Dynamic Resource Scheduling Schemes for W-CDMA Systems

Özgür Gürbüz, Cisco Systems, Inc.

Henry Owen, Georgia Institute of Technology

ABSTRACT

W-CDMA is the strongest candidate for the air interface technology of third-generation wireless communication systems. Dynamic resource scheduling is proposed as a framework that will provide QoS provisioning for multimedia traffic in W-CDMA systems. The DRS framework monitors the traffic variations and adjusts the transmission powers of users in an optimal manner to accommodate different service classes efficiently. Variable and optimal power allocation is suggested to provision error requirements and maximize capacity, while prioritized queuing is introduced to provision delay bounds. A family of DRS algorithms has been devised along these dimensions for obtaining different levels of QoS. The DRS schemes are discussed in terms of queuing and bandwidth allocation with an emphasis on their impact on delay QoS.

INTRODUCTION

Third-generation wireless networks will support heterogeneous traffic, consisting of voice, video, and data (i.e., multimedia). Quality of service (QoS) is the major issue for these applications in any kind of networking environment, and wireless QoS is a complex problem due to the time-varying characteristics of the channel and user mobility. In this research, we are investigating an efficient QoS provisioning framework to be applied to third-generation wireless networks. We are considering wideband code-division multiple access (W-CDMA) because it is the strongest candidate for the air interface technology. In particular, we refer to the Japanese W-CDMA standard by Association of Radio Industries and Businesses ARIB [1], proposed for International Mobile Telecommunications 2000 (IMT-2000) [2].

In [3], we proposed dynamic resource scheduling (DRS) as an adaptive power assignment and control mechanism for W-CDMA systems. DRS applies a power optimizing approach to the wireless QoS problem by coordinating simultaneous transmissions. It prevents severe interference situations with the ultimate objective of provisioning bandwidth and error QoS while minimizing the total power. One major contribution is flexible resource allocation considering variable-rate users. The second major contribution is traffic classification and prioritized power scheduling that provide guaranteed delay profiles. Applying both a prioritized and a rate-adaptive power control scheme fully utilizes the statistical multiplexing gain for CDMA capacity [4].

In this article we extend DRS to a family of algorithms and examine these DRS algorithms from the delay perspective for temporal QoS. The rest of the article is organized as follows. We discuss the queuing dimension of DRS with single- and multiple-queue implementations, emphasizing the advantages of classification and prioritization. We then describe the fixed- and variable-rate architectures, and their impact on delay. We present our simulation results for comparison of DRS schemes. Following that, we present our conclusions.

SINGLE- VS. MULTIPLE-QUEUE IMPLEMENTATION

DRS is motivated by the scheduled CDMA (S-CDMA) scheme that has been proposed as a medium access control protocol for wireless asynchronous transfer mode (ATM) networks. S-CDMA is a hybrid time-division multiple access (TDMA)/CDMA scheme that coordinates the timing and power levels of packets in the wireless ATM environment [5].

The considered Japanese W-CDMA system utilizes pure CDMA, without any time multiplexing. Hence, the time-slot-based S-CDMA scheme is not directly applicable in this architecture. In the DRS framework, the S-CDMA scheme has been modified to be applied to the W-CDMA physical layer and enhanced for flexible multimedia services. Still, simultaneous transmissions should be coordinated in order to preserve the required signal-to-interference ratio (SIR) for everyone, but there is no need for a complex time scheduler as in S-CDMA. Our proposed architecture, DRS, consists of a

