Utility Optimization and Fairness Guarantees for Multimedia Traffic in the Downlink of DS-CDMA Systems

Xiang Duan, Zhisheng Niu, and Junli Zheng
States Key Labs on Microwave and Digital Communications,
Department of Electronic Engineering,
Tsinghua University, Beijing 100084, CHINA

Abstract—Radio resource control is one of the key technologies for providing quality-of-service (QoS) in wireless communication systems, where power and data rate are schedulable resources. In this paper, we propose a utility-based resource control scheme for multimedia traffic in the downlink of DS-CDMA systems. The goal is to achieve the social optimal resource utilization. Fairness is guaranteed by allocating each user with a minimal transmission data rate. Simulation results show that the proposed resource control scheme is flexible and efficient for the downlink of multimedia DS-CDMA systems.

I. INTRODUCTION

Radio resource control is one of the key technologies for providing quality-of-service (QoS) guarantees for multimedia traffic in the emerging wireless communication systems. Although most previous studies have interpreted user’s QoS requirements into specific technical measurements such as equivalent bandwidth, delay and loss [1] [2], QoS itself is a perception which represents the satisfaction level that a user or an upper layer application obtains when using the service provided by the network. Fairness is guaranteed by allocating each user with a minimal transmission data rate. With the concept of utility, QoS requirements of multimedia traffic are interpreted as “soft” objects, which are tunable according to system load and channel conditions.

In direct-sequence code-division multiple access (DS-CDMA) systems where power and data rate are schedulable resources, power control and data rate allocation are two important components for resource control. In [3] and [4], authors proposed distributed utility-based power control schemes for uplink and downlink in CDMA systems, respectively. Ref. [5] proposed a distributed power control scheme based on a N-person non-cooperative game model, and [6] improved the scheme by introducing cost to utility functions.

The references mentioned above considered only the problem of power control. In DS-CDMA systems, however, transmission data rates is also schedulable via the variation of spreading gain, which is not considered in [3] - [6]. In [7], the author presented a utility-based resource control scheme for CDMA networks carrying elastic traffic, which concerns enhancing the resource utilization by the incentive of congestion price. However, the proposal was not feasible for achieving the social optimal resource utilization in the downlink of CDMA systems. We have proposed [8] a utility-based joint power and rate allocation algorithm aiming at maximizing the system overall utilities (defined as the sum of utilities of all users in the system). Fairness is guaranteed by allocating each user with a minimal transmission data rate.

In this paper we extended our work in [8] and propose a utility-based resource control scheme for multimedia traffic in the downlink of DS-CDMA systems. The goal of the algorithm is to obtain the social optimal resource utilization which maximizing the system overall utilities (defined as the sum of utilities of all users in the system). Fairness is guaranteed by allocating each user with a minimal transmission data rate.

The rest of the paper is organized as follows. In Section II, we describe the utility functions and the system model of the utility-based resource control problem. Then we look for the social optimal resource allocation by finding the optimum solution in Section III. Also in this section, a thorough description of the resource control algorithm is given. Section IV presents the numerical results. Finally we conclude the paper in Section V.

II. SYSTEM MODEL

Consider the downlink of a typical cell in a multimedia DS-CDMA systems with N users. Denoted by Ri the transmission rate, by Pi the transmit power, and by hi the channel gain of from base station to user i, and let $R = \{R_1, R_2, \cdots, R_N\}$ and $P = \{P_1, P_2, \cdots, P_N\}$. Let $W$ represent the system bandwidth, and $I_i$ the background interference at the location of user i. Then, the received signal-to-interference ratio (SIR) of user i is given by [7]:

$$\gamma_i = \frac{W}{R_i} \frac{h_i P_i}{\theta d h_i (\sum_{j \neq i} P_j) + I_i}$$

where $\theta_d$ is the downlink orthogonality factor with the typical values fall in the range of [0.1, 0.6] [9].

We focus on elastic traffic, where a user evaluates the received QoS by its data throughput [10], which is the product
of transmission rate $R_i$ and the transmission efficiency. Transmission efficiency, given by the efficiency function $E_s(\gamma)$, describes the percentage of successfully transmitted information bits to overall bits transmitted [8]. It is a function of received SIR, and is determined by physical layer technologies. Users are referred to [8] for more detailed information about efficiency functions.

