Optimized Resource Allocation and Scheduling in Downlink for Multimedia CDMA Wireless Systems

H. Cahit Akin
School of Electrical Engineering
University of California at San Diego
La Jolla, California 92093
Email: hakin@ucsd.edu

Kimberly M. Wasserman
Cisco Systems
Research Triangle Park
North Carolina 27709
Email: kwasserm@cisco.com

I. Extended Abstract

A. Introduction

CDMA based technologies are becoming a standard for Wireless Wide Area Networks. With the introduction of higher peak rate availabilities, various real-time applications with latency and high bandwidth requirements are expected to become widespread besides voice applications. Therefore it is crucial to have spectrum and power efficient wireless technologies.

In this paper, we are introducing a power allocation and scheduling algorithm. As a consequence of the way the problem is set up, the proposed algorithm always outperforms pure CDMA and pure TDMA based schemes.

B. System Model and Analysis

We will focus on the downlink (forward link) in a CDMA based cellular Wireless Wide Area Network (WWAN). We will assume that every wireless station is assigned to a single base station and stays assigned to that base station throughout its connection lifetime. Every wireless station has some latency sensitive data to transmit with different QoS requirements during its connection lifetime. We will further assume without loss of generality that every wireless station has a single connection and the QoS requirement of that connection is not changing over time.

Let M be the number of wireless stations in a particular cell site of interest within a larger network. Let **p(t)** $= (p_1(t), p_2(t), \ldots, p_M(t))$ be the powers of the transmitted signals to the wireless stations from the basestation with time dependencies. Also define **g(t)** $= (g_1(t), g_2(t), \ldots, g_M(t))$ as the downlink channel gains vector from the base station to the wireless stations.

Define **P** $= (P_1, P_2, \ldots, P_M)$ as the maximum allowed transmission powers vector, where P_i is the maximum allowed power from the basestation to wireless station i. Notice that such a power restriction non-explicitly exists for real-time connections with latency requirements. As an example consider a 9.6Kbps realtime connection with 100msec latency tolerance (i.e. voice connection). Then, every 100msecs, there is only 960 bits to be sent. If the smallest granularity of packets is 10 msecs then the maximum needed peak rate is limited by 96Kbps which in tern puts a restriction on the peak power level needed.

Let I_i denote the inter-cell interference (interference caused by the neighbour basestations in the adjacent cells) plus the background noise experienced by the wireless station i. In general the value of I_i will be time dependent, but for the time being for short time durations of our interest, we will assume that I_i is a constant.

There is a unique mapping from the BER requirement of a downlink connection to the required E_b/N_0 value at the wireless station, where E_b is the energy per bit and N_0 is the total noise experienced at the wireless station for that connection. This mapping depends on factors such as the modulation scheme, interleaving method and error-correction scheme. Therefore we will assume the wireless stations define their QoS requirements in terms of their E_b/N_0 needs. Let $\kappa = (\kappa_1, \kappa_2, \ldots, \kappa_M)$ be the QoS requirements vector.

Then the SNR equation for the wireless station i is:

$$(\frac{E_b}{N_0})_i = \frac{W}{R_i} \frac{g_i p_i}{\sum_{j \neq i} g_j p_j + I_i} = \frac{W}{R_i} \frac{p_i}{\sum_{j \neq i} p_j + \frac{I_i}{g_i}}, \quad \forall i \in \{1, 2, \ldots, M\} \tag{1}$$

where R_i is the throughput of the i^{th} wireless station for a unit time duration ($rate$) and W is the bandwidth of the downlink. The last equation follows because $g_i = g_j$ for all $j \in \{1, 2, \ldots, M\}$ in a downlink from the base station to the wireless station.

