Scheduling Algorithms for Downlink Rate Allocation in Heterogeneous CDMA Networks

Aikaterini C. Varsou and H. Vincent Poor

Abstract: The downlink rate scheduling problem is considered for CDMA networks with multiple users carrying packets of heterogeneous traffic (voice/audio only, bursty data only or mixed traffic), with each type having its own distinct quality of service requirements. Several rate scheduling algorithms are developed, the common factor of which is that part of the decision on which users to serve is based on a function of the deadline of their head-of-line packets. An approach of Andrews *et al.*, in which the basic Earliest-Deadline-First algorithm is studied for similar systems, is extended to result in better performance by considering a more efficient power usage and by allowing service of more than one user per timeslot if the power resources permit it. Finally, the performance of the proposed schemes is compared through simulations.

Index Terms: Rate allocation, CDMA networks, heterogeneous traffic, quality-of-service requirements, deadline.

I. INTRODUCTION

An inevitable issue in the design of wireless networks is that of energy conservation, making power control and power allocation two important factors that must be addressed in such design. Since power and rate are essentially equivalent notions (as one can be derived from the other), power considerations can be addressed by examining the problem of rate allocation.

Emerging wireless networks, such as third-generation cellular networks, anticipate a combination of data and voice/audio services in the same wireless system. In such networks, different traffic classes can ask for diverse Quality of Service (QoS) characteristics, thus provoking the study of rate scheduling techniques for heterogeneous classes of traffic with diverse QoS needs.

A scheduling principle establishes the order in which packets will be transmitted or equivalently the way in which bandwidth will be shared among different users, and therefore influences directly the QoS guarantees. Specifically, given the traffic descriptions, such as burstiness and throughput, and the performance requirements (delay, packet loss probability) of a user and its packet stream, the network must determine how much bandwidth and rate should be allocated to this user. Several such

allocation, or rate scheduling, schemes, can be found in the literature. This paper is concerned with the allocation of rate among the users of a Code-Division Multiple Access (CDMA) network with heterogeneous traffic.

Supporting the heterogeneous delay requirement of applications with extensively varying characteristics requires scheduling algorithms that are more sophisticated than "First In First Out", which transmits packets in the same order as they were received by the scheduler. Recent work in this area is found in [1], in which the authors consider CDMA systems with delay-tolerant bursty data users. They examine the problem of scheduling the rate of CDMA data users on the downlink such that certain QoS requirements are met. They evaluate the performance of several different scheduling schemes with the main conclusion being that it is optimal for the base station to transmit to only one user at a time by operating either at zero or at full power. As a result, their best rate scheduling scheme is the "Earliest-Deadline-First" (EDF) algorithm (or its weighted version WEDF). Similar versions of EDF have appeared often in the literature under the names "Earliest Due Date" (EDD), "Stochastic Earliest Deadline" (SED), etc., and have been shown to perform better than other algorithms for similar systems. According to EDF, at each timeslot the user carrying in its head-of-line the packet with the earliest deadline is served, thereby assigning all the power resources to this single user. Hence only one bursty-data user can be served per timeslot.

The need for data and voice/audio services in the same wireless system, motivates us to consider downlink rate allocation for such heterogeneous cases in this paper. We do so by expanding the model of [1] to that in [2], which includes three different cases: Users carrying data packets only, voice/audio (constant-bit-rate, CBR) only, or mixed traffic. We consider an algorithm that we call "Powered Earliest-Deadline-First" (PEDF) which, like EDF, finds at the beginning of a timeslot the user with the closest deadline for the head packet and assigns all the power to it. However, if after serving that packet (user) there is still some power remaining, it serves the user with the *next* earliest deadline. The algorithm then continues serving the complete head packets of the chosen users according to the available power resources. If a point is reached where there is not enough power left to serve a chosen user, it serves part of the corresponding packet and moves to the next timeslot. Hence, this approach differs substantially from the one of [1]. Firstly, more complex, heterogeneous traffic input is considered with CBR voice and bursty data traffic having differing delay requirements. Secondly, PEDF is more power efficient because of the new idea of serving more than one user per timeslot, power permitted.

We further consider three additional new algorithms for rate

scheduling of heterogeneous traffic. In the first one (see also [3]), the head packet of each user is assigned a pseudoprobability, which is a function of its deadline. The pseudoprobabilities are then normalized by the packet lengths; and the packet having the maximal normalized quantity is served first. If there is power remaining in the system after the packet is served, the procedure is repeated with the *next* maximum normalized pseudoprobability packet, and so on. Even though this scheme is similar in philosophy to PEDF, the serving criterion is different, thus allowing dramatic improvement in performance, and consequently more users to be supported in the system, as we will see in the sequel.