As a result, the utility function of user $i$ is given by $U_i(R, E_s(\gamma_i))$. The form of $U_i(\cdot)$ should be carefully selected in order to properly reflect the nature of satisfaction level. Some literature [3] [4] has pointed out that $U_i(\cdot)$ should be a non-decreasing function, satisfying $U_i(0) = 0$ and $U_i(\infty) = U_i^{\max} < +\infty$. For data users who have no stringent QoS requirements on transmission rate, we assume an increasing concave function. For data and multimedia applications that have minimum rate requirements, Sigmoidal-like functions are used, as shown in Fig. 1.

Assume that $P_{\max}$ is given constraint on the transmit power of base station. Mathematically, our resource control problem is formulated as an optimization model aiming at maximization of the system overall utilities:

$$\max_{\mathbf{P}, \mathbf{R}} \sum_{i=1}^{N} U_i(R_i, E_s(\gamma_i))$$

s.t. $$\begin{align*}
\gamma_i &= \frac{W h_i P_i}{\theta_i h_i (\sum_{j \neq i} P_j) + I_i}, \\
\sum_{i=1}^{N} P_i &\leq P_{\max}
\end{align*}$$

\tag{1}

The optimum solution to the problem is called the social optimal resource utilization of the utility-based resource control problem.

Eq. (1) is a nonlinear optimization problem with $2N$ decision variables, $\mathbf{R}$ and $\mathbf{P}$. The following proposition simplifies the problem.

**Proposition 1:** Assume $\mathbf{P}$ and $\mathbf{R}$ are the optimum solution to eq. (1), then $\mathbf{P}$ and $\mathbf{R}$ must satisfy

$$\gamma_i = \gamma^*, \quad (i = 1, 2, \ldots, N)$$

where $\gamma^*$ is the optimal received SIR such that

$$\gamma^* = \arg \max_{\gamma} E_s(\gamma)$$

**Proof:** We can prove this proposition by contradiction. Suppose the assumption in the proposition is not true, i.e., there exists $\mathbf{P}$ and $\mathbf{R}$ that are the optimum solution to eq. (1), and for some $i$,

$$\gamma_i = \frac{W h_i P_i}{R_i - I_i} \neq \gamma^*_i$$

where $I_i = \theta_i h_i (\sum_{j \neq i} P_j) + I_i$ is the interference received by user $i$. It is clear that $I_i$ keeps unchanged when the value of $R_i$ changes. Therefore, we let $\mathbf{P'} = \mathbf{P}$, $\mathbf{R'} = \{R_1, R_2, \ldots, R'_i, \ldots, R_N\}$, where

$$R'_i = \frac{W h_i P_{i}}{\gamma^* - \frac{I_i}{R_i}}$$

Therefore, $\gamma'_i = \gamma^*_i$. Since $\gamma^* = \arg \max_{\gamma} E_s(\gamma)$, we have

$$R'_i E_s(\gamma'_i) = \frac{W h_i P_i E_s(\gamma^*_i)}{I_i - R_i E_s(\gamma^*_i)} > \frac{W h_i P_i E_s(\gamma_i)}{I_i - R_i E_s(\gamma_i)} = R_i E_s(\gamma_i)$$

while $R_i E_s(\gamma_i)$ keeps unchanged for $j \neq i$. As a result, $\sum_{i=1}^{N} U_i(R'_i E_s(\gamma'_i)) > \sum_{i=1}^{N} U_i(R_i E_s(\gamma_i))$. This contradicts with the fact that $\mathbf{P}$ and $\mathbf{R}$ are the optimum solution to system model (1). So the assumption in the above proposition must be true.

If $E_s(\gamma)$ is unimodal, the value of $\gamma^*$ is the solution of

$$E_s(\gamma^*_i) = \gamma^*_i E_s(\gamma^*_i)$$

The value of $\gamma^*_i$ can be obtained by simulation or measurements in real systems.

