Resource Allocation for Wireless CDMA Systems Supporting Heterogeneous Multimedia Traffic *

Yee Sin Chan and J. W. Modestino

Center for Image Processing Research
and
Electrical, Computer, and Systems Engineering Department
Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Ph.:+1 518 276 6823, Fax:+1 518 276 6261, Email:chanys,modestin@ecse.rpi.edu

ABSTRACT

We investigate the resource allocation problem for future third-generation (3G) wireless CDMA cellular networks supporting heterogeneous multimedia applications. More specifically, we demonstrate that the network resources consumed by an individual user in a direct-sequence spread-spectrum CDMA network can be taken as the product of the allocated source coding rate, R_s*, and the energy per bit normalized to the multiple-access interference noise density,* E_b/N_0*. We propose a joint source coding and power control (JSCPC) approach for allocating these two quantities to an individual user, subject to a constraint on the total available bandwidth, to simultaneously maximize the per-cell capacity while maximizing the end-user, application-specific, quality-of-service (QoS) for different multimedia services.*

1 INTRODUCTION

Code division multiple-access (CDMA) is receiving considerable attention as the core multiple-access technology in the development of upcoming third-generation (3G) wireless cellular networks. The anticipated advantages of CDMA in cellular applications include improved channel capacity, asynchronous transmission capability, simple call admission control (CAC) procedures, soft-capacity limits and reduced sensitivity to multipath fading [1]. Future 3G wireless networks are expected to integrate different types of multimedia traffic, such as voice, data, and compressed images and video. However, the associated high bit-rate and latency requirements have made the reliable transport of multimedia particularly difficult in the error-prone mobile wireless environment. This paper investigates the resource allocation problems associated with wireless CDMA networks supporting heterogeneous multimedia traffic where individual users have different QoS requirements.

In our previous work [2] and [3], we have shown that network resources consumed by an individual user in a wideband wireless CDMA system can be taken as the

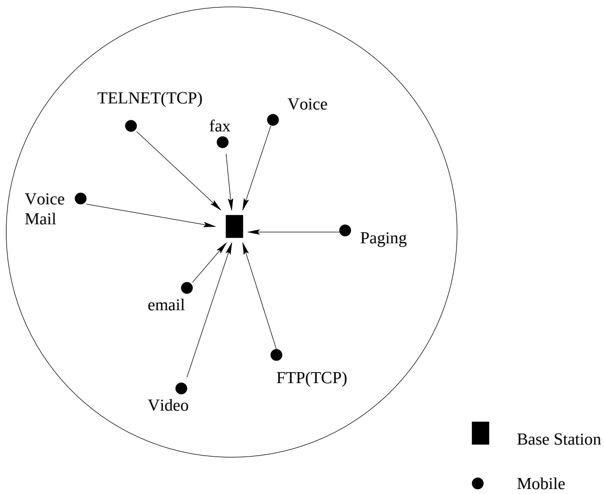

Figure 1: Reverse link of a single cell CDMA supporting multimedia services.

product of the source data rate, R_s, and the energy per bit normalized to the multiple-access interference ratio, E_b/N_0. Based on this observation, we proposed a joint source coding and power control approach (JSCPC) to allocating these two quantities to an individual video user, subject to a constraint on the total available bandwidth, to simultaneously maximize the quality of the delivered video while minimizing the network resource requirements of an individual user. More specifically, based on the network resources allocated to an individual user ($R_s \times E_b/N_0$), rather than taking R_s and E_b/N_0 as independent and unrelated parameters, we consider the optimum tradeoff between R_s and E_b/N_0.