0163-6804/00/$10.00 © 2000 IEEE

simple buffer and a *power scheduler* (Fig. 1). The user requests are collected in a first come first served (FCFS) fashion, and *connection admission control* and *optimized power control* in the power scheduler is applied to the backlogged requests. The optimized power for user i is given as [5, 6]

$$P_i = \frac{g_i.\eta_0.W}{h_i.\left(1-\sum_{j=1}^{N} g_j\right)}, \tag{1}$$

where N is the admitted number of users, and g_i denotes the *power index* of user i as the required radio resource from the network. The feasibility of the optimal solution complements the following connection admission control (CAC) test [5, 6]:

$$\sum_{j=1}^{N} g_j < 1 - \frac{\eta_0.W}{\min_i\left(\frac{\tilde{P}_i.h_i}{g_i}\right)}. \tag{2}$$

Despite the optimal power assignment, single-queue DRS may not be desirable for mixed traffic scenarios. Since all the requests are collected in a FCFS fashion, the mobile stations geographically close to the base station will be considered earlier in the admission control test in Eq. 2. This way, closer mobiles can block admission of farther mobiles, which is similar to the *near-far effect* in CDMA systems [7]. The classical near-far effect due to path loss is further complicated for the multimedia case, due to the differences in resource requirements. Path loss compensation, h_i, in the optimal solution in Eq. (1) is not sufficient to overcome this effect. Blocking of requests can cause unwanted delays, delay jitters, and packet losses for users. Non-real-time data traffic is carried by advanced transport protocols like TCP, which employ error, flow, and congestion control. Hence, losses are recoverable, and delay is tolerable for those sources. Meanwhile, the real-time traffic flows employ fast transport mechanisms like UDP, so losses and delay performance can be critical. Therefore, a classifier is essential to distinguish and separately treat different classes of traffic in order to provision QoS requirements.

We introduce *prioritized queuing* to solve the new near-far problem in DRS. Figure 2 depicts our prioritized DRS architecture. The requests are classified and backlogged in two separate queues: *guaranteed* and *best effort*. The classification is according to the traffic characteristics of the requested service. The guaranteed queue holds requests that have been promised to be served at a predefined rate. Considering ATM technology as an example, this involves real-time and non-real-time variable bit rate (rt/nrt-VBR), constant bit rate (CBR), and minimum-rate available bit rate (ABR) services. The best effort queue enqueues unspecified bit rate (UBR) and excess ABR requests. The admission control test is first applied on the requests at the head of the guaranteed queue. The best effort queue is served afterward, as long as power resources are available. This way, the best effort services use the leftover capacity. Resource availability is checked through the CAC test in Eq. 2.

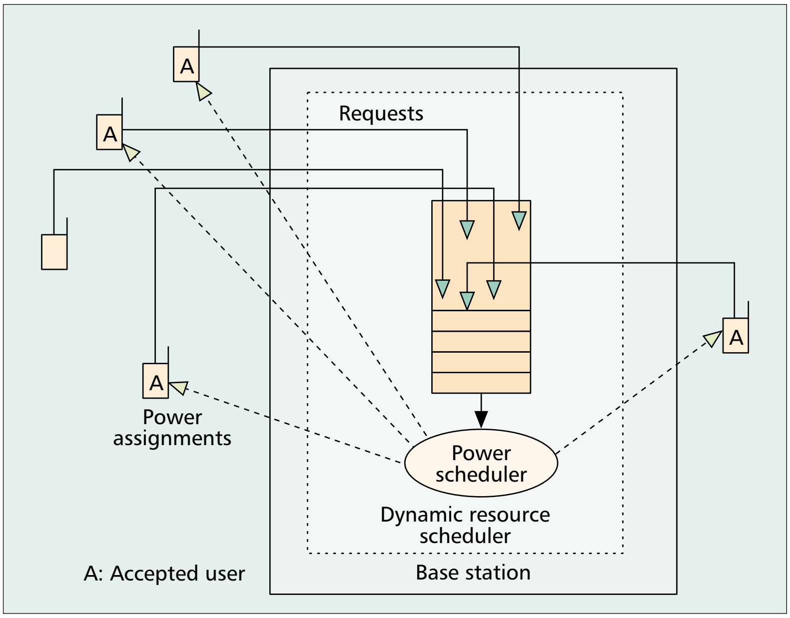

Figure 1. *DRS with a single queue.*

As an alternative to the multiple-queue scheme, a time scheduler can be employed to order the requests in the buffer according to their delay tolerances (as in S-CDMA). Consequently, the best effort services will be forced toward the tail of the queue to be served after the guaranteed requests. Nevertheless, this scheduler would require complexity and delay for the extra processing, while prioritized multiple queuing is fairly simple and fast.

