Hierarchical SIR and Rate Control for CDMA Data Users on the Forward Link *

Lei Song and Narayan B. Mandayam
Wireless Information Network Laboratory (WINLAB)
Department of Electrical & Computer Engineering
Rutgers University, Piscataway, NJ 08854

Abstract - In this paper, we consider SIR and rate control algorithms for data users on the CDMA forward link under constraints on delay (i.e., limited allowable retransmissions), data error rate and total power. The QoS for data users is mapped onto a family of utility functions that reflect the tradeoff between total system throughput and fairness (similarity in data rates of users). The proposed algorithm uses an hierarchical control structure that is simple to implement in a distributed manner. Specifically, the data rates are adjusted jointly by the base station while the SIR targets are adjusted by the mobiles using information specific to the individual users.

1 Introduction

In CDMA networks, system resources such as signal to interference ratio (SIR) and data rate are controlled to maintain overall system feasibility and optimize network performance measures such as delay and throughput for data. Most of the research on SIR and rate control for integrated voice/data are based on simplifications such as characterizing a type of data service with a minimum SIR requirement or assuming an errored frame will be retransmitted repetitively until it goes through [1–4]. In [5], a hierarchical SIR and rate control algorithms is proposed for the reverse link of practical systems such as [6,7]. In these systems, a limited retransmission protocol called Radio Link Protocol (RLP) is implemented for data applications. In this paper, we present a similar hierarchical algorithm for the forward link. Unlike the reverse link which is usually interference limited, the forward link capacity is often limited by the power. Similar to the approach in [5], we characterize the QoS measure for a data user by a utility function that is a measure of the achieved throughput. The proposed algorithm maximizes the utility functions under constraints on the post-RLP error rate (PRER) and average total power. The algorithm can be implemented in a distributed fashion and requires very limited exchange of information between the base station and mobile stations.

2 System Model

In the third generation (3G) CDMA systems, on the forward link [7], each mobile selects its own SIR target to achieve a certain frame error performance. After the mobile selects its SIR target, it then issues a sequence of power control bits to instruct the base station to either increase or decrease its power to meet the SIR target. Such an up/down power control loop was shown to be able to achieve the SIR target [8]. The SIR target is determined by each mobile based only on the local channel information for a given user. In this paper, we assume the local information is given in the form of the forward link frame error rate for a SIR target, i.e., for a SIR γi, the FER f(γi) is assumed to be known. We further assume that the function f(γ) be analytic, monotonically decreasing and convex.

The data rates for the users are decided jointly by the base station to meet an average total power constraint. Such a power limit is necessary not only to protect the base station amplifier but also to reduce the interference to the neighboring cells. We assume all the data users are synchronous at the frame level and data is transmitted in segments as shown in Figure 1. The data rate for a user remains the same throughout a segment which lasts for multiple frames. The RLP retransmits an errored frame up to M times. For the original transmission of a frame from user i, the mobile sets a target SIR γi,1. If the frame is received in error, the mobile sends a NAK to the base station using the reverse link control channel asking for retransmission of this frame. For the jth retransmission of a frame, the SIR target is set at γi,j. The SIR vector γ = {γi,1, ..., γi,M} represents the SIR targets to be used in the data segment. The data rate Ri remains the same throughout the whole data segment. We further require that the segments are synchronized for all data users.

The RLP protocol limits the delay for data by only allowing a limited number of retransmissions. Such an RLP protocol does not guarantee error free data transmission. The residual error, which we call Post-RLP Error Rate (PRER), must be within a tolerable limit to ensure proper data transmis-
sion. Given that the delay and PRER constraints are met, we choose the QoS for data users to reflect the throughput and fairness.

There are many ways to define this QoS. Recently, there has been a lot of work on using utility functions to describe QoS for data in wireline [9] and wireless [10] communications. In this paper, we adopt the utility function based approach. We can define the utility function $U_i(R_i)$ for user $i$ as a function of the data rate $R_i$. We assume the utility function to be concave and monotonically increasing. By adjusting the shape of $U_i(\cdot)$, a flexible trade-off can be achieved between the overall optimality (total system data rate) and fairness (data rate discrepancy) (see Figure 4).

3 Problem Formulation

Consider a CDMA system with $N$ data users. The control algorithm will find a set of SIR and rate parameters for each data user at the beginning of a data segment such that the total utility is optimized for the segment. The constraints are that (a) the average total forward power for all the data users should be within a limit $C$; and (b) for each individual user, the PRER should be within a limit $\epsilon_0$.