Figure 1 shows a single-cell CDMA system supporting multimedia services with a wide range of transmission rates and various error-delay performance requirements. Based on the network requirements, we classify the services provided into 3 classes. Class I consists of real-time connections with zero delay tolerance, such as voice and compressed video. In this paper we specifically concentrate on the resource allocation problem associated with real-time video transmission which is considered to be one of the most important multimedia applications in future 3G networks. The appropriate end-to-end QoS in

*This work was supported in part by CenSSIS, the Center for Subsurface Sensing and Imaging Systems, under the Engineering Research Centers Program of the National Science Foundation under award number EEC-9986821.

this case is defined as the end-user average Peak-Signal-to-Noise Ratio (PSNR). Class II includes nonreal-time, delay sensitive, error-free services, such as Telnet, interactive games and similar applications associated with the use of the transport control protocol (TCP). Class III consists of message-oriented, delay-tolerant, error-free services, such as email, fax and data file transfer. The classifications are based on the characteristics of the anticipated services to be supported in future 3G systems and is similar to the classification used in IMT-2000 proposals [4, 5, 6], in which three classes of data services are considered: low delay data bearer services (LDD), long constrained delay data bearer services (LCD), and unconstrained delay data bearer services (UDD).

We consider only channel errors due to uncorrelated bit errors which can be achieved by sufficient interleaving to randomize the burst errors at the channel decoder output. This corresponds to representing the path from the input of the channel encoder to the output of the channel decoder as a memoryless binary symmetric channel (BSC). Although this does not fully capture channel error effects on many real-world wireless channels, it provides considerable insights into the joint interactions and the effects of both video coding rate selection and channel error control on end-user PSNR. We assume the use of binary convolutional codes on the additive white Gaussian noise (AWGN) channel so that the bit-error probability can be accurately approximated by the well-known transfer-function bounds [7]. In our analysis, we assume perfect power control. By perfect power control, we mean that the received power of each user is adjusted so that the received E_b/N_0 is fixed at the desired level.

2 Problem Formulation

Using the same approach as in [8] for determining the capacity of a single-cell CDMA cellular system supporting voice traffic, we first extend the results for perfectly power-controlled homogeneous voice users to heterogeneous multi-rate traffic with different power levels. Assuming there are N users in the system and ignoring the background noise due to spurious interference as well as thermal noise, user i with information bit-rate R_i operates with power level $S_i = E_i R_i$, where E_i is the corresponding energy per information bit. The bit-energy to multiple-access interference (MAI) ratio of user i will then be

$$\frac{E_i}{N_0} = \frac{\frac{S_i}{R_i}}{\sum_{j \neq i}^{N} \frac{S_j}{W_T}}; i = 1, 2, 3, \ldots, N, \tag{1}$$

where W_T is the total bandwidth in Hz and $N_0/2$ is the two-sided noise power spectral density due to MAI in watts/Hz. It follows that

$$\frac{E_i R_i}{N_0} = \frac{S_i W_T}{\sum_{j \neq i}^{N} E_j R_j}. \tag{2}$$

If we assume a large number of users and that the MAI experienced by any user due to the total traffic from all the other users is approximately equal, summing (2) for N users yields

$$\sum_{i=1}^{N} \sum_{j \neq i}^{N} \frac{E_j}{N_0} R_j = (N-1) \sum_{i=1}^{N} \frac{E_i}{N_0} R_i = N W_T, \tag{3}$$

so that for large N,

$$\sum_{i=1}^{N} \frac{E_i}{N_0} R_i \approx W_T. \tag{4}$$

The last expression represents the approximately optimal solution (in term of resource allocation) of the necessary and sufficient conditions for the existence of a feasible transmit power assignment to all users derived by Sampath *et al.* [9] in a single-cell CDMA system[1].

We define the resource requirement of an individual user operating with source rate R_i and bit energy to MAI ratio E_i/N_0 as the *equivalent bandwidth*, $W_{eq}^{(i)} \triangleq R_i E_i/N_0$, with units of Hz. Basically, it is the total bandwidth resource allocated to an individual user in a spread-spectrum CDMA network. The capacity of a CDMA system is often referred to as interference-limited, while other systems like FDMA or TDMA are bandwidth-limited [10]. As a result, the majority of research efforts on improving CDMA system performance have been concerned with power control or simply interference suppression techniques. However, (4) clearly shows that CDMA systems are both bandwidth-limited as well as interference-limited.