With $\gamma^*$, $\mathbf{P}$ and $\mathbf{R}$ are not independent, and the system model (1) is equivalent to

$$\max_{\mathbf{P}, \mathbf{R}} \sum_{i=1}^{N} U_i(R_i, E_s(\gamma_i))$$

s.t. $$\begin{align*}
\frac{W h_i P_i}{\theta_i h_i (\sum_{j \neq i} P_j) + I_i} &= \gamma^*, \\
\sum_{i=1}^{N} P_i &\leq P_{\max}
\end{align*}$$

\tag{2}

As a result from [11], the system model can be simplified to

$$\max_{\mathbf{R}} \sum_{i=1}^{N} U_i(R_i, E_s^*)$$

s.t. $$\sum_{i=1}^{N} \delta_i g(R_i) \leq 1$$

where $E_s^* = E_s(\gamma^*)$ is the corresponding transmission efficiency of $\gamma^*$, and

$$\delta_i = 1 + \frac{P_i}{P_{\max}}$$
The downlink equivalent background interference \( I_i^b \) is the downlink equivalent background interference \( I_i^b = \frac{1}{\gamma_{R_i} \gamma_{o_i}} \) \[11\]. The power index \[2\] \( g(R_i) \) is a function of \( R_i \), given by

\[
g(R_i) = \frac{1}{\sum\limits_{i=1}^N \gamma_{R_i} \gamma_{o_i} + 1}, \quad (i = 1, 2, \ldots, N)
\]

The equivalent system model \( \text{(2)} \) has \( N \) independent decision variables, \( \mathbf{R} \). Generally finding optimum solution by normal algorithms, such as steepest decent method or gradient projection method \[12\], is a non-trivial task. In the next section we focus solving this problem with lower computational complexity.

### III. FINDING THE SOCIAL OPTIMAL RESOURCE UTILIZATION

#### A. Analysis of System Model

Define the Lagrangian of the equivalent system model \( \text{(2)} \):

\[
\mathcal{L}(\mathbf{R}, \lambda) = \sum\limits_{i=1}^N U_i(R_i E^*_i) - \lambda \sum\limits_{i=1}^N \delta_i g(R_i) - 1
\]

\[
= \sum\limits_{i=1}^N \{U_i(R_i) - \lambda \delta_i g(R_i)\} + \lambda
\]

We have the following lemma according to \[12\].

**Lemma 1:** Let \( f : \mathbb{R}^N \mapsto \mathbb{R} \) and \( h : \mathbb{R}^N \mapsto \mathbb{R} \) be arbitrary functions. Let

\[
\mathcal{L}(\mathbf{X}, \lambda) = f(\mathbf{X}) - \lambda h(\mathbf{X})
\]

and

\[
\hat{\mathbf{X}}(\lambda) = \arg \max_{\mathbf{X}} \mathcal{L}(\mathbf{X}, \lambda)
\]

where \( \mathbf{X} = \{x_1, x_2, \ldots, x_N\} \). Then we have \( \hat{\mathbf{X}}(\lambda) \) is a global optimum solution of the following optimization problem:

\[
\max f(\mathbf{X})
\]

s.t. \( h(\mathbf{X}) \leq h(\hat{\mathbf{X}}(\lambda)) \)

Hence, the equivalent system model is decomposed into the following two subproblems:

1) **Lagrangian maximization subproblem:**
   This subproblem is to find such \( \mathbf{R}^* \) that \( \mathcal{L}(\mathbf{R}^*, \lambda) = \max_{\mathbf{R}} \mathcal{L}(\mathbf{R}, \lambda) \);

2) **Price matching subproblem:**
   Here \( \lambda \) is called the resource price. The task of this subproblem is to find such \( \lambda^* \) that \( \sum_i \delta_i g(R_i) = 1 \).

Lemma 1 implies that if we find such a pair of \( \lambda^* \) and \( \mathbf{R}^* \) that solves the two subproblems at the same time, \( \mathbf{R}^* \) is the optimum solution to \( \text{(2)} \). If \( \sum_i \delta_i g(R_i) \approx 1 \), \( \mathbf{R}^* \) is a good approximation to the optimum solution to \( \text{(2)} \) satisfying the feasibility condition.

Eq. \( \text{(3)} \) illustrates that the Lagrangian maximization subproblem is further decomposed into \( N \) independent sub-subproblems given by

\[
\max\limits_{R_i} \{U_i(R_i E^*_i) - \lambda \delta_i g(R_i)\}, \quad (i = 1, 2, \ldots, N)
\]
B. Algorithm Description

Assume the following system parameters are given or calculated in advance when the system is set up or a user is admitted to access the network: the base station power constraint \( P_{\text{max}} \), optimal SIR \( \gamma^* \) and corresponding efficiency \( E_i^* \), the forms of utility functions \( U_i(\cdot) \) and inverse characteristic functions \( f_{\lambda_i}(\cdot) \) of each user, and the corresponding \( R_{\text{min}} \) and \( \lambda_{\text{max}} \). According to discussions above, the algorithm of the resource control module works periodically at each resource control work point (may be the beginning of each time frame) can be described as the following steps.