The second and third new algorithms are better versions of PEDF and the above pseudoprobability approach, with the additional feature of fairness queueing. "Fairness queueing" is considered in [4] in the context of an ATM switch. This technique is implemented by dividing the system bandwidth equally among all users. If a user cannot satisfy its needs with the given bandwidth allocation, it is characterized as bottlenecked. Otherwise it is marked as satisfied. It may be the case that some satisfied users will not need all of their allocated bandwidth. In this situation excess bandwidth from the satisfied users can be used to ease up the bottlenecked ones. Again this excess is divided equally among all bottlenecked users. This approach guarantees max-min fairness. Since the notions of available bandwidth and available power are equivalent, we can combine the PEDF and normalized pseudoprobability schemes with the idea of fairness queueing. We do so here and show through simulations that these combinations can lead to better performance when compared to the PEDF and pseudoprobability algorithms.

This paper is organized as follows. In Section II, we introduce the main model, while in Section III, we present sketches of the new algorithms. A mathematical analysis is presented in Section IV. In Section V, we provide simulation examples to demonstrate the performance of the algorithms proposed in this paper. Section VI contains our conclusions.

II. MAIN MODEL

As in [1], we concentrate on a single-cell model in which the total available transmission power is constant in time and normalized to unity. There are K active users in the system, and the downlink is modelled as a multiple server queue. Traffic is generated for every user for a specific time period. The packets generated can be either CBR packets only, or bursty data traffic packets only, or mixed traffic ones. The bursty data case is generated by an on-off fluid source.

Let us concentrate on a specific timeslot and denote by W_i the waiting time for the head packet of user i and by T_i the delay threshold, which is a predefined value that will be different for each kind of traffic. At each timeslot the queues of all users are examined to decide which clients to serve. It is assumed that every user has to meet a delay QoS requirement regardless of traffic type; i.e., we must have

$$P(W_i \leq T_i) \geq \delta_i, \tag{1}$$

where δ_i is the minimum probability with which the constraint must be met. Since the situation where we have different values

of δ_i for each user is quite complicated, we will consider the simpler case where the δ_i are the same for all users.

The Earliest Deadline First (EDF) algorithm, which is the main rate scheduling scheme considered in [1], assigns all available power to the user for which the deadline of the packet at the head of the queue is earliest, i.e., for which the difference $T_i - W_i$ is smallest. Therefore only one user is served in each time slot.

Powered Earliest Deadline First (PEDF), proposed in [2], looks at the beginning of a timeslot for the user whose head packet has the earliest deadline and assigns all the power to it. Then, if after serving that user's packet there is still power left, the head packet of the user with the *next* earliest deadline is served. This process continues with the full head packets of chosen users according to the power resources. If a point is reached where there is not enough power left to serve an entire packet, part of the next packet is served and attention switches to the next timeslot. Therefore, this approach serves more than one user when enough power to do so is available. A disadvantage of choosing multiple users in this way, is that multiuser interference arises among them. But, this disadvantage is often outweighed by the tendency to make use of more of the available power. Thus PEDF can outperform EDF, as will be verified below through simulations.

III. SKETCHES OF THE ALGORITHMS

A. HOLPRO

All the algorithms presented in this paper refer to the complex heterogeneous traffic input discussed above. In this section we develop a modification of PEDF, which we call "Head Of Line PseudopRObability" (or HOLPRO), which serves more than one user at a time according to the same notion used in PEDF. HOLPRO however uses a different serving criterion than PEDF. In particular, in HOLPRO each user is assigned a pseudoprobability p_i and the user with the maximal normalized pseudoprobability is chosen as long as power resources remain.

More specifically, in each timeslot each user's head packet is assigned a pseudoprobability p_i, given by

$$p_i = \frac{\frac{1}{(T_i - W_i)^3}}{\sum_{k \in H} \frac{1}{(T_k - W_k)^3}}, \tag{2}$$

where W_i and T_i are the waiting time and delay threshold as before, and where H contains the indices of all the users that coexist in the given timeslot. Although $\sum_i p_i = 1$, the p_i's are not necessarily non-negative, and hence the term "pseudoprobability". The pseudoprobabilities p_i are normalized by the size of the head packet, which is denoted by l_i, and the user with the maximum normalized probability is served first. Clearly an important feature of HOLPRO is that it gives preference to shorter packets and lower priority to packets that have already missed their due date. Note that when packets of different users are served, multiuser interference may result. However, different packets of the same user do not interfere with one another. The algorithm is summarized as follows.

Algorithm 1: HOLPRO

  1. *Given T_i and W_i for each user, compute p_i.*
  2. *For each time slot in the time window.*
     (a) *Start with full transmission power ($P = 1$).*
     (b) *While power $P > 0$.*
        i. *Determine the user with $\max_i \frac{p_i}{l_i}$.*
        ii. *Compute the power P' this user needs to transmit its packet in this time slot.*
        iii. *Compute the marginal power P'' required by all previous packets that were to be transmitted during this slot, to account for the interference due to this new packet i.*
        iv. *If the power allocated to the newly chosen packet is not enough to serve it completely (i.e., $P' > P - P''$), serve as much of it as possible in this time slot, set $P = 0$, and move to the next time slot.*
        v. *Otherwise ($P' \leq P - P''$).*
           • *serve the entire packet, and*
           • *compute the power remaining in the system ($P = P - P' - P''$).*