Mathematically, the optimization problem can be formulated as choosing the vector $S_i = \{\gamma_i, R_i\} = \{\gamma_{i,1}, ..., \gamma_{i,M}, R_i\}$ for $i = 1, ..., N$, such that

$$\max_{S_1, ..., S_N} \sum_{i=1}^{N} U_i(R_i) \quad \text{(1)}$$

subject to

$$\sum_{i=1}^{N} P_i \leq C \quad \text{(2)}$$

$$PRER_i \leq \epsilon_0 \quad \text{for} \quad i = 1, ..., N \quad \text{(3)}$$

where $h_i$ is the link gain and $I_i$ is the other-cell interference and thermal noise for user $i$. $W$ is the spreading bandwidth. Note that we do not assume perfect orthogonality on the forward link which eliminates the interference from the users in the same sector.

Since the up/down power control loop can achieve the SIR targets, $X_i$ is also related to the SIR target assignments as follows:

$$X_i \approx \sum_{m=1}^{M} \gamma_{i,m} q_i(m), \quad \text{(5)}$$

where $q_i(m)$ is the probability that a frame from user $i$ is being retransmitted for the $m$th time. $q_i(m)$ can be found as:

$$q_i(m) = \frac{\prod_{j=1}^{m-1} f_i(\gamma_{i,j})}{1 + \sum_{j=1}^{m-1} \frac{\prod_{k=1}^{j} f_i(\gamma_{i,k})}{\prod_{k=1}^{j} f_i(\gamma_{i,k})}}. \quad \text{(6)}$$

For a SIR and rate assignment, the power vector $P$ can be found from equation (4) to equation (6). The Post-RLP error rate for user $i$ is given as:

$$PRER_i = \prod_{m=1}^{M} f_i(\gamma_{i,m}). \quad \text{(7)}$$

In addition, the SIR and rate assignment should be within some limits. $\gamma_{\min} \leq \gamma_{i,m} \leq \gamma_{\max}$ for $i = 1, ..., N$, and $m = 1, ..., M$. $R_{\min} \leq R_i \leq R_{\max}$ for $i = 1, ..., N$.

4 Hierarchical SIR and Rate Control

The problem (1) - (3) is an constrained optimization problem, which can be solved using standard search algorithms [11]. However, a direct search approach is not practical since it requires tremendous computational complexity to search through a very high dimensional space. Further, it also lacks a distributed structure that can be fit in the mobile station-base station architecture in CDMA systems. An effective approach to solve the large-scale optimization problem is the hierarchical control method which often provides simple distributed solutions. In this paper, we will apply classical hierarchical control techniques to derive the SIR and rate control algorithm. First, we introduce two theorems to simplify the optimization problem (1) - (7) into two sub-problems.
Theorem 1 Let \( S_i^* = (\gamma_i^*, R_i^*) \) for \( i = 1, \ldots, N \) be the solution of (1) - (7). Then at least one of the following statements is true:

1. \( R_i^* = R_{\text{max}} \) for all \( i = 1, \ldots, N \), or,

2. the total power constraint in equation (2) is met with equality, i.e., \( \sum_{i=1}^{N} P_i = C \), and the PRER constraints in equation (3) are met with equality, i.e., PRER\(_i\) = \( \epsilon_0 \) for \( i = 1, \ldots, N \).

Theorem 1 implies that the optimal SIR and rate vector \( R_i^* \) should result in the PRER being equal to \( \epsilon_0 \) and a total power of \( C \) except for the case \( R_i^* = R_{\text{max}} \) for all \( i \). Intuitively, it means that if there is room in system power or PRER constraints, we should increase the rate \( R_i \) until \( R_{\text{max}} \) or the constraints are met with equality. Since the case \( R_i^* = R_{\text{max}} \) for all \( i = 1, \ldots, N \) is trivial, we will focus on the case where all users can not be assigned with the maximum data rate.

Theorem 2 Let \( \gamma_i^* = (\gamma_{i1}, \ldots, \gamma_{im}) \) be the vector that minimizes the required \( X_i \) for user \( i \) (see equation (5)), subject to the constraint that \( \text{PRER}_i = \epsilon_0 \). Suppose that the system is not maximally feasible, i.e., there exists some user \( j \) such that \( R_j^* < R_{\text{max}} \). Then \( \{\gamma_{i1}^*, R_i^*\} \) for \( i = 1, \ldots, N \) is the solution of the optimization problem (1) - (7) for some \( R_i \in [R_{\text{min}}, R_{\text{max}}] \).