To provide a better understanding of this issue, consider the following example: given a fixed $W_{eq}^{(i)}$, if we increase the bit-rate R_i, the quantity E_i/N_0 must be decreased accordingly so as to achieve the same overall capacity. However, decreasing E_i/N_0 also means an increase in bit-error rate (BER) for the corresponding user which may ultimately lead to decreased end-to-end performance. Indeed, often the end-to-end performance for larger BER degrades rapidly with increasing bit-rate as we demonstrate in what follows. Conceptually, the problem can be solved by using a more powerful channel coding scheme which can effectively improve the overall capacity (in terms of the number of users supported) of a spread-spectrum CDMA network by decreasing the required E_i/N_0 as shown in [11], [12], [13]. Practically, however, the system complexity and delay constraints for different multimedia applications often limit the choice of channel coding schemes that can be employed. For example, Viterbi decoding becomes too complex to implement for constraint lengths much beyond 10, which is often required to achieve low bit-error probability with convolutional codes. Hence, the crucial issue here is, using a practical channel coding scheme, how to maximize

[1] In particular, [9] shows that the necessary and sufficient conditions for the existence of a feasible transmit power assignment to all users satisfying both the bit-rate, r and QoS in terms of energy-per-bit to MAI density ratio γ is $\sum_{j=1}^{N} \frac{1}{\frac{W_T}{r_j \gamma_j}+1} < 1$, where W_T is the total bandwidth. Clearly, this can be reduced to $\sum_{j=1}^{N} r_j \gamma_j \leq W_T$, if W_T is sufficiently large with respect to individual requirements and equality holds for optimal allocations.

the performance of a single-cell CDMA system supporting heterogeneous multimedia traffic by allocating both the bit-rate and the power level to an individual user subject to a constraint on the allocated resource, the equivalent bandwidth $W_{eq}^{(i)}$.

3 JOINT SOURCE CODING-POWER CONTROL METHODOLOGY

3.1 Reliable Delay-Tolerant Data Services

The QoS of this class of services is defined by the throughput. We assume that each user generates a sequence of packets of fixed-length L, and reliability of data communications is guaranteed through error detection and retransmission (ARQ) in conjunction with the FEC coding. Hence, the source rate of a data user is [2]

$$R_s = G \times LR_c, \tag{5}$$

where G is the average number of packets generated per second which consists of both newly generated packets and retransmitted packets that were unsuccessfully received during a previous transmission. L is the length of a packet and R_c is the channel coding rate in bits/channel use. Hence, the QoS of data users, defined in terms of the number of successfully transmitted information bits (including overhead) , or the *throughput*, S is

$$S(R_s, E_b/N_0) = G \times L \times R_c \times P_s, \tag{6}$$

where P_s is the packet success probability which is equal to[3]

$$P_s = (1 - P_b(E_b/N_0))^{LR_c}. \tag{7}$$

Hence, we define the optimal allocation for a single user with equivalent bandwidth requirement W_{eq}, in Hz, to be given as

$$\begin{aligned} S(W_{eq}) &= max\ S(R_s, E_b/N_0) \\ &= max\ R_s(1 - P_b(E_b/N_0))^{LR_c} \\ &= max\ \frac{W_{eq}}{E_b/N_o}(1 - P_b(E_b/N_0)^{LR_c}, \end{aligned} \tag{8}$$

where the maximization is taken over all R_s and E_b/N_0 subject to the constraint (10). The E_b/N_0 that maximizes the throughput S is easily computed by setting the derivative of (8) with respect to E_b/N_0 equal to zero. The optimal allocation $(R_s^*, E_b^*/N_0)$ must then satisfy

$$1 - P_b(E_b/N_0) + P_b'(E_b/N_0)LR_c = 0, \tag{9}$$

with constraint

$$W_{eq} = R_s \times \frac{E_b}{N_0}. \tag{10}$$

From (9), it can be observed that the optimal E_b^*/N_0 is independent of W_{eq}, or equivalently the optimal source