1) Channel prediction or measuring: for any user \( i \) (\( i = 1,2,\ldots,N \)), predict or measure current channel gain \( h_i \) and background interference \( I_i \), then calculate \( I_i^0 \) and \( \delta_i \). Since small-scaled fast channel fading are partly neutralized by fast power control, \( h_i \) is assumed to be the large scale path loss and is kept constant in the period.

2) Adaption setup: Let \( m \) represent the iteration step. Set \( m = 1 \), initialize \( \lambda \) and \( \bar{\lambda} \) to the maximal and minimal possible price, respectively, and set \( \lambda^{(1)} = 0.382(\lambda + \bar{\lambda}) \).

3) Rate adaption: Calculate \( R_i^{(m)} \) by
   
   \[
   R_i^{(m)} = \begin{cases} 
   f_{\lambda_i}^{-1}(\delta_i \lambda_i(m)), & \text{if } \delta_i \lambda_i(m) \leq \lambda_{\text{max}} \\
   R_{\text{min}}, & \text{if } \delta_i \lambda_i(m) > \lambda_{\text{max}} 
   \end{cases}
   \]
   and calculate \( g(R_i^{(m)}) \) (\( i = 1,2,\ldots,N \)).

4) Price adaption: Adjust \( \lambda \), \( \bar{\lambda} \) and \( \lambda \) by the golden section search: if \( \sum \delta_i g(R_i^{(m)}) < 1 \), \( \lambda = \lambda^{(m)} \), and \( \lambda^{(m+1)} = 0.382(\lambda + \bar{\lambda}) \); otherwise \( \bar{\lambda} = \lambda^{(m)} \), and set \( \lambda^{(m+1)} = 0.618(\lambda + \bar{\lambda}) \).

5) Terminate condition: If \( m = M \) or \( \bar{\lambda} - \lambda \leq \varepsilon \), stop; otherwise \( m = m + 1 \) and goto step 3. Here \( \varepsilon \) is the required precision [12], and \( M \) is the maximal iteration times. According to our simulation, \( M=6 \) is good enough for the algorithm to converge.

6) Normalization of resource allocation: Calculate the normalized rate \( R_i \) by
   
   \[
   R_i = g^{-1}\left(\frac{g(R_i^{(m)})}{\sum_j \delta_j g(R_j^{(m)})}\right)
   \]
   where \( g^{-1}(\cdot) \) is the inverse function of \( g(\cdot) \). By normalization, \( \sum_j \delta_j g(R_j) = 1 \) can be satisfied, which means all radio resources are utilized.

7) Transmit power: Transmit power assigned to user \( i \) is given by
   
   \[
   P_i = \frac{g_i \sum_j \delta_j R_j}{1 - \sum_j g_j} + g_i \bar{I}_i^0
   \]
   where for convenience we use \( g_i \) to represent \( g(R_i) \).

IV. SIMULATION RESULTS

In this section, we present simulation results. We simulate a typical cell in DS-CDMA systems with 40 mobile users. Each user is equipped with one type of traffic among the four considered types: voice, data and two types of multimedia traffic. Parameters of these traffic are given in Table I. Their utility functions and characteristic functions are shown in Figs. 1 and 2, respectively. Other system parameters are assumed to be: \( W = 3.84 \text{MHz}, P_{\text{max}} = 24 \text{W}, I_i \sim (\mu, \sigma^2) \text{W} \) with \( \mu = 3e - 6 \) and \( \sigma^2 = 2.5e - 13 \) and \( \theta_d = 0.2 \). QPSK modulation, (511,175,46) BCH coding (coding rate 1/3) and simple ARQ scheme are assumed as the physical layer technologies (as a result, \( \gamma^* = 1.55 \) and \( E_i^* = 0.34 \)).

We simulate 50 seconds in real time. The time interval between two consecutive resource control work point is 50 ms. In order to analyze the performance, we also consider two other proposals given in [7] and compared them with our algorithm. One proposal is a resource control algorithm of congestion pricing for downlink (CPD), which is formulated as an optimization problem:

\[
\max_{R_i} U_i(R_i, E_i(\gamma_i)) - \lambda P_i
\]

where \( w_i \) is the user \( i \)'s willingness-to-pay [7]. Here we use the value of \( R_{\text{min}} \) to represent it.

