Resource Control for Loss-Sensitive Traffic in CDMA Networks

Vasilios A. Siris
Institute of Computer Science, FORTH and Department of Computer Science, University of Crete
P. O. Box 1385, GR 711 10, Heraklion, Crete, Greece
Email: vsiris@ics.forth.gr

Costas Courcoubetis
Department of Informatics
Athens University of Economics and Business
47A Evelpidon & 33 Lefkados, GR 113 62, Athens, Greece
Email: courcou@aueb.gr

Abstract—We investigate the problem of efficient resource control for loss-sensitive traffic in CDMA networks, using an economic modelling framework that takes into account the joint control of the transmission rate and the signal quality. Although the corresponding global optimization problem has a non-trivial structure, and we cannot in general guarantee that a solution can be found using the Lagrangian method, we have strong experimental evidence that this is possible for a wide range of user utilities. In this case the global optimum can be achieved in a decentralized manner, using shadow prices to influence the individual user resource requests. Based on this evidence, the main contribution of the paper is to discuss how existing rate control and outer-loop power control procedures can obtain a simple and attractive form that takes into account, through shadow prices, the level of demand and supply in order to achieve efficient resource utilization. Moreover, we describe and evaluate approximations of the proposed resource control model that can simplify its application, and we present extensions of the model for the case where the packet success ratio depends on the transmission rate in addition to the signal quality, and for network paths containing multiple wireless links.

I. INTRODUCTION

Flexible and efficient resource control in wireless networks is an increasingly important issue, due to the limited capacity of wireless networks, and their increasing use for delay and loss-sensitive applications, in both private and commercial environments. In this paper we investigate the application of economic modelling for resource control in Code Division Multiple Access (CDMA) wireless networks carrying loss-sensitive traffic. Although our approach is generally applicable to CDMA-based systems, we focus our discussion on Wideband CDMA (WCDMA), which is the most widely adopted third generation (3G) air interface technology [1], and identify how the proposed model affects existing control procedures in WCDMA systems. In WCDMA, data bits are spread over the entire spectrum that is available for transmission, and unique digital codes are used to separate the signals from different mobiles. WCDMA supports variable bit rate transmission with the use of variable spreading factors, which determine how much a data bit is spread in time, and with the use of multiple codes. In addition to its use in cellular systems, CDMA is receiving increasing attention for use in wireless LANs.

A unique property of wireless networks, and CDMA networks in particular, is that resource usage is determined by the transmission rate and the transmission power, which can be different for different mobile users. In WCDMA, resource control procedures include rate control, outer-loop power control, and fast closed-loop power control. Rate control is responsible for adjusting the variable bit spreading factor and the number of digital codes. Outer-loop power control is responsible for adjusting the target signal quality, determined by the bit-energy-to-noise-density ratio, in order to achieve a given packet error rate; this target signal quality is then given to fast closed-loop power control, which is responsible for adjusting the transmission power in order to achieve the target. Fast closed-loop power control operates on a faster time-scale, approximately 1500 Hz, compared to rate control and outerloop power control, which operate at a frequency typically less than 100 Hz.

Economic modelling has been identified as a flexible framework for efficient resource control in wired networks; e.g., see [2], [3], [4] and the references therein. Recently, economic modelling has also been applied for resource control in wireless networks [5], [6], [7], [8], [9], [10], [11], [12], [13]; we present a brief overview of this work later. Our work differs from this work in one or more of the following: First, we consider the joint optimization of the transmission rate and the bit-energy-to-noise-density ratio (signal quality), and investigate how the two control procedures for adjusting these quantities need to be modified in order to achieve efficient resource utilization. It is important to note that our model does not affect closed-loop power control, which can be performed as currently done in existing systems. Moreover, our approach results in mobile users selecting target bit-energy-to-noise-density ratios that lead to feasible power allocations, hence protects fast closed-loop power control from diverging. Second, our work is based on the economic modelling framework of social welfare maximization, for which we consider the particular network constraints of CDMA networks. Third, our work considers particular expressions for the user utility that capture its dependence on the average throughput and the data loss rate. In this direction we define loss-insensitive traffic to be traffic whose utility depends solely on the average throughput, and is independent of the loss rate. An example of
loss-insensitive traffic is bulk data transfer, where delays due to retransmissions are not important. For loss-sensitive traffic, on the other hand, the utility depends on both the average throughput and the loss rate. Examples of loss-insensitive traffic are real-time and interactive applications, which are sensitive to delay, hence to losses which result in retransmissions. Also, TCP traffic can be considered loss-sensitive, since its achievable throughput depends on the packet loss rate.

In this paper we present a new model for resource control in CDMA networks carrying loss-sensitive traffic, based on an economic modelling framework that takes into account the wireless resource constraints and the joint control of the transmission rate and the signal quality. The corresponding social welfare optimization problem has a non-trivial structure, and we cannot in general guarantee that a solution can be found using the Lagrangian method. For this reason we have resorted to numerical experimentation, which indicate, to our initial surprise, that for a wide range of user utilities the Lagrangian method can be applied, hence a shadow price exists and the social welfare maximum can be achieved in a decentralized manner, by having each user solve a local optimization problem.