B. PEDFF

We also consider allocation according to a combination of the PEDF algorithm with the fairness queueing technique of [4]. In particular, as long as enough power resources remain, the closest deadline user is served if we can serve its complete head packet and then the remaining power is divided among the other active users. Again more than one user per timeslot is served, but the system is now "fairer" since all the users other than the one with the earliest deadline packet receive their equal power share. The algorithm in this case, which we call "Powered Earliest Deadline First Fair" (PEDFF) is thus:

Algorithm 2: PEDFF

  1. *For each time slot in the time window.*
     (a) *Start with full transmission power ($P = 1$).*
     (b) *While power $P > 0$.*
        i. *Given W_i, T_i for each user, find the one with the earliest deadline $T_i - W_i$.*
        ii. *Compute the power P' needed to transmit its packet in this time slot.*
        iii. *Compute the marginal power P'' required by all previous packets that were to be transmitted during this slot, to account for the interference due to this new packet i.*
        iv. *If the power allocated to the newly chosen packet is not enough to serve it completely (i.e., $P' > P - P''$), serve as much of it as possible in this time slot, set $P = 0$ and move to the next time slot.*
        v. *Otherwise ($P' \leq P - P''$).*
           • *serve the entire packet,*
           • *compute the power remaining in the system ($P = P - P' - P''$),*
           • *determine how many users have packets in the queue now and divide the remaining power equally among them,*

• *for each one of these users, serve as much as possible of its head packet according to the power allocated, keeping in mind that the interferences may need to be updated if there are packets served in the same time interval from the same user, and*
• *update the power remaining in the system.*

C. HOLPF

As we already mentioned, it will become clear through simulations that HOLPRO outperforms PEDF. Also PEDFF is fairer than PEDF and turns out to have better performance. This suggests a combination of the fairness queueing notion of PEDFF with the idea of the HOLPRO algorithm. This produces the following scheduling scheme, which we call "Head Of Line Pseudoprobability Fair" (HOLPF):

Algorithm 3: HOLPF

  1. *Given T_i and W_i for each user, compute p_i.*
  2. *For each time slot in the time window.*
     (a) *Start with full transmission power ($P = 1$).*
     (b) *While power $P > 0$.*
        i. *Determine the user with $\max_i \frac{p_i}{l_i}$.*
        ii. *Compute the power P' this user needs to transmit its packet in this time slot.*
        iii. *Compute the marginal power P'' required by all previous packets that were to be transmitted during this slot, to account for the interference due to this new packet i.*
        iv. *If the power allocated to this user is not enough to serve its packet completely ($P' > P - P''$), serve as much as possible in this time slot, set $P = 0$, and move to the next time slot.*
        v. *Otherwise ($P' \leq P - P''$).*
           • *serve the entire packet,*
           • *compute the power remaining in the system ($P = P - P' - P''$),*
           • *determine how many users have packets in the queue now and divide the remaining power equally among them,*
           • *for each one of these users, serve as much as possible of its head packet according to the power allocated to it, keeping in mind that the interferences may need to be updated if there are packets served in the same time interval from the same user, and*
           • *update the power remaining in the system.*

IV. MATHEMATICAL ANALYSIS

In analyzing and simulating the performance of the above rate-allocation schemes, we will consider the performance in a central cell surrounded by interfering cells. Using the models of [5], [6], we have the following relation (see also [2]) for the power $\Gamma_{j,i}$ that mobile i receives from the base of the j-th cell:

$$\Gamma_{j,i} = P_T \left(\frac{d_{j,i}}{d_0}\right)^{-4} 10^{0.1\sigma_s Y_j}, \tag{3}$$

where P_T is the total power that each cell transmits on the downlink, $d_{j,i}$ is the distance between base station j and mobile i, d_0 is a reference distance, σ_s is the deviation of log-normal shadow fading, and Y_j is a zero-mean unit variance Gaussian random variable. The i-th mobile will then choose the cell corresponding to the highest power.

In the simulations, users are placed in space and all $\Gamma_{j,i}$'s are computed. If the maximum power $\max_j \Gamma_{j,i}$ is attained for j corresponding to the central cell, then user i is placed in that cell. Otherwise, the user is regenerated, so that all our users are placed in the central cell.

Assuming that the total cell site power is divided between the users/subscribers and a pilot signal, let β denote the fraction of the power devoted to the traffic channels (i.e., $1 - \beta$ is devoted to the pilot) and let P_i denote the fraction of this devoted to user i. Define the following parameters:

E_b/N_o: Bit energy to noise ratio required at the receiver.
G : Spreading gain.
η: Noise power in the downlink receivers.
α: Probability that each user is ON (it is assumed that users turn on and off independent of one another).
ρ: An "orthogonality loss" factor, where $\rho = 0$ corresponds to perfect orthogonality and $\rho = 1$ corresponds to random codes.