Figure 3 shows the extended family of DRS algorithms. *Prioritized Fixed Spreading Gain* (*P-FSG*) is the first DRS scheme which implements S-CDMA with multiple queues. Both S-CDMA and P-FSG apply constant bandwidth allocation for users. Bandwidth flexibility is possible in both single and multiple queue dimensions to

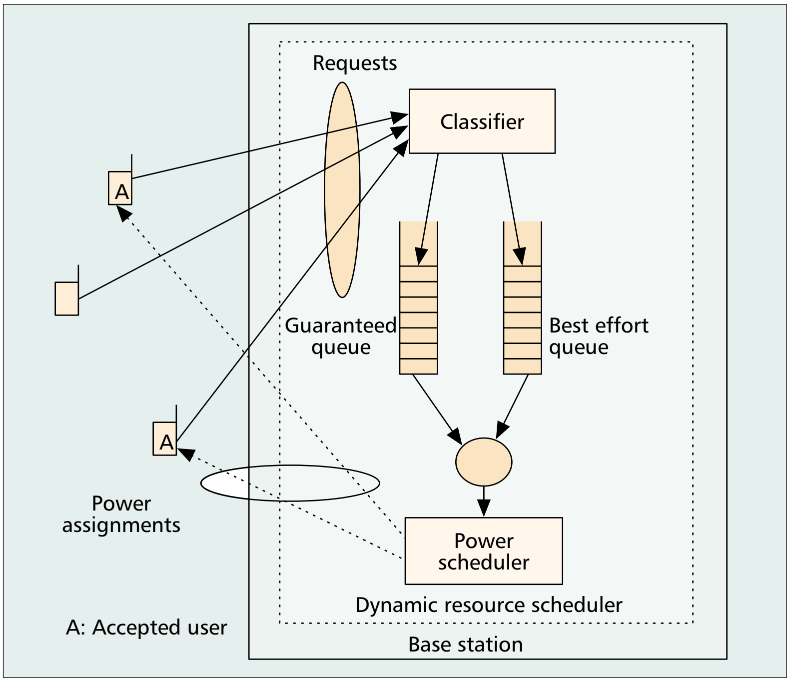

Figure 2. *DRS with prioritized queuing.*

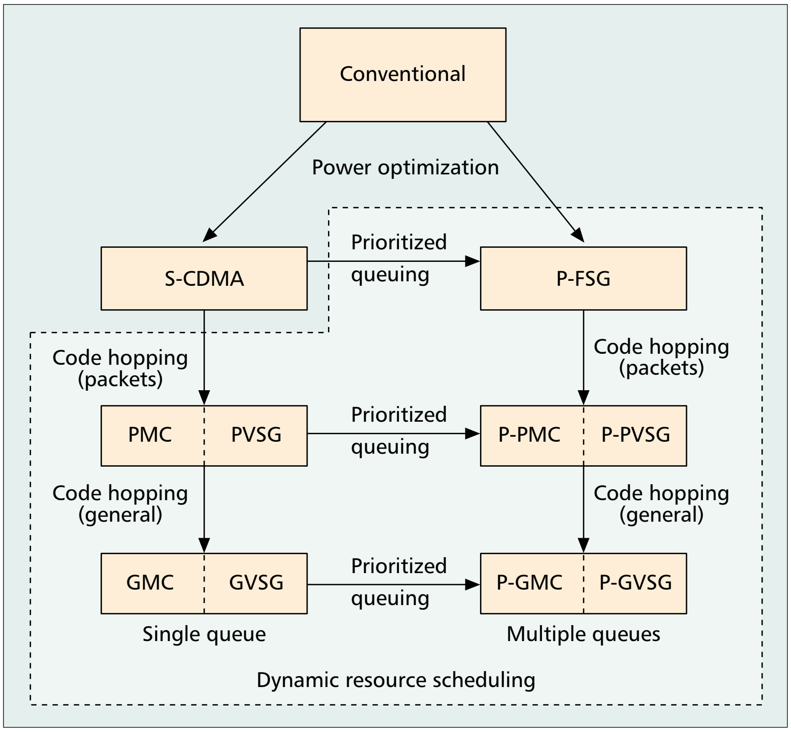

Figure 3. *The family of DRS algorithms.*

support variable rate sources efficiently. The next section discusses the fixed vs. variable bandwidth allocation dilemma in DRS.

FIXED VS. VARIABLE BANDWIDTH ALLOCATION

DRS algorithms are based on radio frame timing, rather than time slots as in S-CDMA. Since the service data units in W-CDMA are radio-frame-based, DRS fits perfectly into the W-CDMA infrastructure. The input data rate is subject to fluctuations, especially for multimedia traffic, and it is possible to track the variations in input traffic rate in a W-CDMA system on a radio frame basis [1].

We have suggested variable bandwidth allocation by assigning a set of codes to potential variable rate services and perform *code hopping* [3, 4]. Code hopping was previously proposed as a diversity technique for increasing voice capacity in CDMA systems [8]. We apply it as a means to switch to a code with an appropriate spreading gain to support the instantaneous data rate of the multimedia traffic. This is referred to as the *variable spreading gain* (*VSG*) implementation. W-CDMA utilizes orthogonal variable spreading factor (OVSF) codes, by means of which a wide range of data rate values are easily obtained [1]. An alternative way for high rates and flexible bandwidth is the *multicode* (*MC*) approach, where the number of code streams are varied instead of the spreading gain [9]. In both implementations, the CAC test in Eq. 2 takes the