Based on the aforementioned experimental evidence, the main contribution of the paper is to discuss how existing control procedures, namely rate control and outer-loop power control, can obtain a simple and attractive form that takes into account, through shadow prices, the level of demand and supply in order to achieve efficient resource utilization. In this direction, we show that for loss-sensitive traffic where the utility depends on the loss rate in addition to the average throughput, unlike the case of loss-insensitive traffic where the utility depends only on the average throughput, the optimal signal quality is no longer independent of the level of resource demand and supply. Hence, both rate control and outer-loop power control need to take into account shadow prices. A second contribution of the paper is to discuss and evaluate an approximation that can simplify the application of the proposed resource control model to discrete rate control and for supporting class-based service differentiation in terms of both throughput and loss-sensitivity. Finally, we extend the model to the case where the packet success ratio depends on the transmission rate, in addition to the signal quality, and to networks containing multiple wireless links.

The rest of the paper is structured as follows. In Section II we summarize results on resource control in CDMA networks carrying loss-insensitive traffic. In Section III we present and investigate a model for resource control in the case of loss-sensitive traffic, and in Section IV we discuss various issues regarding the application of the proposed model. In Section V we present extensions to the proposed model. Finally, in Section VI we present a brief overview of related work and in Section VII we conclude the paper.

II. RESOURCE CONTROL FOR LOSS-INSENSITIVE TRAFFIC

In this section we summarize background work on resource control for loss-insensitive traffic in CDMA networks, as developed in [14]; for such traffic a user’s utility depends solely on his average throughput. Consider the uplink of a single CDMA cell. Note, however, that the results can be extended to the multiple cell case by considering the inter-cell interference coefficient [15]. Let $W$ be the rate. The bit-energy-to-noise-density ratio at the base station is given by [15], [16]

$$
\frac{E_b}{N_0} = \frac{W}{r_i \sum_{j \neq i} g_{ij} p_j + \eta},
$$

where $r_i$ is the transmission rate, $p_i$ is the transmission power, $g_i$ is the path gain between the base station and mobile $i$, and $\eta$ is the power of the background noise at the base station. The ratio $W/r_i$ is the spreading factor or processing gain for mobile $i$.

The value of the bit-energy-to-noise-density ratio $(E_b/N_0)_i$ corresponds to the signal quality, since it determines the bit error rate, $BER$ [15], [16]. Under the assumption of additive white Gaussian noise, $BER$ is a non-decreasing function of $E_b/N_0$, that depends on the multipath characteristics, and the modulation and forward error correction (FEC) algorithms.

Let $\gamma_i$ be the target bit-energy-to-noise-density ratio required to achieve a target bit or packet error rate; the adjustment of $\gamma_i$ is performed by outer-loop power control. The target $\gamma_i$ is given to closed-loop power control, which adjusts the transmission power in order to achieve it. If we assume perfect power control, then $(E_b/N_0)_i = \gamma_i$. Solving the set of equations (1) for each mobile $i$, and assuming that there is a large number of mobiles, each using a small portion of the available resources, one can show that the wireless resource constraint can be approximated by [16], [17], [1]

$$
\sum_i r_i \gamma_i < W.
$$

Hence, the amount of resources used by a mobile is given by the product of the transmission rate and the target bit-energy-to-noise-density ratio (signal quality). In actual systems, due to the limited transmission power of the mobile hosts, imperfect power control, shadowing, and inter-cell interference, the total load must be well below one. Indeed, in radio network planning [1], all the above factors are used to determine an interference margin (or noise rise) $I_{\text{margin}}$, based on which the wireless resource constraint becomes

$$
\sum_i r_i \gamma_i < \rho^{UL} W \quad \text{where} \quad \rho^{UL} = \frac{I_{\text{margin}} - 1}{I_{\text{margin}}}. \quad (2)
$$

In the case of loss-insensitive traffic, where users value only the average throughput of successful data transmission, which is given by the product $r_i P_i(\gamma_i)$ of the transmission rate and the probability of successful packet transmission, the utility is given by

$$
U_i(r_i, \gamma_i) = U_i(r_i P_i(\gamma_i)) .
$$
Consider the global problem of maximizing the aggregate utility (social welfare)

\[
\text{maximize } \sum_i U_i(r_i P_i(\gamma_i)) \\
\text{over } \sum_i r_i \gamma_i < \rho^{W}. \\
\text{subject to } \sum_i r_i \gamma_i < \rho^{s}. \\
\]

By setting \( z_i = r_i \gamma_i \), and assuming \( \gamma_i > 0 \), the above problem can be written as

\[
\text{maximize } \sum_i U_i \left( z_i P_i(\gamma_i) \right) \\
\text{over } z_i \geq 0, \gamma_i > 0 \\
\text{subject to } \sum_i z_i < \rho^{s}. \\
\]

which is equivalent to

\[
\text{maximize } \sum_i U_i \left( z_i \max_{\gamma_i > 0} \frac{P_i(\gamma_i)}{\gamma_i} \right) \\
\text{over } z_i \geq 0 \\
\text{subject to } \sum_i z_i < \rho^{s}. \\
\]