Then, according to [2] and [6], the E_b/N_o's and rates for user i are given by

$$(E_b/N_o)_i$$
$$= \frac{\alpha\beta\Gamma_{0,i}P_iG}{\rho\Gamma_{0,i}(1 - \beta + \alpha\beta\sum_{\forall j \neq i} P_j) + (1 - \beta + \alpha\beta)(\sum_{n\neq 0}\Gamma_{n,i} + \eta)},$$
$$(4)$$

and

$$R_i = \min\{1, \bar{R}_i\}, \tag{5}$$

where $\bar{R}_i$ has the form of (6) at the bottom of this page and for simplicity the interference $\sum_{\forall j \neq i} P_j$ is denoted by I_i.

Here the rates are assumed to be normalized by the system bandwidth (i.e., the units are bits/sec/Hz), which means that we must multiply the above rates by the system bandwidth B to get rates in bits per second (bps).

As a reminder, the technique used in [2] and [6] is similar to the one in [5]. In [5] however, the authors generate a simulated histogram for the fraction $\frac{\sum_{n\neq 0}\Gamma_{n,i}}{\Gamma_{0,i}}$ and then the outage is calculated through a Chernoff bound, which accounts for voice activity. As Chernoff bounds tend to be loose, we have used a Monte Carlo approach instead. To include voice traffic in the

Monte Carlo simulaition, the formulae in [5] have been modified so that the $\Gamma_{j,i}$'s contain a pilot fraction without activity and a traffic fraction with activity. As a result, (4) was derived.

Furthermore, we see from (6) that the rate needed for a user's packet is a function of the power needed for that packet and of the interference. Some care must be taken in computing interference when more than one packet from the same user is served in the same timeslot. When a given user is having its first packet served in a given timeslot, its rate is a function of the needed power, P_{needed}, and the interference power. That interfering power is roughly calculated to be equal to $1 - P_{needed}$. If in the same timeslot, the given user subsequently has a further packet served, then this packet's power was included in the previous calculations of the interfering power for the first packet, which is inaccurate since packets belonging to the same user do not interfere with each other.

Hence we can solve the equation for the required power by distinguishing the following two cases.

A. Case 1: The Packet to Serve Belongs to a User not Served in the Same Timeslot

In this case, it is easy to conclude (as in [2] and [3]) that $I_i = 1 - P_i$ and therefore

$$P_i = \frac{R_i(E_b/N_o)_i(1 - \beta + \alpha\beta)(\rho\Gamma_{0,i} + \Sigma_i)}{\alpha\beta\Gamma_{0,i}(B + \rho R_i(E_b/N_o)_i)}, \tag{7}$$

where, in order to simplify notation, we have denoted $\sum_{n\neq 0}\Gamma_{n,i}$ by Σ_i.

B. Case 2: The Packet to Serve Belongs to a User Already Served in the Same Timeslot

Suppose that n packets from user i were previously served in this timeslot, with required rates $R_{i,1}, \cdots, R_{i,n}$. The packet in question is then the $(n + 1)$-st and its interference is $I_{i,n+1} = 1 - P_{i,n+1} - P_{i,1} - \cdots - P_{i,n}$, where the $P_{i,j}$'s are the quantities we are trying to compute, i.e, $P_{i,j}$ is the updated power of the j-th packet of user i. Then, denoting $(E_b/N_o)_i$ by E_i, we can write

$$\alpha\beta\Gamma_{0,i}(B + \rho E_i R_{i,n+1})P_{i,n+1}$$
$$+ \alpha\beta\Gamma\rho E_i \ R_{i,n+1}(P_{i,1} + \cdots + P_{i,n})$$
$$= E_i R_{i,n+1}(1 - \beta + \alpha\beta)(\rho\Gamma_{0,i} + \Sigma_i). \tag{8}$$

By generating similar equations for the other packets we arrive at the system of (9) given at the bottom of next page for the desired powers, where $C = E_i(1 - \beta + \alpha\beta)(\rho\Gamma_{0,i} + \Sigma_i)$.

The solution of this system is the vector of desired powers that are used to calculate the marginal power P'' in the algorithms.