Notice that the right-hand side of the equation is the processing gain multiplied by the user's received power divided by the total noise power that user is experiencing. $(E_b/N_0)_i$ and R_i are inversely proportional, therefore the QoS requirements should be met with equality, $(E_b/N_0)_i =$

κ_i, for throughput maximization; since we can always lower the E_b/N_0, increase R_i and keep every other value constant in equation (1) as long as the E_b/N_0 requirement is satisfied. Therefore the throughput of the i^{th} wireless station in the $[0,t]$ time interval is given by

$$h_i(0,t) = \int_0^t R_i(t)dt = \frac{W}{\kappa_i}\int_0^t \frac{p_i(t)}{\sum_{j\neq i} p_j(t) + I_i/g_i(t)}dt \tag{2}$$

Let $\rho = (\rho_1, \rho_2, \ldots, \rho_M)$ be the minimum required rates vector. Define the sets $\Phi = \{\mathbf{R} \mid \frac{\int_0^t R_i(t)dt}{t} \geq \rho_i, \forall i = 1, 2, \ldots, M\}$ and $\Psi = \{\mathbf{p} \mid \mathbf{0} \leq \mathbf{p} \leq \mathbf{P}\}$ as the required rates and feasible powers vector sets respectively[1]. Assume that the connection topology stays the same during $[0,t]$ time duration and further assume that $[0,t]$ is short enough so that the channel gains are constant. If we denote $H(0,t) = \sum_i h_i(0,t)$ as the total throughput then the total downlink throughput maximization problem of the cell is given by:

$$\sup_{\mathbf{R}\in\Phi,\ \mathbf{p}\in\Psi} H(0,t) = \sup_{\mathbf{R}\in\Phi,\ \mathbf{p}\in\Psi} \sum_i \frac{W}{\kappa_i}\int_0^t \frac{p_i(t)}{\sum_{j\neq i} p_j(t) + I_i/g_i}dt \tag{3}$$

Clearly the set that defines the domain of the optimization (3) is infinite and there is no clear method of finding the best solution other than trial and error which is infeasible.

For a practical system, the transmitted powers are selected from a discrete (quantized) set of power levels. Therefore with a dense enough quantization, the continuous time power allocation and scheduling can be approximated with a power allocation and scheduling with fixed power levels, with arbitrary proximity. For quantization level k (i.e., a transmitted power can take one of the k discrete values), there are only k^M different power allocations possible. Let's associate the time durations $\Gamma_n = (t_n, t_{n+1})$, $n = 1, 2, \ldots, k^M$ to each such distinct power allocation and find the optimum set of Γ_n's which maximizes the total throughput of the cell site, i.e.:

$$\max_{\mathbf{R}\in\Phi,\ \mathbf{p}\in\Upsilon(B)} H(0,t) \cong \max_{\mathbf{R}\in\Phi,\ \mathbf{p}\in\Upsilon(B),\ \sum|\Gamma_n|=t} \sum_i \frac{W}{\kappa_i}\sum_{n=1}^{K} \frac{p_{in} \mid \Gamma_n \mid}{\sum_{j\neq i} p_{jn} + \frac{I_i}{g_i}} \tag{4}$$

where the set $\Upsilon(\alpha) = \{\mathbf{p} \mid \mathbf{0} \leq \mathbf{p} \leq \mathbf{P} \text{ and } \mathbf{1}\cdot\mathbf{p} = \alpha\}$ and B is the maximum total power the base station is allowed to transmit, $\mid \Gamma_n \mid = t_{n+1} - t_n$ and p_{in} is the fixed transmitted power level of wireless station i in the n^{th} subinterval Γ_n, i.e. $p_i(t) = p_{in}$ for $t \in \Gamma_n$. Γ_n's form a partition of the $[0,t]$ time interval. Notice that the above maximization problem is done over all

[1]Vector relations are componentwise.

possible sets of Γ_n's and the p_{in} values are constants. Therefore the condition $\mathbf{R} \in \Phi$ becomes the constraint on the decision variables Γ_n's.