Hence, the global problem of maximizing the social welfare can be decomposed into two separate problems, the first gives an optimal signal quality \( \gamma_i^* \) for each user \( i \) that satisfies

\[
P_i^*(\gamma_i^*) = P_i(\gamma_i), 
\]

which is independent of the user utility, and the second problem is an optimization over \( \{z_i\} \)

\[
\text{maximize } \sum_i U_i \left( z_i P_i(\gamma_i^*) \right) \\
\text{over } z_i \geq 0 \\
\text{subject to } \sum_i z_i < \rho^{s}. 
\]

If \( U_i(\cdot) \) is differentiable and strictly concave, then the Lagrangian method can be applied, hence the global optimization problem is equivalent to each rational user solving the problem

\[
\text{maximize } U_i(r_i P_i(\gamma_i)) - \lambda r_i \gamma_i \\
\text{over } r_i \geq 0, \gamma_i \geq 0, 
\]

where \( \lambda \) is the shadow price for the wireless resource constraint \( (2) \). In practise, the shadow price can be iteratively adjusted based on a tatonnement process, so that the aggregate demand equals the available wireless resources.

The above results allow the decoupling of the two problems of selecting the optimal \( \gamma_i^* \) and the optimal transmission rate \( r_i^* \). Moreover, the optimal \( \gamma_i^* \) is independent of the level of resource demand and supply (shadow price); this is not the case for loss-sensitive traffic as we will see in the next section.

### III. Resource Control for Loss-Sensitive Traffic

In this section we first propose a utility for loss-sensitive traffic that takes into account the dependence on the average throughput and the loss rate, followed by a model for efficient resource control in CDMA networks based on an economic modelling framework. Then we present and evaluate an approximation of the proposed resource control model that can simplify its application.

#### A. Utility for loss-sensitive traffic

Consider a user whose utility depends, in addition to the average throughput, on the loss rate he is experiencing. If the latter affects the utility in a subtractive way, we have the following expression for user \( i \)’s utility

\[
U_i(r_i, \gamma_i) = U_i(r_i P_i(\gamma_i)) - V_i(r_i (1 - P_i(\gamma_i))), 
\]

where \( U_i(x_i) \) represents user \( i \)'s valuation for average throughput \( x_i \), and \( V_i(y_i) \) represents the decrease of the user’s utility for data loss rate equal to \( y_i \).

The global problem of maximizing the aggregate utility (social welfare) is

\[
\text{maximize } \sum_i \left[ U_i(r_i P_i(\gamma_i)) - V_i(r_i (1 - P_i(\gamma_i))) \right] \\
\text{over } r_i \geq 0, \gamma_i \geq 0 \\
\text{subject to } \sum_i r_i \gamma_i < \rho^{s}. 
\]

The above problem has a non-trivial structure, and involves a non-convex constraint over the two variables \( r_i, \gamma_i \). Hence, unlike the case for loss-insensitive traffic, we cannot in general guarantee that a solution can be found using the Lagrangian method. Nevertheless, we have strong experimental evidence that this is possible for a wide range of user utilities. Here we present a subset of these experiments with both concave exponential and logarithmic functions for \( U_i \), and convex exponential functions for \( V_i \). Figs. 1 and 2 show, for various mixes of up to three different utility types, the optimal social welfare \( SW^* \) in \( (6) \) as a function of the total resource constraint \( \rho^{s} \). The Lagrangian method can be applied when \( SW^*(\rho^{s}) \) is a concave function of the total available resources \( \rho^{s} \); the figures show that this is indeed the case.

#### B. Resource control model

Considering the utility function for loss-sensitive users given by \( (5) \), and assuming as our experimental evidence indicates that the Lagrangian method can be applied, the user optimization problem is

\[
\text{maximize } U_i(r_i P_i(\gamma_i)) - V_i(r_i (1 - P_i(\gamma_i))) - \lambda r_i \gamma_i \\
\text{over } r_i \geq 0, \gamma_i \geq 0 , 
\]

where as in the case of loss-insensitive traffic, \( \lambda \) is the shadow price for the wireless resource constraint \( (2) \).
If there exists \( r_1^* > 0 \) and \( \gamma_1^* > 0 \) that achieve the maximum in (7), then from the first order conditions we have

\[
P_i'(\gamma_1^*) \gamma_1^* = P_i(\gamma_1^*) \frac{\lambda \gamma_1^*}{\lambda \gamma_1^* + V_i'(r_i^*(1 - P_i(\gamma_1^*))}).
\] (8)

The last equation is a generalization of (3), since for loss-insensitive traffic we have \( V_i'(\cdot) = 0 \). Hence, whereas for loss-insensitive traffic the optimal signal quality depends only on \( P_i(\gamma_1) \), for loss-sensitive traffic it also depends on the shadow price and the user’s sensitivity to losses. It is interesting to observe, nevertheless, that the optimal signal quality does not depend on the user’s valuation for his average throughput.