instantaneous demands in terms of power indices, g_i, and the power solution in Eq. 1 is updated according to instantaneous load. Employing variable resource allocation and adaptive power solutions, we obtained power savings and increased capacity [4]. We compared the two spreading strategies, VSG and MC, for implementing variable rate, and investigated their performance for semantic QoS in [10].

In this article we introduce variable rate on two levels: packet and general. In *Packet Variable Rate DRS*, we propose adaptive reservation for ABR and UBR flows; the only limitation on maximum rate is the physical spread bandwidth. In Fig. 3, packet variable schemes are abbreviated as *PVSG* and *PMC*, employing either VSG or MC spreading, respectively. In *Generalized Variable Rate DRS* schemes, code hopping is applied for all types of services. The two spreading strategies are referred as Generalized VSG (GVSG) and Generalized MC (GMC). The difference between packet- and general-level variable rate is that the generalized schemes provide more bandwidth flexibility at a cost of greater signaling overhead and computational complexity in the base station.

Variable rate can be adapted in multiple-queue architectures. *Prioritized Packet VSG/MC* (*P-PVSG/P-PMC*) and *Prioritized General VSG/MC* (*P-GVSG/P-GMC*) schemes are packet- and general-level variable rate schemes we developed for multiple queues.

DELAY PERFORMANCE ANALYSIS

The total delay is composed of *access delay*, *queuing delay*, *propagation delay*, and *transmission delay*. The propagation delay is small enough to be ignored among the other delay components, and the access delay is beyond our current scope since we assume a constant channel access scheme. Being a resource management framework, DRS schemes impact the queuing and transmission delays, due to different queuing architectures and variable bandwidth allocation.