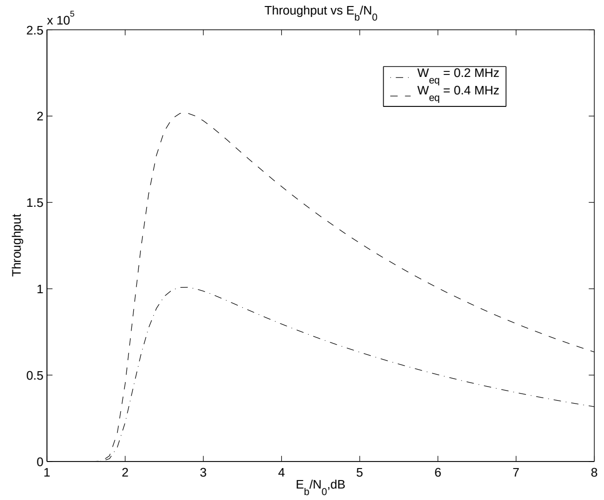

Figure 2: Throughput vs. E_b/N_0 with $R_c = 1/2$, $L = 256$, $K = 9$.

coding rate increases linearly as W_{eq} increases. This means that when the W_{eq} of an individual user is increased due to a reduction in competing traffic, the most effective way of increasing the throughput is to increase the data rate while keeping the target E_b/N_0 the same. This gives an alternative interpretation of the results in [14] that the optimal transmission rate decreases inverse linearly as the MAI level increases. Figure 2 shows the throughput vs. E_b/N_0 with packet length $L = 256$ using binary convolutional codes[4] with $R_c = 1/2$ and constraint length $K = 9$. It can be noted that there is an optimal setting for E_b/N_0. Below the optimal value, the throughput is limited by the low successful transmission probability. Above this value, although the probability of successful transmission increases due to higher E_b/N_0 and thus lower BER, the throughput is limited by the low data rate (R_s). Also note that the optimal, E_b^*/N_0 for the different W_{eq} are the same.

3.2 Reliable Delay-Sensitive Data Services

For delay sensitive services such as typical Telnet and interactive data retrieval services, the average delay has to be less than a maximum acceptable time (D_T) defined by the application-specific QoS requirements, or equivalently,

$$D \leq D_T, \tag{11}$$

where average packet delay D is defined as the average time from when a packet is generated until it is successfully received. Hence, the optimal allocation can be defined the same as (8) subject to both the constraints (10) and (11).

The average delay normalized to the packet transmission time, assuming immediate acknowledgment, is given as [15]

$$D = 1 + d + (N_{tr} - 1) \times (1 + \delta + 2d), \tag{12}$$

where δ is the mean retransmission delay and $2d$ is the round-trip propagation delay. $N_{tr} - 1$ is the average num-

[2] Since we concentrate on a single user in this section, we will drop the indexing on the i^{th} user and write W_{eq} in place of $W_{eq}^{(i)}$ and use R_s and E_b/N_0 in place of R_i and E_i/N_0, respectively.

[3] Here $P_b(E_b/N_0)$ represents the bit-error probability as a function of E_b/N_0.

[4] In this case, we make use of transfer-function bounds to evaluate $P_b(E_b/N_0)$.

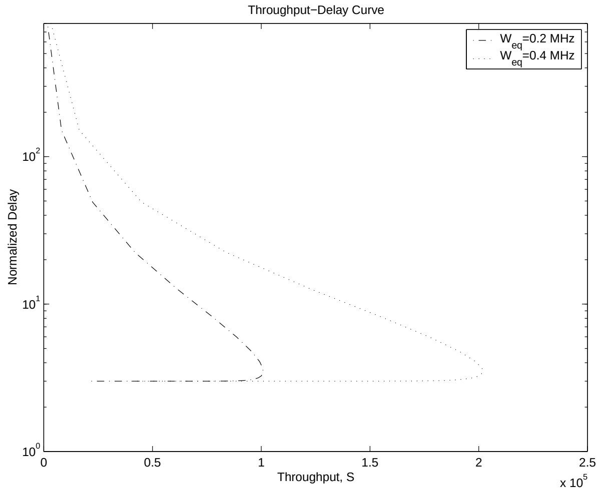

Figure 3: Throughput-Delay curve with $R_c = 1/2$, $L = 256$, $K = 9$, $\delta = 5$, $d = 2$.

ber of retransmissions needed for a packet to be successfully received, and is given as

$$N_{tr} = \sum_{k=1}^{\infty} k P_s (1 - P_s)^{k-1}. \quad (13)$$

Here we assume positive acknowledgment packets are always correctly received.