Problem 1 (Mobile Algorithm): For each user \( i \), choose \( \gamma_i = (\gamma_{i1}, \ldots, \gamma_{im}) \), such that

\[
\min_{\gamma_i} \quad X_i(\gamma_i)
\]

subject to

\[
\text{PRER}_i = \epsilon_0
\]

\[
\gamma_{\text{min}} \leq \gamma_{im} \leq \gamma_{\text{max}} \quad \text{for} \quad m = 1, \ldots, M
\]

The second sub-problem can be derived as follows. According to Theorem 1, except for under-loaded systems where the power is not limiting, the total power should be set at the limit \( C \), i.e., \( \sum_{i=1}^{N} P_i = C \). Then, equation (4) can be rewritten as:

\[
X_i = \frac{W}{R_i} \cdot \frac{P_i}{C + I_i/h_i}
\]

which results in

\[
P_i = \frac{X_i R_i}{W} (C + I_i/h_i)
\]

Summing up equation (12) for all users, we can translate the power constraint into a constraint on SIR and data rate:

\[
C = \sum_{i=1}^{N} P_i = \sum_{i=1}^{N} \frac{C + I_i/h_i}{W} X_i R_i := \sum_{i=1}^{N} \alpha_i X_i R_i,
\]

where \( \alpha_i = \frac{C + I_i/h_i}{W} \). Now we can state the second sub-problem as:

Problem 2 (Base Station Algorithm): Given the optimum \( X^* = \{X_1^*, \ldots, X_N^*\} \) from Problem 1, choose \( B = \{R_1, \ldots, R_N\} \), such that

\[
\max_{B} \quad \sum_{i=1}^{N} U(R_i)
\]

subject to

\[
\sum_{i=1}^{N} \alpha_i X_i^* R_i = C
\]

\[
R_{\text{min}} \leq R_i \leq R_{\text{max}} \quad \text{for} \quad i = 1, \ldots, N
\]

4.1 Mobile Algorithm

In this section, we present an efficient 2-level iteration algorithm to solve Problem 1 at the mobile stations. This algorithm is identical to the one presented in [5]. A diagram of algorithm is shown in Figure 3.

Level 2 Algorithm: We update \( \lambda(k+1) \) iteratively given \( \gamma(k) \) using a gradient search algorithm. We formulate the error at the iteration \( k \) as:

\[
e(k) = \phi'(\lambda(k)) = \epsilon_0 - \prod_{m=1}^{M} f_i(\gamma_{im}(k))
\]

and then update \( \lambda(k+1) \) as:

\[
\lambda(k+1) = \lambda(k) + ae(k)
\]
where $a$ is the step size.

**Level 1 Algorithm:** We update $\gamma_i(k)$ based on $\lambda(k)$ obtained from Level 2. The values of $\gamma_i^*(k)$ is found by optimizing $L(\gamma_i, \lambda(k))$ and can be given in closed-form as follows. Let

$$\gamma_i^* = \begin{cases} 
\gamma_{\min} & \text{if } Y_i^M < f'(\gamma_{\min}) \\
\gamma_{\max} & \text{if } Y_i^M > f'(\gamma_{\max}) \\
d_{\gamma_i}^{-1}(-Y_i^M) & \text{if otherwise},
\end{cases}$$

(18)

where

$$Y_i^M = \frac{1}{\lambda(k)}$$

(19)

Note that (18) gives the optimal SIR value only for the $M$th retransmission. The other SIR values $m = 1, ..., M-1$, $\gamma_{i,m}^*(k)$ can be obtained as:

$$\gamma_{i,m}^* = \begin{cases} 
\gamma_{\min} & \text{if } -\frac{1}{\lambda(k)} < f'(\gamma_{\min}) \\
\gamma_{\max} & \text{if } -\frac{1}{\lambda(k)} > f'(\gamma_{\max}) \\
d_{\gamma_i}^{-1}(-\frac{1}{\lambda(k)}) & \text{if otherwise},
\end{cases}$$

(20)

where

$$Y_i^m = \frac{1}{\lambda(k)} \prod_{j=m+1}^{M} f_i(\gamma_{i,j}^*(k))$$

(21)

Similar to [5], we can show that the mobile algorithm (17) - (21) converges to a local optimum solution of the optimization problem (8).

### 4.2 Base Station Algorithm

After the 2-level algorithm converges, the mobile calculates the minimum required loading $X_i^*$ from equation (5) using the optimal SIR set $\{\gamma_{i,1}^*, ..., \gamma_{i,M}^*\}$ and sends it to the base station. The base station uses $\{X_1^*, ..., X_N^*\}$ from all the data users to choose the values of $R = \{R_1^*, ..., R_N^*\}$ to optimize Problem 2. Again, an hierarchical algorithm can efficiently calculate the solution and has guaranteed convergence.