$$\bar{R}_i = \frac{1}{G} = \frac{\alpha\beta \ \Gamma_{0,i}P_i}{(E_b/N_o)_i \ (\rho\Gamma_{0,i}(1 - \beta + \alpha\beta I_i) + (1 - \beta + \alpha\beta)(\sum_{n\neq 0}\Gamma_{n,i} + \eta))}, \tag{6}$$
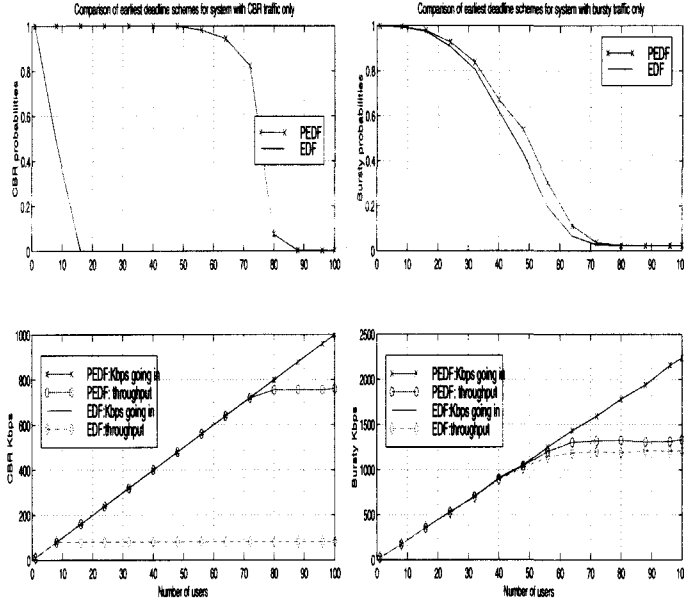
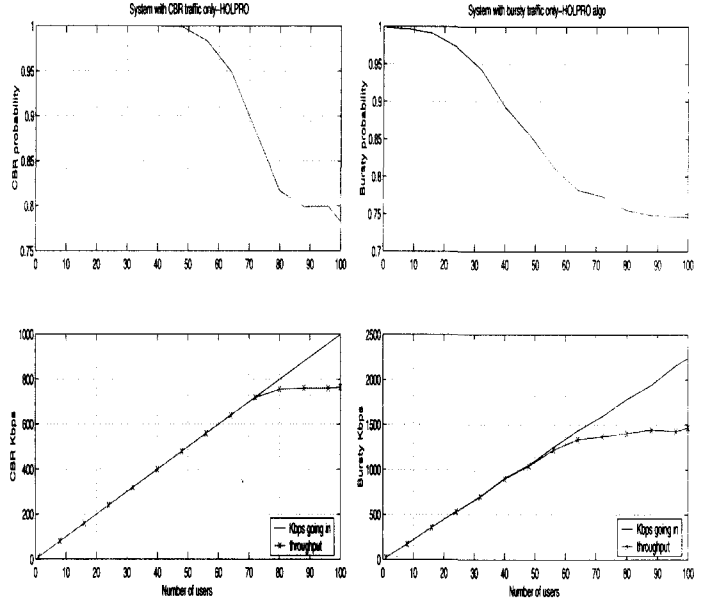Fig. 1. PEDF vs. EDF – (a) CBR, (b) Bursty.

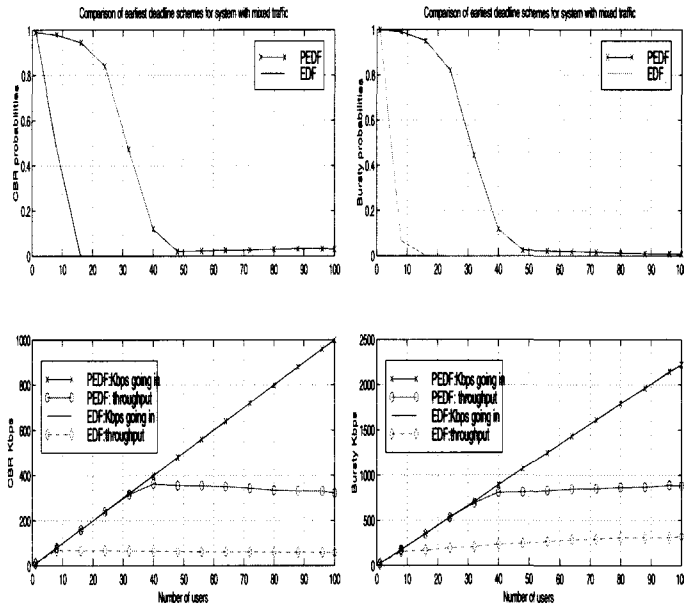Fig. 3. HOLPRO –(a) CBR, (b) Bursty.

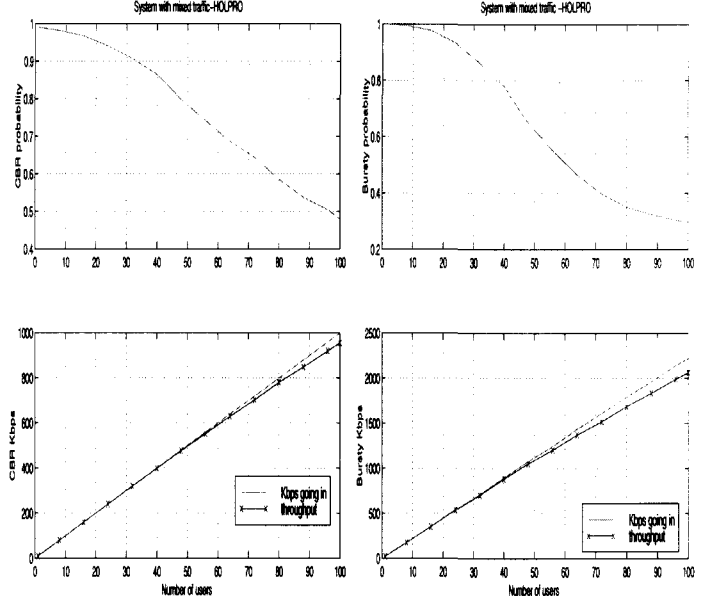Fig. 2. PEDF vs. EDF in mixed traffic – (a) CBR, (b) Bursty.