Figure 3 illustrates the tradeoff between the average delay D, given by (11), and the throughput S, given by (6), for different values of W_{eq}. Here, the curves for a fixed W_{eq} are plotted with E_b/N_0 as a parameter and R_s determined from the constraint (10). Given a fixed D_T, only throughputs corresponding to $D \leq D_T$ are achievable. Hence, if an individual user has a tight delay constraint, he may have to operate below the point with the highest throughput. This corresponds to transmitting at a higher E_b/N_0 with a subsequent decrease in source coding rate R_s due to the constraint (10).

3.3 Real-Time Video Traffic

We demonstrate the efficacy of this approach using the ITU-T H.263+ video source coder, although the approach is generally applicable to other source coding schemes as well. Due to the real-time requirement, we only consider FEC coding for video transmission, i.e., we specifically rule out ARQ techniques. Errors at the receiver end are handled by a passive error concealment scheme [16] to avoid the delay due to retransmissions. The bit-rate of the H.263+ source coder is adjusted through the choice of the corresponding quantization parameters (QPs). Smaller QPs result in finer quantization and consequently larger bit-rate. For the single-layer scheme considered here, the QPs for the $I-$ and $P-$ pictures are denoted as QP_I and QP_P, respectively. In H.263+, QP_I and QP_P may be individually drawn from the values $\{1, 2, 3, \ldots, 31\}$. The overall end-to-end performance of the delivered video will be measured as the average PSNR over a sequence of N_f consecutive frames and includes channel error effects as well as source coding losses. For a single-layer video system, as considered in this work, we define the optimal operational distortion-rate characteristic for a single user with equivalent bandwidth requirement W_{eq}, in Hz, to be given as

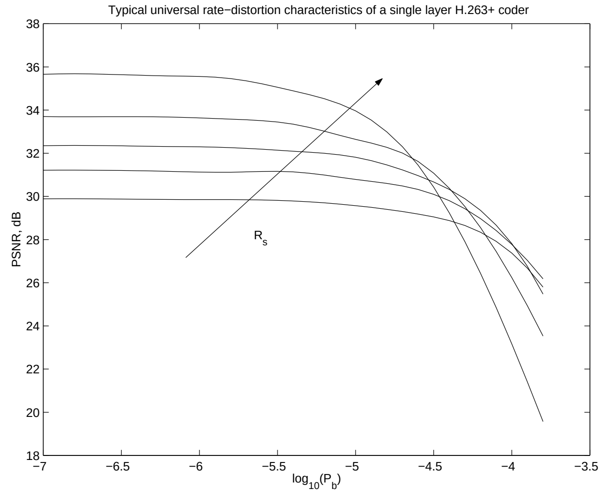

Figure 4: Typical universal rate-distortion characteristics, $PSNR(R_s, P_b)$, for a single-layer H.263+ encoder.

$$PSNR(W_{eq}) = max\ PSNR(R_s, E_b/N_0), \quad (14)$$

where the maximization is performed over all R_s and E_b/N_0 of interest, subject to the constraint given by (10). In (14) the quantity $PSNR(R_s, E_b/N_0)$ represents the $PSNR$ for specified values of source coding rate R_s and E_b/N_0. Exact numerical evaluation of this quantity is analytically intractable although, at least in principle, it could be evaluated through simulations. Unfortunately, this requires explosive computational effort if $PSNR(R_s, E_b/N_0)$ is to be reliably evaluated over a wide range of source coding rates, choice of channel codes and channel conditions, represented here by E_b/N_0.