In the case of loss-sensitive traffic, as expected, the optimal signal quality \( \gamma_2^* \) is higher than in the case of loss-insensitive traffic, since the ratio on the right-hand side of (8) is smaller than one, and the packet success ratio as a function of \( \gamma \) has a sigmoid shape, Fig. 3. Indeed, the above ratio is the ratio of the marginal charge due to the wireless constraint, \( \lambda \gamma_2^* \), and the total marginal cost incurred by the user, which includes the marginal charge due to the wireless constraint and the marginal cost due to losses, expressed by \( V_i'(\cdot) \). Finally, observe that a higher sensitivity to losses corresponds to a smaller ratio on the right-hand side of (8), hence to a higher optimal signal quality.

Next we present an approximation of the resource control model presented above that can simplify its application. The approximation is motivated by the fact that, for different modulation schemes, see Fig. 4, the optimal target signal quality is such that the packet success ratio obtains large values, typically larger than 80%. Moreover, as discussed above, loss-sensitive traffic has a larger optimal target signal quality compared to loss-insensitive traffic.

Based on the previous discussion, which suggests that in practical situations \( P_i(\gamma_1^*) \) will be close to one, we can consider the following approximation for (8) that is expected to be accurate when \( r_i(1 - P_i(\gamma_1)) \) is close to zero and the derivative of \( V_i(y_i) \) for loss rates \( y_i \) close to zero does not change.
where expression in (7), and find the optimal packet loss rate [18]. Indeed, if we substitute the last equation to TCP’s well-known inverse square root dependence on the corresponding average throughput has a behavior similar to the numerator of the right-hand side, then the average throughput depends on the packet loss rate. If we assume that a TCP segment fits in a single physical layer packet, then we can define the implicit utility for TCP to be
\[
U_{tcp}(r, \gamma) = -\frac{2}{T^2 r P(\gamma)} - r(1 - P(\gamma)),
\]
where \(T\) is the round trip time. For the above implicit utility, the corresponding average throughput has a behavior similar to TCP’s well-known inverse square root dependence on the packet loss rate [18]. Indeed, if we substitute the last expression in (7), and find the optimal \(r^*, \gamma^*\), then we obtain the following average throughput
\[
r^* P(\gamma^*) = \frac{1}{T} \sqrt{\frac{2P(\gamma^*)}{1 - P(\gamma^*) + \lambda \gamma^*}},
\]
with \(\gamma^*\) satisfying
\[
P'(\gamma^*) \gamma^* = P(\gamma^*) \frac{\lambda \gamma^*}{\lambda \gamma^* + 1}.
\]
If \(P(\gamma^*)\) in (11) is assumed close to one to one and is omitted from the numerator of the right-hand side, then the average throughput, \(r^* P(\gamma^*)\), is inversely proportional to the square root of the sum of the loss probability, \(1 - P(\gamma^*)\), and the congestion price, \(\lambda \gamma^*\).

If a single TCP segment is broken into \(k\) physical layer packets, then the loss rate is \(r(1 - P(\gamma)^k)\), which using the Taylor expansion can be approximated by \(r k (1 - P(\gamma))\). Hence, TCP’s implicit utility in this case can be defined as
\[
U_{tcp,k}(r, \gamma) = -\frac{2}{T^2 r P(\gamma)} - r k (1 - P(\gamma)).
\]

\[\text{Fig. 5. Loss-sensitivity } V(y) = d(e^{cy} - 1). \text{ The loss-sensitivity increases from right to left: a higher value of } c \text{ in (13) corresponds to a higher loss-sensitivity. } d = 0.05.\]

From the last equation observe that the effect of losses is higher when a TCP segment is broken into multiple physical layer packets, compared to the case when a TCP segment fits in a single packet. In the former case, (12) becomes
\[
P'(\gamma^*) \gamma^* = P(\gamma^*) \frac{\lambda \gamma^*}{\lambda \gamma^* + k}.
\]

D. Numerical investigations

In this subsection we investigate the resource control model and its approximation that were presented in the previous subsections. The factor in (5) that gives the valuation for average throughput \(x\) is assumed to be
\[
V(x) = 1 - e^{-bx}.
\]

Although the investigations reported in this paper consider the above throughput utility factor, the qualitative conclusions are similar for other concave functions. The loss-sensitivity factor in (5) is assumed to be
\[
V(y) = d(e^{cy} - 1),
\]
where \(y\) is the loss rate, \(d\) depicts the relative importance of losses compared to the average throughput, and \(c\) reflects the loss-sensitivity, with a larger value of \(c\) corresponding to a higher loss-sensitivity, Fig. 5.