Fig. 4. HOLPRO in mixed traffic – (a) CBR, (b) Bursty.

V. SIMULATIONS

To simulate the algorithms described above, we consider a configuration of 18 cells ordered in two rings around a central cell, in which all the users of interest are placed. We assume the following typical values: E_b/N_o is 5.0 dB, β is 0.8, α is 3/8, the shadow fading has standard deviation 8, and the bandwidth

of the system is 5 MHz (the value proposed for cdma 2000). Constant Bit Rate (CBR) packets are assumed to contain 100 bits and have 10 msec interarrival time, while bursty packets are assumed to have exponentially distributed size and interarrival times with means 10 Kbits and 100 msec analogously. The ON-OFF durations were chosen to be exponential with means 93 msec and 907 msec, the same values as in [1] and [2], for the

$$
\begin{bmatrix}
\alpha\beta\Gamma_{0,i}(B + \rho E_i R_{i,1}) & \alpha\beta\Gamma_{0,i}\rho E_i R_{i,1} & \cdots & \alpha\beta\Gamma_{0,i}\rho E_i R_{i,1} \\
\alpha\beta\Gamma_{0,i}\rho E_i R_{i,2} & \alpha\beta\Gamma_{0,i}(B + \rho E_i R_{i,2}) & \cdots & \alpha\beta\Gamma_{0,i}\rho E_i R_{i,2} \\
\vdots & \vdots & & \vdots \\
\alpha\beta\Gamma_{0,i}\rho E_i R_{i,n+1} & \alpha\beta\Gamma_{0,i}\rho E_i R_{i,n+1} & \cdots & \alpha\beta\Gamma_{0,i}(B + \rho E_i R_{i,n+1})
\end{bmatrix}
\begin{bmatrix}
P_{i,1} \\
P_{i,2} \\
\vdots \\
P_{i,n+1}
\end{bmatrix}
= C
\begin{bmatrix}
R_{i,1} \\
R_{i,2} \\
\vdots \\
R_{i,n+1}
\end{bmatrix}, \quad (9)
$$

Table 1. Standard deviations for the graphs of the main algorithms.

Algorithm	Scenario	Probability plot		Kbps plot	
		Min. std. dev	Max. std. dev	Min. std. dev.	Max. std. dev
PEDF	CBR-only	0.0007	0.0182	5.2456	23.857
	Bursty-only	0.0016	0.0289	21.787	39.694
	Mixed: CBR portion	0.0011	0.052	11.0153	26.7214
	Mixed: Bursty portion	0.0019	0.034	24.3846	42.242
HOLPRO	CBR-only	0.0017	0.0199	13.7477	26.5013
	Bursty-only	0.0028	0.0294	24.8263	47.1514
	Mixed: CBR portion	0.0024	0.0235	14.8277	29.356
	Mixed: Bursty portion	0.0031	0.0697	25.7461	50.184
PEDFF	CBR-only	0.0009	0.0145	9.2175	22.389
	Bursty-only	0.003	0.0415	27.6014	44.5651
	Mixed: CBR portion	0.0028	0.046	10.8031	19.0246
	Mixed: Bursty portion	0.0042	0.0274	30.1403	48.576
HOLPF	CBR-only	0.0026	0.0219	12.427	21.505
	Bursty-only	0.0049	0.0313	32.1650	51.6803
	Mixed: CBR portion	0.0051	0.0246	14.5967	23.3454
	Mixed: Bursty portion	0.0067	0.0839	35.6156	53.242
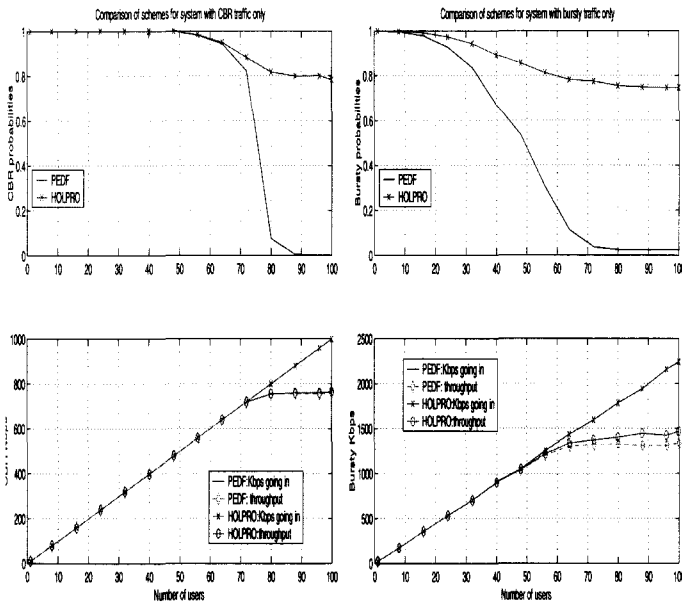
Fig. 5. HOLPRO vs. PEDF – (a) CBR, (b) Bursty.
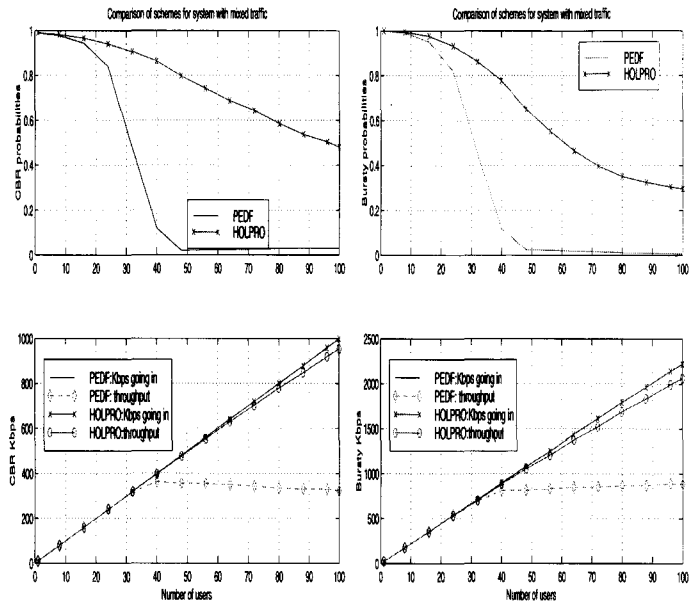
Fig. 6. HOLPRO vs. PEDF in mixed traffic – (a) CBR, (b) Bursty.