In a related problem treated in [17], it was shown that much of the computational complexity in evaluating $PSNR(R_s, E_b/N_0)$ can be reduced through the use of universal distortion-rate characteristics $PSNR\,(R_s, P_b)$. These are families of curves, parameterized by a fixed source coding rate R_s, plotted as a function of the bit-error probability P_b associated with the BSC used to model the serial channel from the input of the channel encoder to the channel decoder output. These characteristics are universal in the sense they need not to be recomputed when the channel coding or the channel conditions change and, as described below, can be used to evaluate $PSNR(R_s, E_b/N_0)$ provided there exists an explicit relationship between P_b and E_b/N_0, or at least tight bounds on bit-error probability performance.

A representative set of universal distortion-rate characteristics is illustrated in Fig. 4. These characteristics were obtained through simulation using $N_f = 120$ frames for the QCIF Susie sequence at frame rate $f_s = 30$ fps. Observe that large bit-error probabilities P_b will favor the use of smaller R_s if the resulting PSNR is to be maximized while the opposite is true for small P_b.

Given a family of universal distortion-rate characteristic curves, together with the appropriate bounds on

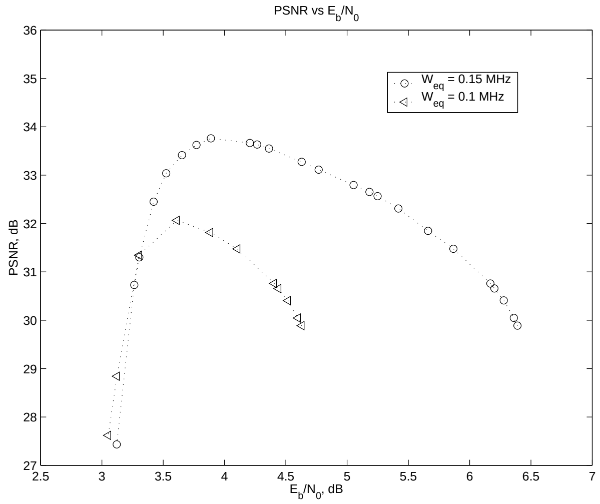

Figure 5: PSNR vs E_b/N_0 with $R_c = 1/2, K = 9$.

bit-error probability behavior for a particular modulation/coding scheme as a function of channel parameters, the corresponding optimal resource allocation represented by (14) can be determined by the following procedure: for a specific class of channel codes and choice of channel coding rate (R_c), there is a one-to-one correspondence between the bit-error probability (P_b) and the signal-to-MAI density per information bit (E_b/N_0). Hence, given a value of $W_{eq} = R_s E_b/N_0$, for each choice of source coding rate R_s, with the constraint $E_b/N_0 = W_{eq}/R_s$, the resulting value of end-user $PSNR(R_s, E_b/N_0)$ can be found from the corresponding universal distortion-rate characteristics curve. The optimal end-to-end distortion $PSNR(W_{eq})$, given by (14) under constraint (10), can then be found by taking the convex hull of all such operational points.

Figure 5 shows a plot of PSNR versus E_b/N_0 under different constraints on W_{eq} using a binary convolution code[5] with $R_c = 1/2$ and constraint length $K = 9$. Each point on the curve represent a different choice of R_s and E_b/N_0. It can be seen that there is an optimal assignment for R_s and E_b/N_0. For low E_b/N_0, or equivalently large R_s, the end-user PSNR performance is limited by the high channel error rates, while at large E_b/N_0 which correspond to low R_s, the performance is limited by the source coding errors.

4 Summary and Conclusions

We demonstrate that the network resources consumed by an individual user in a direct sequence spread-spectrum CDMA network can be taken as the product of the allocated source coding rate, R_s, and the energy per bit normalized to the multiple-access interference noise density, E_b/N_0. We describe a joint source coding and power control approach (JSCPC) for allocating these two quantities to an individual user, subject to a constraint on the total available resources, to simultaneously maximize the per-cell capacity while maximizing the end-user, application-specific, quality-of-service (QoS) for heterogenous multimedia services. This approach is illustrated for 3 distinct service classes and specific results are provided for each service class.