Definition 1 *Vertex: A transmitted powers vector is a vertex in a time interval if $p_i = 0$ or $p_i = P_i$ for all $i = 1, 2, \ldots, M$ in that time interval.*

Definition 2 *Vertex-restricted-by-B: A transmitted powers vector is a vertex-restricted-by-B in a time interval if $p_i = 0$ or $p_i = P_i$ for all $i = 1, 2, \ldots, M$ except one $i = k \in \{1, 2, \ldots, M\}$ for which $0 \le p_k \le P_i$ and $\Sigma_i p_i = B$ in that time interval.*

Proposition 1 *In the solution of the optimization problem (4), the transmitted powers vector in subinterval Γ_i is either a vertex or a vertex-restricted-by-B, for all $i = 1, 2, \ldots, M$.*

Proof: Assume there exists at least one subinterval Γ_j in the optimum solution such that the transmitted power vector is not a vertex nor a vertex-restricted-by-B. This means at least two of the transmitted power values, p_{ij} is neither 0 nor P_i and p_{kj} is neither 0 nor P_k. Assume $p_{ij} > p_{kj}$ without loss of generality. Then one can divide Γ_j into two subintervals such that the new value of p_{ij} is $p^1_{ij} = p_{ij} - q > 0$ and the new value of p_{kj} is $p^1_{kj} = p_{kj} + q$ in the first λ portion of Γ_j and the new values are $p^2_{ij} = p_{ij} - q$ and $p^2_{kj} = p_{kj} + q$ in the remaining $1 - \lambda$ portion of Γ_j. Let all the other transmitted power values stay unchanged for both subintervals. Then the throughput of the downlinks other than i and k are unchanged in the new power allocation scenario. But for the i^{th} and k^{th} downlink we have an increased throughput in the new power allocation scenario if $\lambda \in (\frac{a+b_1-q}{2(a+b_1)}, \frac{c+b_2+q}{2(c+b_2)})$ where $a = p_{ij}$, $b_1 = (\Sigma_{n \ne i,k} p_{nj}) + \frac{I_i}{g_i}$, $b_2 = (\Sigma_{n \ne i,k} p_{nj}) + \frac{I_k}{g_k}$ and $c = p_{kj}$ since:

$$\lambda < \frac{c + b_1 + q}{2(c + b_1)} \Rightarrow \lambda \frac{a - q}{b_1 + c + q} + (1 - \lambda) \frac{a + q}{b_1 + c - q} > \frac{a}{c + b_1} \tag{5}$$

where the intermediate steps are straightforward therefore are skipped. Notice that multiplying each side of the last inequality by $\mid \Gamma_j \mid$ proves that the throughput of the i^{th} downlink is improved in the new scenario. Similarly

$$\lambda > \frac{a + b_2 - q}{2(a + b_2)} \Rightarrow \lambda \frac{c + q}{a + b_2 - q} + (1 - \lambda) \frac{c - q}{a + b_2 + q} > \frac{c}{a + b_2} \tag{6}$$

Again notice that multiplying each side of the last inequality by $\mid \Gamma_j \mid$ proves that the throughput of the k^{th} downlink is improved in the new scenario. Finally one can easily verify that $\frac{a+b_1-q}{2(a+b_1)} < \frac{c+b_2+q}{2(c+b_2)}$ to complete the proof.

■

By proposition 2, the optimization problem in (4) becomes:

$$\max_{\mathbf{R}\in\Phi,\ \sum|\Gamma_n|=t} \sum_i \frac{W}{\kappa_i} \sum_{n=1}^{L} \frac{\tilde{p}_{in} \mid \Gamma_n \mid}{\sum_{j\neq i} \tilde{p}_{jn} + \frac{I_i}{g_i}} \tag{7}$$

where the vectors $\mathbf{Y}_n = (\tilde{p}_{1n}, \tilde{p}_{2n}, \ldots, \tilde{p}_{Mn}) \in \varphi$, $n = 1, 2, \ldots, L$ are all distinct, and φ denotes the set of all possible vertices and vertices-restricted-by-B. Also without loss of generality we renamed the partition in which the vertex $(\tilde{p}_{1n}, \tilde{p}_{2n}, \ldots, \tilde{p}_{Mn})$ is employed, to Γ_n, for all $n = 1, 2, \ldots, L$. Unlike the original optimization problem (over the infinite set $\Upsilon(B)$) , after restricting the solution set considerably (to the finite set φ) , we can now solve the optimization problem with linear optimization techniques like the linear programming method. The output of the optimization algorithm will be the $\Gamma_1, \ldots, \Gamma_L$ values.