1) Effect of loss-sensitivity: The effect of the loss-sensitivity on the optimal \(\gamma^*\) is shown in Fig. 6. The optimal \(\gamma^*\) is given by the intersection of the curve \(P'(\gamma)\) with the curve \(P(\gamma)/\gamma\) in the case of loss-insensitive traffic (3), and with the curve
\[
F(\gamma) = \frac{P(\gamma)}{\gamma} \frac{\lambda \gamma}{\lambda \gamma + V'(r(1 - P(\gamma)))},
\]
in the case of loss-sensitive traffic (8). Observe that the optimal signal quality \(\gamma^*\) in the case of loss-sensitive traffic is larger than in the case of loss-insensitive traffic. Furthermore, a larger loss-sensitivity, i.e. a larger value of \(c\) in (13), results in a higher optimal signal quality \(\gamma^*\). Indeed, for the values considered, see Fig. 3, the values of \(\gamma^*\) for loss-sensitive traffic correspond to packet success ratios \(P(\gamma)\) larger than 90%.
A higher loss-sensitivity results in a higher optimal signal quality $\gamma^*$. $r = 20$ Kbps, $d = 0.05$, $\lambda = 0.002$.

2) Accuracy of the approximate model: Fig. 7 shows, for different loss-sensitivities, the accuracy of (9), which is an approximation of (8). Observe that the approximation is more accurate when the derivative of the loss-sensitivity factor does not change significantly in the neighborhood of loss rates close to zero, which is the case for small values of $c$.

The previous results were for a particular value of the shadow price $\lambda$. Fig. 8 shows the accuracy of approximation (9) for different values of $\lambda$. Observe that the approximation is very accurate for small and large values of $\lambda$, the latter corresponding to high demand. Moreover, observe that for large values of the shadow price, the optimal signal quality for loss-sensitive traffic approaches the optimal signal quality for loss-insensitive traffic; this is expected since for large values of the shadow price the ratio on the right-hand side of (8) approaches one, hence (8) approaches the corresponding equation for loss-insensitive traffic (3).

We end this discussion by noting that a heuristic for (10) is

$$\text{maximize } U_i(r_i) - \lambda r_i \gamma_i^*(\lambda) $$
$$\text{over } r_i \geq 0,$$

which is expected to be accurate when $P_i(\gamma_i)$ is close to one, in which case $r_i P_i(\gamma_i)$ and $r_i (1 - P_i(\gamma_i))$ are close to $r_i$ and zero respectively, and $V_i(y_i)$ for loss rates $y_i$ close to zero is small compared to $U_i(r_i)$. The last expression for the user optimization problem is attractive since it has the same form with the corresponding problem in the case of loss-insensitive traffic (4). The only difference is that now the bit-energy-to-noise-density ratio is not constant, but depends on the shadow price; this dependence is captured through the function $\gamma_i^*(\lambda)$.

IV. APPLICATION

Unlike the case of loss-insensitive traffic, for loss-sensitive traffic the optimal target bit-energy-to-noise density ratio should depend on the shadow price. Hence, in order to achieve efficiency both the outer-loop power control and the rate control procedures need to take into account the shadow price, Fig. 9, in order to solve (9) and (10), respectively. Observe that the optimal target bit-energy-to-noise-density ratio is used as input to both the rate control and the closed-loop power control procedures. For the uplink, outer-loop power control is performed at the Radio Network Controller (RNC) of a WCDMA cell, and rate control can be implemented at the mobile host or at the RNC. An example of the latter case is a particular instantiation of the proposed resource control model, which we describe later in this section, that supports class-based service differentiation and can be implemented solely at the RNC. Note that it is not our objective to discuss in depth all possible implementation details, but rather to highlight some important implementation aspects arguing that the proposed resource control model does not require radically different procedures than those already implemented in existing systems.

A. Outer-loop power control

In current systems, outer-loop power control adjusts the target signal quality in order to achieve a predefined packet error rate, which is typically 10-20% for non-real-time services, and 1% for real-time services. As indicated by the models.
Fig. 9. The proposed resource control model affects only the rate control and the outer-loop power control procedures (shaded boxes), which must take into account the shadow price $\lambda$. The target signal quality $\gamma^*$ is used as input to both rate control and closed-loop power control.

presented in Sections II and III, this does not lead to efficient resource utilization.

Current systems implement outer-loop power control by increasing (or decreasing) the target bit-energy-to-noise-density ratio depending on whether the measured packet error rate is larger (or smaller) than the target packet error rate [1]. Due to the sigmoid shape of $P(\gamma)$, a similar procedure can be used to achieve (9). In particular, the signal quality $\gamma$ is increased with step $\Delta \gamma$ if

$$P'(\gamma) \approx \frac{P(\gamma + \Delta \gamma) - P(\gamma)}{\Delta \gamma} > P(\gamma) \frac{\lambda \gamma}{\lambda \gamma + V'(0)},$$

and is decreased with step $\Delta \gamma$ otherwise.

B. Discrete rate control

In WCDMA, rates can obtain only discrete values, corresponding to spreading factors that are powers of 2. In the uplink, the spreading factor can take values from 256, giving a channel bit rate of 15 Kbps, to 4, giving a channel bit rate of 960 Kbps; higher bit rates are achieved by using up to 6 parallel codes with spreading factor 4 (giving a channel bit rate of 5740 Kbps). Hence, in the uplink, assuming 1/2 rate coding, the user data rates that can be achieved using a single code are 7.5, 15, 30, 60, 120, 240, and 480 Kbps [1].