Fig. 3 shows the comparison of the overall achieved utilities of CPD, WTP and our algorithm in 5 simulation scenarios, and Fig. 4 shows the comparison of achieved utilities of typical users in a typical simulation scenario. In Fig. 4, the value of \( U_i(R_{\text{min}}) \), which is the corresponding utility of \( R_{\text{min}} \), is drawn. From the two figures we clearly observe that:

1) Our proposal achieves more overall utilities than CPD and WTP, as shown in Fig. 3. Especially compared with CPD, the utility gain is generally up to 40%-70%, which is significant. Though WTP performs better than CPD, it still achieves less overall utilities than our proposal.

2) Fairness is well guaranteed by our proposed algorithm, and the condition is similar for CPD, as shown in Fig. 4. For our proposal, all users’ average achieved utilities exceed \( U_i(R_{\text{min}}) \). However, for WTP, data users suffer “service starvation” in that they achieve little utility though other users achieve rather high utilities. Hence, it is totally unfair for data users.

As a conclusion, we see that our proposal has advantages over CPD and WTP in higher overall utilities and better fairness guarantees, therefore it is efficient for resource control in the downlink of DS-CDMA systems.

Next we have a close look at the simulation trace of the four different types of traffic in the system. The system load is designed to be carried by extensive amounts of traffic. Some of the traffic is generated by application-level connections, while some is generated by the system-level control mechanisms. The system-level control mechanisms include the adaptive modulation and coding (AMC) process, the power control process, and the dynamic resource allocation process. The AMC process is responsible for adjusting the modulation and coding schemes according to the channel conditions. The power control process is responsible for adjusting the transmit power of each user according to the channel conditions and the traffic demands. The dynamic resource allocation process is responsible for adjusting the resource allocation according to the traffic demands and the channel conditions.

The simulation trace shows that our proposal is able to achieve better overall utilities and better fairness guarantees than CPD and WTP. This is because our proposal is able to adapt to the changing channel conditions and traffic demands more efficiently than CPD and WTP. The AMC process is able to adjust the modulation and coding schemes according to the changing channel conditions. The power control process is able to adjust the transmit power of each user according to the changing channel conditions and the traffic demands. The dynamic resource allocation process is able to adjust the resource allocation according to the changing traffic demands and the channel conditions.

The simulation trace also shows that our proposal is able to achieve better overall utilities and better fairness guarantees than CPD and WTP. This is because our proposal is able to adapt to the changing channel conditions and traffic demands more efficiently than CPD and WTP. The AMC process is able to adjust the modulation and coding schemes according to the changing channel conditions. The power control process is able to adjust the transmit power of each user according to the changing channel conditions and the traffic demands. The dynamic resource allocation process is able to adjust the resource allocation according to the changing traffic demands and the channel conditions.

As a result, our proposal is able to achieve better overall utilities and better fairness guarantees than CPD and WTP. This is because our proposal is able to adapt to the changing channel conditions and traffic demands more efficiently than CPD and WTP. The AMC process is able to adjust the modulation and coding schemes according to the changing channel conditions. The power control process is able to adjust the transmit power of each user according to the changing channel conditions and the traffic demands. The dynamic resource allocation process is able to adjust the resource allocation according to the changing traffic demands and the channel conditions.
then-decreasing” so that the variation of utilities could be clearly shown. From Fig. 5 we see that the for voice traffic, the achieved utility keeps almost unchanged at different system load; though the achieved utilities of the two types of multimedia traffic are influenced by the variation of system load, they also keep at a relatively high level. On the other hand, the variation of system load has strong influence on the achieved utility of data traffic: when system is heavily loaded, data traffic withdraws from the contention of resources automatically; it is assigned the the lowest data rate and the resources are saved for traffic with higher priority, i.e., data and multimedia traffic.

V. CONCLUSION

In this paper, we have proposed a dynamic utility-based resource control scheme for the downlink of multimedia DS-CDMA systems. The goal is to achieve the social optimal resource utilization, which is the maximization of system overall utilities while the transmit power constraint of base station is satisfied. Simulation results have shown its efficiency and flexibility in that it can allocate radio resources dynamically according to user’s QoS requirements, channel conditions, and current system load.

REFERENCES