An OPNET simulation model was constructed to model the DRS framework for the Japanese W-CDMA system by NTT DoCoMo. For the power and interference calculations, the model incorporates external simulation results for radio pipeline stages in order to represent quadrature phase shift keying (QPSK) modulation and rate 1/3 convolutional codes for error control functions in the receiver. The system bandwidth was 5 MHz, with 4.096 Mchips/s chip rate. Hence, the spreading gain ranged between 4 and 256. The receiver noise figure was taken as 5 dB, and thermal noise density was –174 dBm/Hz.

Four types of traffic sources were considered:

- CBR service: Real-time voice was modeled at 32 kb/s with required SIR of 3 dB. CBR service can be matched to *conversational class* traffic in IMT-2000. The delay tolerance has been specified as 80 ms [2].
- VBR service: A simple coded video pattern IBBPBBI with a frame rate of 30 frames/s was modeled [11]. The matched rates for these different sizes of frames were 1.024 Mb/s, 256 kb/s, and 64 kb/s. The required SIR was 4 dB. VBR service is the *streaming class* of IMT-2000. Its delay tolerance is 500 ms [2].
- ABR service: Variable size packets were created for each session. The packet size

was exponentially distributed. Required minimum rate was 256 kb/s; required SIR was 5 dB. This service corresponds to the *interactive class* of IMT-2000. Delay tolerance is on the order of 1 s [2].

- UBR service: Constant size packets (300 bits, since 300 is the packet data unit, PDU, length for packet services in W-CDMA) arriving in a Poisson fashion were used. The required SIR was 5 dB. UBR traffic falls into the *background class* in IMT-2000. This type of service has no delay requirement [2].

The traffic sources were assumed to employ shaping so that the produced traffic conforms to its specifications. A single cell configuration was considered. We assumed a fixed number of users, so call arrivals and departures were not modeled. For that reason we excluded the fixed spreading gain schemes, S-CDMA and P-FSG, for the time being, since their performance depends on session arrivals and departures. Our experiments aimed to show the effectiveness of single- and multiple-queue, packet-level and generalized variable rate DRS schemes for delay QoS provisioning. In our simulation scenario, an rt-VBR source was located behind a bursty packet source of UBR traffic to create the *near-far effect*.

We first observed the impact of prioritized queuing. We compared single- and multiple-queue DRS schemes considering packet-level and generalized variable rate. In Fig. 4, the delay metrics measured for PVSG, GVSG, P-PVSG, and P-GVSG schemes are illustrated with the x axis indicating the packet arrival rate for the bursty UBR source, and the y axis the average packet delay for the blocked VBR source. In single-queue cases (PVSG and GVSG), the UBR source acquired the power resources before the VBR source. Hence, as the UBR packet arrival rate was increased, the VBR packets started to accumulate in the queue. Since the delay tolerance was limited to 500 ms, packets with higher delay were dropped. Figure 5 presents the percentage of lost packets in the same experiment. The x axis in the figure again indicates the packet arrival rate for the UBR source, and the y axis the percentage of lost packets for the blocked VBR source. The rt-VBR service suffered from blocking by the bursty packet source in single-queue DRS schemes. Meanwhile, the schemes with prioritized architecture (P-PVSG and P-GVSG) perfectly provisioned the delay QoS of the same source to 20 ms, and the loss probability was measured as zero. Throughout the simulations, the other sources (CBR and ABR) were not affected by UBR blocking for this particular geographical configuration.

Our results prove that service classification offers *delay provisioning* for guaranteed services by preventing bulky packet requests. The only drawback is that the best effort packets can suffer from higher delays, but this is not crucial since these flows are more tolerant of delay by nature. This can be improved by variable bandwidth allocation as a higher packet throughput is obtained [4].