Assume that there exists direct feedback of the shadow price from the Base Station/Radio Network Controller (BS/RNC) to the mobile stations, Fig. 10(a). This feedback can be provided over an existing radio control channel or over a higher layer signalling mechanisms, such as Explicit Congestion Notification (ECN) in the IP layer. Based on this feedback, the mobile stations can adjust their transmission rate as follows:

- An initial selection of $r$ is made. Since there are only seven possible rate values, a linear search for the rate that maximizes (10) can be sufficient.
- Neighboring rates are examined to see whether they yield a larger user benefit (utility minus charge) in (10). If a neighboring rate is found to give a higher benefit, then the next rate in the same direction is examined; this procedure is repeated throughout the duration of the connection.

C. Class-based service differentiation

In this section we present a particular instantiation of the model discussed in Section III-B that supports service differentiation based on weights, where different weights can correspond to different classes.

Assume that user $i$’s valuation for his average throughput is given by

$$U_i(r_i P_i(\gamma_i)) = u_i \log(r_i P_i(\gamma_i)),$$  
and his loss-sensitivity factor is

$$V_i(i(1 - P_i(\gamma_i))) = v_i r_i (1 - P_i(\gamma_i)).$$

From the first order conditions of (7) we get

$$u_i = v_i r_i^* (1 - P_i(\gamma_i^*)) + \lambda r_i^* \gamma_i^*.$$  (14)

The right-hand side in the last equation can be interpreted as the cost user $i$ is willing to incur, which includes the cost in terms of lost data and the charge due to the use of wireless resources.

For the utility corresponding to the average throughput and loss-sensitivity factors identified above, the outer-loop power control objective is given by (9) with $V^*_i(0) = v_i$. Based on (14), the rate for user $i$ can be computed from

$$r_i^* = \frac{u_i}{v_i (1 - P_i(\gamma_i^*)) + \lambda \gamma_i^*}.$$  (15)

In practise, a network provider can offer a small set of possible values for $u_i$, each corresponding to a different throughput class, and a small set of values for $v_i$, each corresponding to a different loss-sensitivity class.

Note that the above rate selection can be performed at the Radio Network Controller (RNC), rather than at the mobile stations. After the rate selection, the RNC communicates the rate values to the mobile stations, Fig. 10(b). Such an approach
places more control at the RNC, which in current systems has the intelligence for supporting flexible packet scheduling and load control [1]. Moreover, it does not require the communication of shadow prices to the mobile stations, since the selection of both the target bit-energy-to-noise-density ratio and the transmission rate is performed at the RNC, based on (9), with $V_i'(0) = v_i$, and (15), respectively. On the other hand, it does require the communication of the transmission rates from the RNC to the mobile stations, which however is already supported in current systems. Finally, note that the approach for adjusting discrete rates that was presented in the previous subsection can be combined with the approach for class-based service differentiation that was discussed in this subsection.

V. EXTENSIONS

Next we present two extensions of the resource control model discussed in Section III. First, we extend the model to the case where the packet success ratio depends on the transmission rate, in addition to the bit-energy-to-noise-density ratio. Second, we extend the model to the case of a network containing multiple wireless links.

A. Dependence of packet success ratio on rate

Up to now we have assumed that the packet success ratio $P_i(\gamma_i)$ is a function of only the target bit-energy-to-noise-density ratio. In practice the packet success ratio also depends on the transmission rate, hence we can write $P_i(\gamma_i, r_i)$. Indeed, the required signal quality in order to achieve the same packet success ratio is smaller for higher bit rates. This is because the performance of closed-loop power control depends on the accuracy of the channel and bit-energy-to-noise-density ratio estimation algorithms, which are based on reference symbols carried over the physical control channels. The more power allocated to the control channels, the more accurate the estimation procedure. The power levels for the control channels is typically higher for higher bit rates, hence higher bit rates yield better performance in terms of the packet success ratio, for the same target bit-energy-to-noise-density ratio.

1) Loss-insensitive traffic: First consider the case of loss-insensitive traffic, where the utility depends solely on the average throughput. If the packet success ratio is of the form $P_i(\gamma_i, r_i)$, then the user optimization problem is

$$\text{maximize} \quad U_i(r_i P_i(\gamma_i, r_i)) - \lambda r_i \gamma_i$$

$$\text{over} \quad r_i \geq 0, \gamma_i \geq 0,$$

where $\lambda$ is the shadow price for the wireless resource constraint (2).

A necessary condition for achieving the optimal in (16), assuming that $r_i, \gamma_i > 0$, is the following

$$\frac{\partial P_i}{\partial \gamma} \bigg|_{\gamma_i^*, r_i^*} \gamma_i^* = \frac{\partial P_i}{\partial r} \bigg|_{\gamma_i^*, r_i^*} r_i^*.$$

From the above discussion on the effect of the transmission rate on the packet success ratio, we have $\frac{\partial P_i}{\partial r} > 0$, hence from the last equation the optimal $\gamma_i^*$ is now lower than when there is no dependence of the packet success ratio on the transmission rate.