Let $\mathbf{\Gamma} = (\mid \Gamma_1 \mid, \mid \Gamma_2 \mid, \ldots, \mid \Gamma_L \mid)$ and $\mathbf{A} = ((a_{ij}))$ where $a_{ij} = \frac{W}{\kappa_i} \frac{\tilde{p}_{in}}{\sum_{j\neq i} \tilde{p}_{jn} + \frac{I_i}{g_i}}$ then the optimization problem can be written as:

$$maximize\ \mathbf{1A\Gamma} \text{ with the constraints}^2\text{: } \mathbf{A\Gamma} \geq \rho \tag{8}$$

Notice that although we have found the throughput maximizing scheduling for the duration $(0, t)$, this scheduling will satisfy all the QoS requirements of each wireless user but the extra capacity will be transferred to users with better channel gains. This is unfair and unpractical. If we introduce the additional condition that each user will share the extra capacity proportional to their QoS requirements, then it is easy to see that the throughput maximization problem is equivalent to finding the minimum feasible value t, by when all the QoS requirements are satisfied. Therefore we transform the throughput optimization problem to the following minimization problem:

$$minimize \sum_{i=1}^{L} \mid \Gamma_i \mid \ \text{ with the constraints}^2\text{: } \mathbf{A\Gamma} \geq \rho \tag{9}$$

It is not hard to see that the inequality in the constraint above can be replaced by equality

[2]the inequalities are componentwise

since the optimum solution will satisfy the rate requirements with equality. Let $\mathbf{\Gamma}^* = (\mid \Gamma_1^* \mid, \mid \Gamma_2^* \mid, \ldots, \mid \Gamma_L^* \mid)$ be the solution to the last optimization problem which can be solved by linear optimization methods like linear programming. We skip some details but as a result of our optimization problem, the optimum solution $\mathbf{\Gamma}^*$ will have at least $L - M$ zero values and at most M non-zero values. The efficient linear programming techniques like $simplex\ method$ can be used in order to achieve fast results. Since in the algorithm we would only have $M\ basic\ feasable\ solutions$ at any iteration, we will have $O(ML)$ worst-case time and $O(M^2)$ best-case time. The memory need is only $O(M^2)$.

We can now construct the power allocation and scheduling scheme. The proposed power allocation scheme will work as follows: As soon as one of the values of measured values of **g(t)** or I_i's changes, the base station will run the optimization algorithm and will assign the powers according to the optimization output, meaning for a period of $\mid \Gamma_1^* \mid$, the powers vector $\mathbf{V}_1$ will be transmitted, then for a period of $\mid \Gamma_2^* \mid$, the powers vector $\mathbf{V}_2$ will be transmitted and so forth. As soon as the duration $\mid \Gamma_{2^M}^* \mid$ where the powers vector $\mathbf{V}_{2^M}$ is assigned elapses, the base station will repeat the exact same assignments until either the value of **g(t)** or I_i changes. Remember that there are only at most M non-zero Γ_i^* values, meaning that there will be a time-division round robin between at most M vertices.

Our analytical findings were backed by our extensive simulation results. Specifically, the total throughput per bandwidth per total average power (Kbits/Hz/Watts), was 2.45, 2.78 and 3.08 times the total throughput of regular CDMA systems where all users transmit simultaneously for our 7, 9 and 11 user simulations respectively.

C. Conclusion

We proposed a novel power allocation and scheduling scheme for multimedia CDMA based wireless wide area networks. Unlike traditional CDMA networks, our proposed algorithm transmits to the wireless stations with certain power levels and durations which is a result of an optimization problem we define and solve. The resulting algorithm dynamically adapts to the changes in the channel gains, intercell interference and background noise levels and therefore implements the optimum scheduling for that specific circumstance. Our analytical findings are backed by our simulations results.