in mobile communications systems, random access in mobile radio networks, ad hoc and sensor networks, and energy constrained communications protocols. Dr. Zorzi currently serves on the Editorial Boards of the IEEE Transactions on Communications, the IEEE Transactions on Wireless Communications, the IEEE Transactions on Mobile Computing, the Wiley Journal of Wireless Communications and Mobile Computing and the ACM/URSI/Kluwer Journal of Wireless Networks. He was also the Editor-In-Chief of the IEEE Wireless Communications Magazine and guest editor for special issues in the IEEE Personal Communications Magazine, Energy Management in Personal Communications Systems and the IEEE Journal on Selected Areas in Communications (Multi-media Network Radios).