The outcome of variable bandwidth is observed in Fig. 6. The single-queue schemes, PVSG and GVSG, provide the same amount of bandwidth for UBR, since they both block the VBR source. Having the same bandwidth, their delay performances are similar. When prioritized queues are employed, best effort traffic experiences delays in both cases. Variable rate in the general sense provides lower latency than packet-level variable rate. This is due to the instanta-

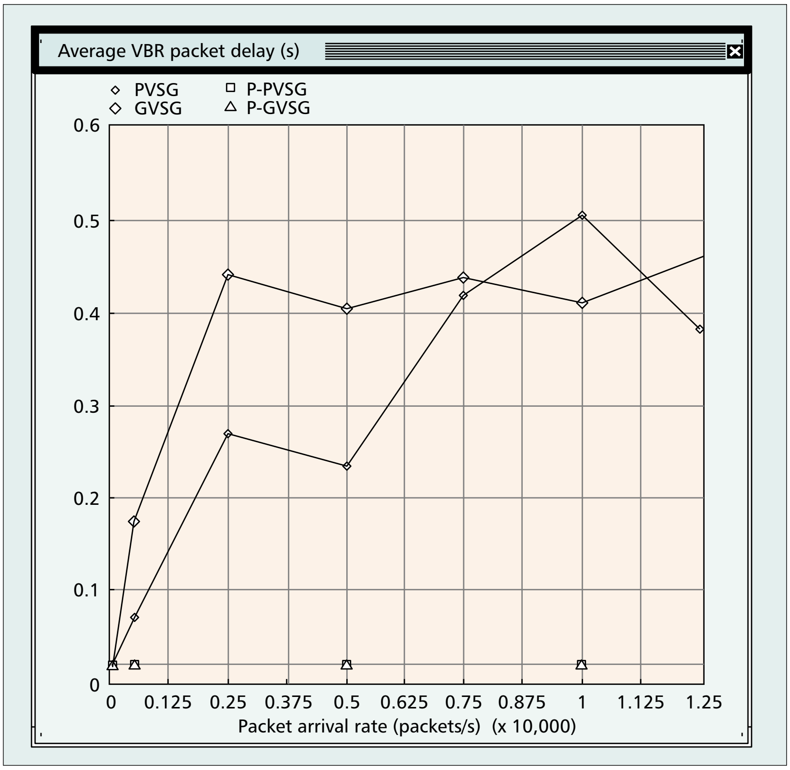

Figure 4. *VBR packet delay in single- and multiple-queue DRS schemes.*

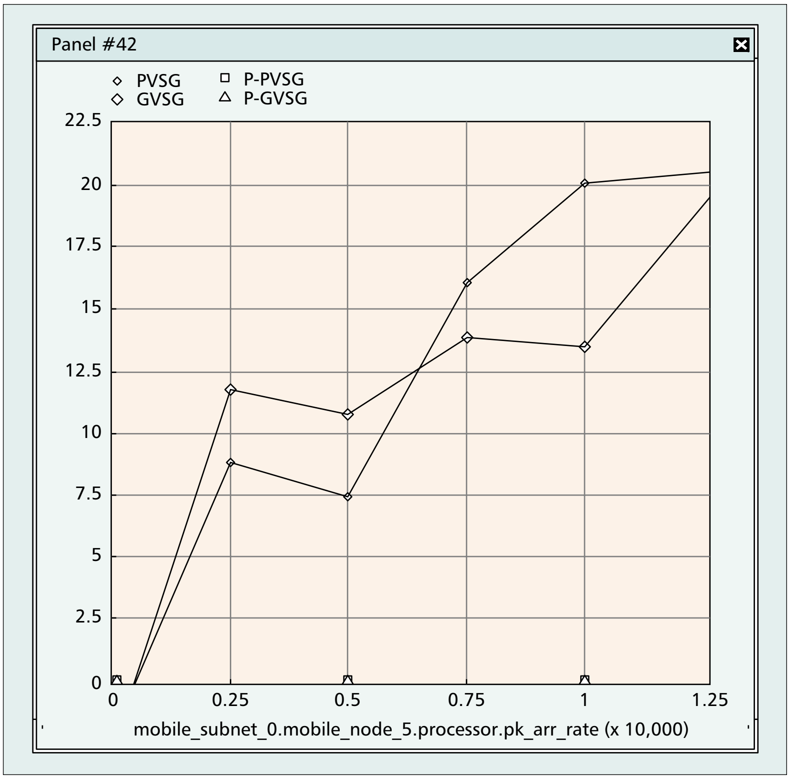

Figure 5. *VBR packet loss in single- and multiple-queue DRS schemes.*

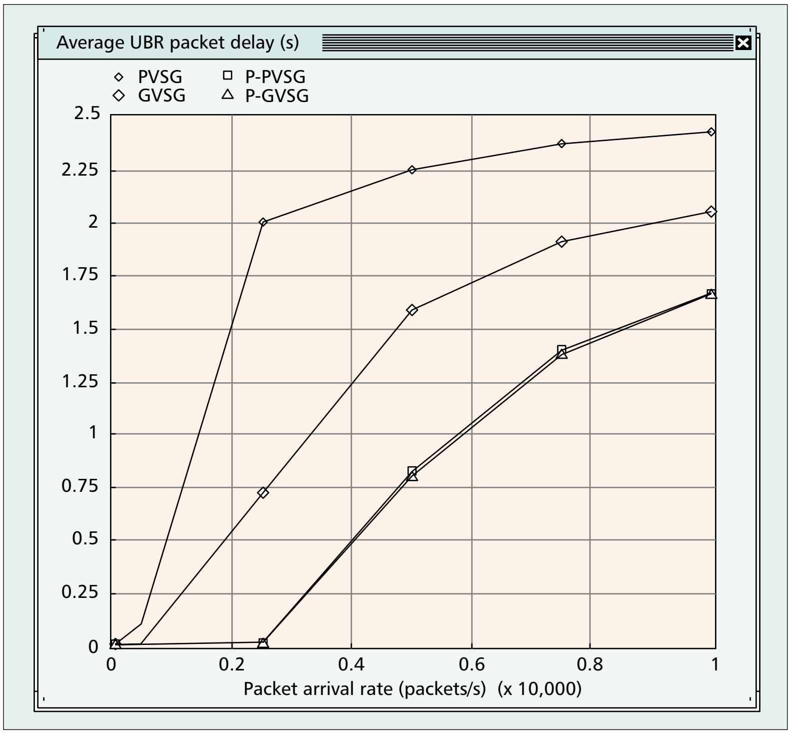

Figure 6. *UBR packet delay in single- and multiple-queue DRS schemes.*

neous capacity created for best effort services due to code hopping in VBR sources [4]. Hence, the UBR delays in the two queue structure can be compensated for to some extent by general variable DRS. Nevertheless, delay performance is not as critical for these best effort flows. They are generally transported by protocols like TCP, which ensure reliable delivery with flow control mechanisms to slow down the traffic rate when detecting packet errors and losses.

Conclusions

The entire DRS framework aims to provide QoS support for multimedia traffic over W-CDMA. It has previously been proven that DRS with variable bandwidth allocation introduces power savings, higher packet throughput, and better semantic QoS [4, 10]. In this article we extend the DRS family and examine the temporal QoS in terms of delays.

Our simulations show that the delay performance can be provisioned for guaranteed services by multiple queues. Hence, multiple-queue DRS schemes are essential for heterogeneous traffic. Variable bandwidth improves the delay performance through the extra capacity gained. The choices of single- or multiple-queue as well as the levels of variable bandwidth allocation depend on the traffic scenario. The DRS framework offers a family of algorithms that can serve a set of scenarios efficiently. Each scheme has its own cost, which would account for the pricing of each type of scenario.