purpose of comparison.

The simulations were run for many different values of CBR and bursty delays for a time period of 50 sec. The graphs considered here depict system behavior over a period of 5 sec, where the maximum delay for the CBR packets is 2 msec and that for the bursty packets is 0.1 sec. These values for the delays were carefully chosen in [2]. It is known that bursty data can tolerate higher delays than CBR traffic, which should be served in a couple of milliseconds at most. The above choice arises from a combination of the fact that the traffic entering the system approximates that exiting for the cases of CBR-only and data-only traffic, and from the fact that the resulting probabilities should not be trivial (i.e., equal to one all the time) but rather should exhibit a smooth decrease.

Our plots show the successful probabilities of meeting the QoS constraints $P(W_i \leq T_i)$ as well as the Kbps entering and exiting the system versus the number of users in the system, for all different types of traffic. Hence CBR (respectively Bursty) probability will denote the probability that the CBR (resp. Bursty) packets meet the desired QoS. In the case of mixed traffic, the term CBR (resp. Bursty) will refer to the CBR (resp. Bursty) portion of the traffic. Since all the numbers in the plots are produced by simulation, we will also give the minimum and maximum standard deviation for each algorithm. Table 1 shows the values of the minimum and maximum standard deviation for each different scenario of each algorithm that we plot.

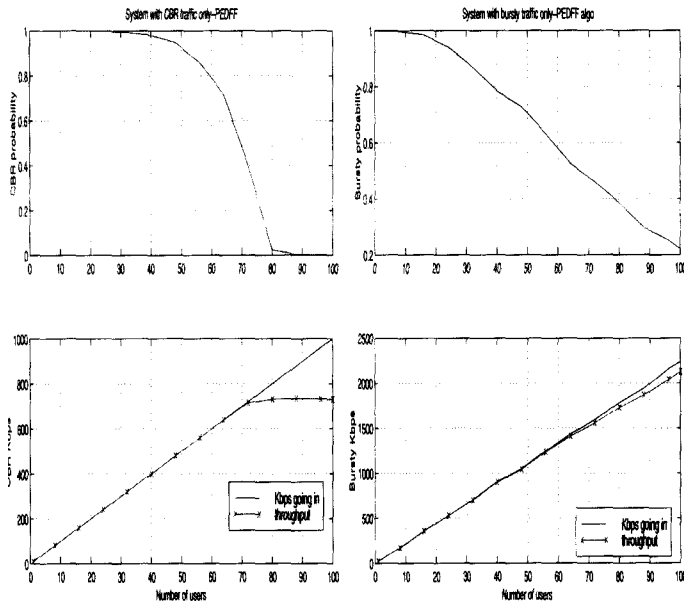In Figs. 1 and 2, we reproduce a comparison of EDF and

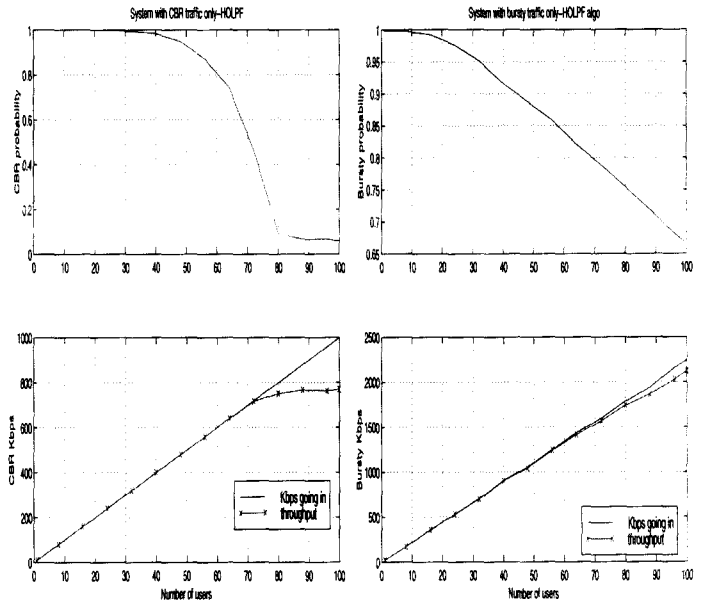Fig. 7. PEDFF – (a) CBR, (b) Bursty.