2) Loss-sensitive traffic: In the case of loss-sensitive traffic, where the user utility is given by (5), the user optimization problem is

$$\text{maximize} \quad U_i(r_i P_i(\gamma_i, r_i)) - V_i(r_i(1 - P_i(\gamma_i, r_i))) - \lambda r_i \gamma_i$$

$$\text{over} \quad r_i \geq 0, \gamma_i \geq 0,$$

where $\lambda$ is the shadow price for the wireless resource constraint (2). From the first order conditions we have

$$\frac{\partial P_i}{\partial \gamma} \bigg|_{\gamma_i^*, r_i^*} = \frac{\partial P_i}{\partial r} \bigg|_{\gamma_i^*, r_i^*} r_i^*,$$

which is a combination of (8) and (17).

B. Extension to networks with multiple wireless hops

Next we present the extension of the proposed resource control model to networks containing multiple wireless hops. This would be the case of two mobile stations both communicating through WCDMA access links. As we will see, the optimal signal qualities on the wireless links are no longer independent. Although the models we present in this section consider only two wireless links, the extension to more than two links can be done in a straightforward way, although the equations become more lengthy. Note that, for loss-insensitive traffic, the extension to networks containing fixed links was presented in [14].

For the uplink, the wireless resource constraint is given by (2). In the downlink, the constrained resource is the total transmission power at the base station, which can lead to a resource control model that is different from the one in the uplink. Nevertheless, one can show, see [1], that for resource dimensioning and network planning the downlink constraint can be approximated by an inequality similar to (2)

$$\sum_i r_i \gamma_i \leq P_i^\text{base},$$

where $P_i^\text{base} \leq 1$ depends on the total base station power, the noise, the average attenuation from the base station to the mobiles, the average downlink orthogonality factor, and the average inter-cell interference.

1) Loss-insensitive traffic: Consider a network with two wireless links, with packet success ratio $P_1(\gamma_1)$ and $P_2(\gamma_2)$, respectively, where $\gamma_1$ and $\gamma_2$ are the target signal qualities on the two links. The user optimization problem for loss-insensitive traffic is the following (to simplify the notation, we drop the subscripts identifying different users)

$$\text{maximize} \quad U(r P_1(\gamma_1) P_2(\gamma_2)) - \lambda_1 r \gamma_1$$

$$\text{over} \quad r \geq 0, \gamma_1 \geq 0, \gamma_2 \geq 0,$$
where $\lambda_1$, $\lambda_2$ are the shadow prices for the constraints at the two wireless links. Observe that the rate on the second link is $rP_2(\gamma_1)$, since some packets will be lost on the first wireless link. From the first order conditions of the above optimization problem, and after some manipulation one can show that the optimal $\gamma_1^*$ and $\gamma_2^*$, assuming that they exist and $r^*, \gamma_1^*, \gamma_2^* > 0$, satisfy
\[
P'_1(\gamma_1^*)\gamma_1^* = P_1(\gamma_1^*),
\]
\[
P'_2(\gamma_2^*)\gamma_2^* = \frac{\lambda_2 P_1(\gamma_1^*)\gamma_2^*}{\lambda_1 \gamma_1^* + \lambda_2 P_1(\gamma_1^*)\gamma_2^*} P_2(\gamma_2^*).
\]
Hence we observe that in the case of two wireless links, for loss-insensitive traffic, the optimal target bit-energy-to-noise-density ratio for the first link is the same as that in the single link case. Observe that the ratio on the right-hand side of the last equation for $\gamma_2^*$ denotes the percentage of the total charge due to the second wireless link. Hence, the optimal target bit-energy-to-noise-density ratio for the second link is larger than what it would be if it were the only link. Moreover, if the percentage of the charge due to the second link decreases, then $\gamma_2^*$ increases.

2) Loss-sensitive traffic: In the case of loss-sensitive traffic, the user optimization problem is
\[
\text{maximize} \quad U(rP_1(\gamma_1)P_2(\gamma_2)) - V(r(1 - P_1(\gamma_1)P_2(\gamma_2))) - \lambda_1 r\gamma_1 - \lambda_2 rP_1(\gamma_1)\gamma_2 \quad \text{over} \quad r_1 \geq 0, \gamma_1 \geq 0, \gamma_2 \geq 0,
\]
where $\lambda_1$, $\lambda_2$ are the shadow prices for the constraints at the two wireless links. From the first order conditions of the above optimization problem, and after some manipulation one can show that the optimal $r^*$, $\gamma_1^*$, and $\gamma_2^*$, assuming that they exist and are positive, satisfy
\[
P'_1(\gamma_1^*)\gamma_1^* = \frac{\lambda_1 \gamma_1^*}{\lambda_1 \gamma_1^* + V'(r^*(1 - P_1(\gamma_1^*)P_2(\gamma_2^*)))} P_1(\gamma_1^*),
\]
which is the same as the case for a single wireless link, (8). For the second wireless link we have
\[
P'_2(\gamma_2^*)\gamma_2^* = \frac{\lambda_2 P_1(\gamma_1^*)\gamma_2^*}{\lambda_1 \gamma_1^* + \lambda_2 P_1(\gamma_1^*)\gamma_2^* + V'(r^*(1 - P_1(\gamma_1^*)P_2(\gamma_2^*)))} P_2(\gamma_2^*),
\]
where the ratio on the right-hand side is the percentage of the total charge that is due to the second wireless link, assuming that the total charge is the sum of the marginal charge of the two links and the cost due to losses $V'(\cdot)$.