Single-queue DRS schemes are efficient for a single class/type of traffic. In this case, buffer space is preserved as one queue, and overhead with classification is avoided. Still, power control and variable bandwidth allocation ensure efficient and optimal use of W-CDMA resources. PVSG and PMC schemes that apply variable rate at the packet level can readily be applied to users which require only background and interactive classes, such as e-mail and Web access applications. GVSG and GMC schemes can be devoted to users that require conversational and streaming type services, like CBR voice and compressed video.

Multiple-queue DRS algorithms provide delay provisioning at the cost of extra buffer space and complexity. However, for real-time services that cost is inevitable because of stringent QoS requirements. P-FSG can be assigned for constant rate voice (conversational) and data (background) users. P-PVSG and P-PMC will easily serve for conversational and interactive class users (e.g., simultaneous voice and Web applications). P-GVSG and P-GMC schemes are the most expensive DRS algorithms, requiring the greatest signaling overhead and queue processing. They can support all classes: conversational, streaming, background, and interactive services.

Acknowledgment

This article is based on our previously published material from 3Gwireless 2000 (ISSN No. 1529-2592) by DELSON GROUP.

References

[1] Experimental W-CDMA Description, NTT DoCoMo, Japan.

[2] 3GPP, "Technical Specification Group Services and System Aspects, QoS Concept," 3G TR 23.907, v.1.1.0.

[3] Ö. Gürbüz and H. Owen, "Dynamic Resource Scheduling for Variable QoS Traffic in W-CDMA," *Proc. IEEE ICC'99*, vol. 2, June 1999, pp. 703–7.

[4] Ö. Gürbüz and H. Owen, "A Resource Management Framework for QoS Provisioning in W-CDMA Systems," *Proc. IEEE VTC'99*, vol. 1, May 1999, pp. 407–11.

[5] M. A. Arad and A. Leon-Garcia, "Scheduled CDMA: Hybrid Multiple Access for Wireless ATM Networks," *Proc. PIMRC '96*, vol. 3, Oct. 1996, pp. 913–17.

[6] A. Sampath, P.S. Kumar, and P. Holtzman, "Power Control and Resource Management for a Multimedia CDMA Wireless System," *Proc. PIMRC '95*, vol. 1, Sept. 1995, pp. 21–25.

[7] R. Yates, "A Framework for Uplink Power Control in Cellular Radio Systems," *IEEE JSAC*, vol. 13, no. 7, Sept. 1995, pp. 1341–47.

[8] B. Unal and Y. Tanik, "Capacity Improvement by Code-Hopping in S-CDMA Systems," *Proc. ICC '98*, vol. 2, June 1998, pp. 989–93.

[9] I. Chih-Lin and D. G. Richard, "Multi-Code CDMA Wireless Personal Communications Networks," *Proc. ICC '95*, vol. 2, June 1995, pp. 1060–64.

[10] Ö. Gürbüz and H. Owen, "Dynamic Resource Scheduling Strategies for QoS in W-CDMA," *Proc. IEEE GLOBECOM '99*, vol. 1a, Dec. 1999, pp. 183–87.

[11] M. Krunz, "Bandwidth Allocation Strategies for Transporting Variable-Bit-Rate Video Traffic," *IEEE Commun. Mag.*, vol. 37, no.1, Jan. 1999, pp. 40–46.

Biographies

Özgür Gürbüz (ogurbuz@cisco.com) received her B.S. and M.S. degrees in electrical and electronics engineering from the Middle East Technical University in 1992 and 1995, respectively. She received her Ph.D. in electrical and computer engineering from the Georgia Institute of Technology in 2000. She is currently with Cisco Systems, and her research interests include wireless networks, wireless QoS, and network performance analysis.

Henry Owen (henry.owen@ece.gatech.edu) received his B.S.E.E. in 1979, M.S.E.E. in 1983, and Ph.D. in 1989 from the Georgia Institute of Technology. He is presently an associate professor in the School of Electrical and Computer Engineering at Georgia Institute of Technology. His research interests include QoS in the Internet.