### 5 Numerical Results

In this section, we present some numerical results. The FER function $f(\cdot)$ is approximated as:

$$f_i(\gamma) = 1 - e^{-\frac{\gamma}{\alpha}}$$

where $F_0$ is the parameter representing the fade margin [12]. To reflect the difference between the channels, the fade margin can be written as

$$F_i = \beta_i F_0 \gamma_i$$

where $F_0$ is a constant and $\beta_i$ is the parameter reflecting the power control efficiency [13]. For perfect power control, $\beta_i = 1$. In the simulation, we assume there are 6 users with $\beta_i = 0.2, 0.3, ..., 0.7$. $F_0 = 7.5$. The lower limit of SIR $\gamma_{\min} = 2$ and upper limit of SIR $\gamma_{\max} = 80$. The power limit $C$ is set at 8 Watts. The link gains are $[h/32 h/16 h/8 h/4 h/3 h]$ and the interference levels are $[I/2 I/4 I/8 I/16 I/32]$, where $h = 0.01$ and $I = 0.045$. Note that user 6 has the best channel and user 1 has the worse channel. The spreading bandwidth is 1M bits/sec. The minimum rate for a data user $R_{\min}$ is 20K bits/sec and the maximum $R_{\max}$ is 200K bits/sec. The RLP will retransmit an errored frame at most 3 times before passing it to the higher layer, i.e., $M = 3$. To ensure proper operation of the higher layer protocols, the Post-RLP error rate is kept less than $5 \times 10^{-6}$. The utility function we suggest is as follows:

$$U(R_i) = \left( \frac{R_i}{R_{\max}} \right)^{\rho}$$

(22)

The convexity can be guaranteed by choosing $\rho \leq 0$. The

![Utility function U( ) for different values of ρ](image)

517
We can see that users with better channels require lower SIR targets to meet the PRER constraints. The resulting rate assignments from the base station using different values of $\rho$ is given in Table 2, where we can see that a lower value of $\rho$ results in a fairer system but with lower total system performance. A flexible tradeoff between the total data rate and fairness can be achieved by adjusting the value of $\rho$.

Table 1: SIR assignments by the mobile

<table>
<thead>
<tr>
<th>user</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_1$</td>
<td>9.6</td>
<td>5.5</td>
<td>3.7</td>
<td>2.8</td>
<td>2.2</td>
<td>2.0</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>73.8</td>
<td>37.6</td>
<td>23.2</td>
<td>16.0</td>
<td>11.8</td>
<td>8.1</td>
</tr>
<tr>
<td>$\gamma_3$</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 2: Rate assignment by the base station

<table>
<thead>
<tr>
<th>$\rho$</th>
<th>$R_1^*$</th>
<th>$R_2^*$</th>
<th>$R_3^*$</th>
<th>$R_4^*$</th>
<th>$R_5^*$</th>
<th>$R_6^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20.0</td>
<td>25.8</td>
<td>200</td>
<td>200</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>-1</td>
<td>25.4</td>
<td>48.3</td>
<td>87.9</td>
<td>175.6</td>
<td>200</td>
<td>200</td>
</tr>
<tr>
<td>-2</td>
<td>31.1</td>
<td>47.6</td>
<td>68.4</td>
<td>96.3</td>
<td>135.2</td>
<td>195.4</td>
</tr>
<tr>
<td>-3</td>
<td>34.1</td>
<td>46.5</td>
<td>60.9</td>
<td>78.1</td>
<td>98.8</td>
<td>123.5</td>
</tr>
<tr>
<td>-5</td>
<td>36.4</td>
<td>44.8</td>
<td>53.0</td>
<td>63.2</td>
<td>73.8</td>
<td>85.3</td>
</tr>
<tr>
<td>-10</td>
<td>38.4</td>
<td>42.9</td>
<td>47.4</td>
<td>51.8</td>
<td>56.4</td>
<td>61.0</td>
</tr>
<tr>
<td>-15</td>
<td>39.0</td>
<td>42.2</td>
<td>45.1</td>
<td>48.0</td>
<td>50.9</td>
<td>53.7</td>
</tr>
</tbody>
</table>

6 Conclusion

In this paper, we proposed a SIR and rate control algorithm for CDMA data users under constraints on delay and total power. The QoS for data users is mapped onto a family of utility functions that reflect the tradeoff between total system throughput and fairness (similarity in data rates of users). The proposed algorithm uses an hierarchical control structure that is simple to implement in a distributed manner. Specifically, the data rates for the users are adjusted by the base station while the SIR targets are adjusted by the mobiles. The mobiles make SIR adjustments only based on information specific to each user. The base station makes the rate assignments jointly based on the feedback from the mobiles. The algorithm requires very limited information exchange between the base station and mobiles. The distributed structure makes it a very attractive candidate for resource control in high speed CDMA data systems.

References