The results of this section identify how the optimal target signal qualities on the two wireless links are related. An interesting open issue is how the two links can communicate in order to coordinate their selections.

VI. Related work

Next we present a brief overview of related research, identifying its differences with the work presented in this paper.

The work in [5] considers a utility that is interpreted as the number of information bits transmitted per unit of energy. It is shown that the non-cooperative game, where mobiles adjust their power to maximize their utility, has a unique Nash equilibrium, which however is inefficient. With the introduction of prices [6], Pareto improvements are achieved, but not the social welfare optimal. The work in [10] considers a utility that is a monotonically increasing concave function of the bit-energy-to-noise-density ratio and a monotonically decreasing concave function of the mobile’s power, and prove the existence of a Nash equilibrium. The work in [13] considers a weighted logarithmic utility function of the signal quality and a price proportional to the transmitted uplink power, and prove the existence of a Nash equilibrium.

The work in [7] considers a utility that is a function of the transmission rate, and investigate the problem of maximizing the sum of all utilities in the forward link (downlink), under constraints on the total transmission power at the base station, and constraints on the maximum error rate for each user. Our work considers user utilities that depend on the loss rate, in addition to the average throughput. The work in [9], [11], [12] also considers the downlink power control taking into account the constraint on the total transmission power. The work in [9] considers a user utility that is a step function of the bit-energy-to-noise-density ratio, and a mobile’s charge contains a constant term (price per code) and a term linear in the transmitted power from the base station. The work in [11] focuses on revenue maximization through pricing, where prices are per unit of time. The work in [12] investigates social welfare maximization in a multi-class CDMA system, and proposes an algorithm that can obtain a Pareto optimal allocation, which is a good approximation of the social optimal power allocation.

The work in [8] considers a utility that is a function of the bit-energy-to-noise-density ratio, which can have a sigmoid shape, and formulate a utility-based distributed power control algorithm where each user seeks to maximize his net utility, and charges are proportional to the power. For a constant price per unit of power, it is proved that the power update algorithm converges.

Our work differs from the above in that it considers the joint optimization of the signal quality and transmission rate, and takes into account the particular resource constraints in the uplink. Moreover, we discuss the form of the utility function for loss-sensitive traffic, taking into account both the user’s valuation for his average data throughput and his sensitivity to losses. The aforementioned work is geared towards mechanisms for power control; on the other hand, our work deals with control mechanisms, namely rate control and outer-loop power control, that operate on a slower timescale, hence on top of fast closed-loop power control, without modifying it. Finally, in the proposed resource control model there is no differentiation of mobile users based on their position. On the other hand, in the approaches of [5], [6], [8], [13], [11], mobile users far from the base station that encounter high path loss face a higher charge and receive less resources, compared
to users close to the base station; this is termed ‘near-far unfairness’ in [8].

VII. CONCLUDING REMARKS

We have presented a new model for resource control in CDMA networks carrying loss-sensitive traffic, based on an economic modelling framework that takes into account the wireless resource constraints and the joint control of the transmission rate and the signal quality. The corresponding social welfare optimization problem has a non-trivial structure, for which we cannot in general guarantee that a solution can be found using the Lagrangian method. Nevertheless, we have strong experimental evidence that this is possible for a wide range of user utilities, in which case the social welfare maximum can be achieved in a decentralized manner by having each user solve a local optimization problem. Based on the above evidence, the main contribution of the paper is to discuss how existing control procedures, namely rate control and outer-loop power control, can obtain a simple and attractive form that takes into account, through shadow prices, the level of demand and supply in order to achieve efficient resource utilization. Moreover, we describe and evaluate an approximation of the proposed resource control model that can simplify its application, without requiring radically different procedures than those implemented in existing systems. One application scenario, which can be implemented solely at the Radio Network Controller (RNC) of a WCDMA network, can support service differentiation based on throughput and loss-sensitivity classes.

Ongoing work includes investigating in depth the particular characteristics of the social welfare maximization (6), to understand why and when the Lagrangian method can be applied for the various user utilities that we have experimentally investigated. Our initial findings suggest that this is related to the fact that the optimal signal quality for loss-sensitive traffic is higher than for loss-insensitive traffic, and corresponds to packet success ratios close to one. Other related work includes the application of economic modelling for resource control and service differentiation in wireless LANs based on the IEEE 802.11 standard.

ACKNOWLEDGMENT

The authors would like to thank Bob Briscoe and Dave Songhurst for helpful discussions on parts of this work.

REFERENCES