[5]Again we make use of transfer-function bounds to evaluate P_b as a function of E_b/N_0.

References

[1] R. Prasad and T. Ojanpera, "An overview of CDMA evolution toward wideband CDMA," *IEEE Communications Surveys, http://www.comsoc.org/pubs/surveys*, vol. 1, no. 1, pp. 2–29, Fourth Quarter 1998.

[2] Y. S. Chan and J. W. Modestino, "A joint source coding and power control approach for maximizing the capacity of CDMA wireless networks supporting heterogeneous video traffic," in *Proc. of 11th Packet Video Workshop*, Kyongju, Korea, May 2001, pp. 52–61.

[3] Y. S. Chan and J. W. Modestino, "Transport of scalable video over CDMA wireless networks: A joint source coding and power control approach," in *Proc. of the International Conf. on Image Processing*, Thessaloniki, Greece, October 2001, vol. 2, pp. 973–976.

[4] ARIB IMT-2000 Study Committee, *Japan's Proposal for Candidate Radio Transmission Technology on IMT-2000: W-CDMA*, Association of Rdio Industries and Businesses (ARIB), Japan, June 1998.

[5] ETSI/UTRA, *The ETSI UMTS Terrestrial Radio Access (UTRA) ITU-R R-TT candidate Submission*, European Telecommunication Standard Institution (ETSI), June 1998.

[6] TIA/TIA45.5.4, *The cdma2000 ITU-R RTT candidate submission (0.18)*, Telecommunications Industry Association (TIA), USA, 1998.

[7] J.G. Proakis, *Digital Communications*, McGraw-Hill Book Company, 2nd edition, 1989.

[8] K. S. Gilhousen et al., "On the capacity of a celluar CDMA system," *IEEE Trans. on Vehicular Tech.*, vol. 40, no. 2, pp. 303–312, May 1991.

[9] A. Sampath, P. S. Kumar, and J. M. Holtzman, "Power control and resource management for a multimedia wireless CDMA system," in *Proc. PIMRC'95*, Toronto, Canada, September 1995, pp. 21–25.

[10] T. S. Rappaport, *Wireless Communications: Principles & Practice*, Prentice Hall, Upper Saddle River, NJ, 1996.

[11] J. Y. N. Hui, "Throughput analysis for code division multiple accessing of the spread spectrum channel," *IEEE J. on Select. Areas Commun.*, vol. SAC-2, no. 4, pp. 482–486, July 1984.

[12] A. J. Viterbi, "Spread spectrum communications-myths and realities," *IEEE Communications Magazine*, pp. 11–18, May 1979.

[13] A. J. Viterbi, "Very low rate convolutional codes for maximum theoretical performance of spread-spectrum multiple-access channels," *IEEE J. on Select. Areas Commun.*, vol. 8, no. 4, pp. 641–649, May 1990.

[14] S. J. Oh and K. M. Wasserman, "Dynamic spreading gain control in multiservice CDMA networks," *IEEE J. on Select. Areas Commun.*, vol. 19, no. 5, pp. 918–927, May 2000.

[15] L. Kleinrock and F. A. Tobagi, "Packet switching in radio channels: Part I - carrier sense multiple-access modes and their throughput-delay characteristics," *IEEE Trans. on Communications*, vol. 23, no. 12, pp. 1400–1416, December 1975.

[16] ITU-T/SG16/Q15, "Error resilience for mobile using GOB sync and error concealment," in *Fifth Meeting: Whistler*, British Columbia, Canada, July 1998.

[17] M. Bystrom and J. W. Modestino, "Combined source-channel coding for video transmission over an Additive White Gaussian Noise Channel," *IEEE J. on Select. Areas Commun.*, vol. 18, no. 6, pp. 880–890, June 2000.