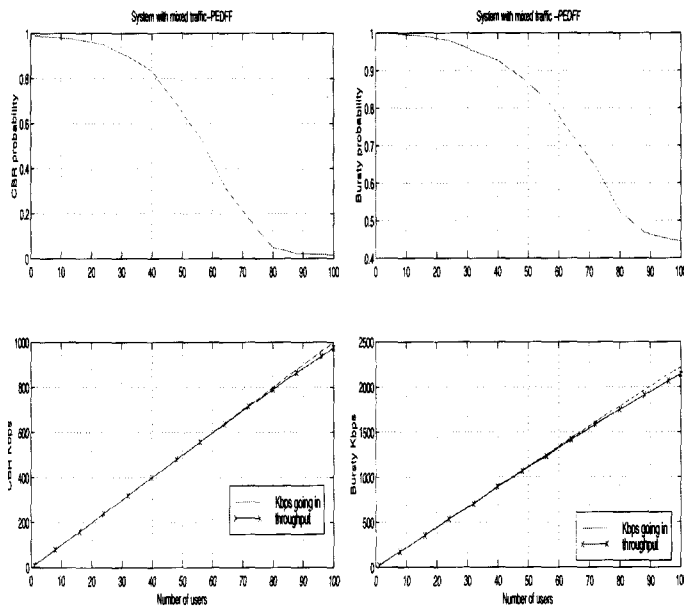Fig. 9. HOLPF – (a) CBR, (b) Bursty.

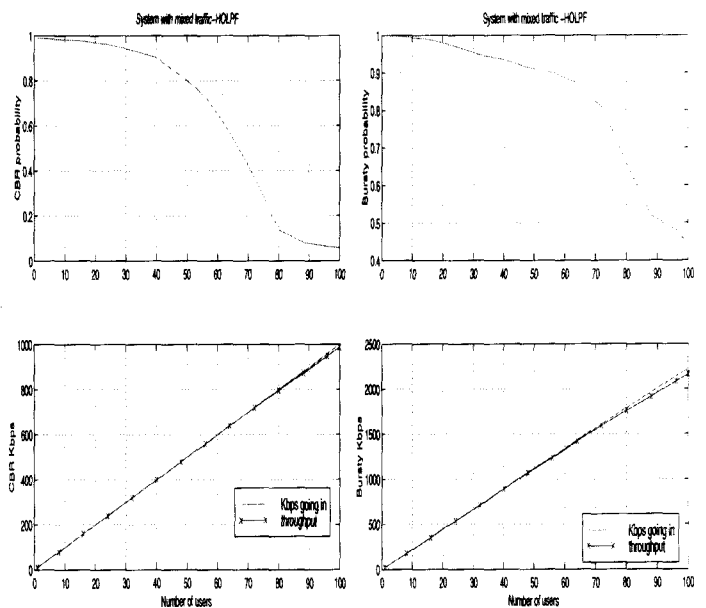Fig. 8. PEDFF in mixed traffic – (a) CBR, (b) Bursty.

Fig. 10. HOLPF in mixed traffic – (a) CBR, (b) Bursty.

PEDF from [2] for the purpose of comparison. In the case where the waiting times of the users exceed their delay constraints, the algorithm in [2] serves the packets anyway; but of course those packets served do not contribute to the calculation of the success probabilities, since the constraints are clearly violated. This is the main reason why we also include plots of the Kbits entering and exiting the system. As we can clearly observe from these figures, PEDF performs almost as well as EDF in the cases of data-only packets, while it outperforms it in all other cases.

Figs. 3 and 4 depict results of a simulation of the HOLPRO algorithm [1], while Figs. 5 and 6 compare PEDF to HOLPRO.

[1] The success probabilities should decrease as the number of the users in the system increases. Thus, in the case of CBR-only graph, we believe that two of the simulation points for more than 80 users are slightly displaced due to

Clearly these results indicate that HOLPRO outperforms PEDF. From the EDF-PEDF-HOLPRO comparisons we can see that if, for example, we want more than 90% of users to satisfy QoS, then for the given data EDF supports only one user with CBR-only traffic, 24 with bursty-only traffic and one user with mixed traffic, while PEDF gives around 64 users for CBR, 24 for bursty and 16 for the mixed case. On the other hand, HOLPRO can support 70 users for CBR, 38 for bursty, and approximately 30 for mixed traffic.

Figs. 7(a) and (b) show PEDFF's performance for a system with CBR-only or bursty-only traffic, while Fig. 8 depicts PEDFF's performance for the mixed-traffic case. The analogous

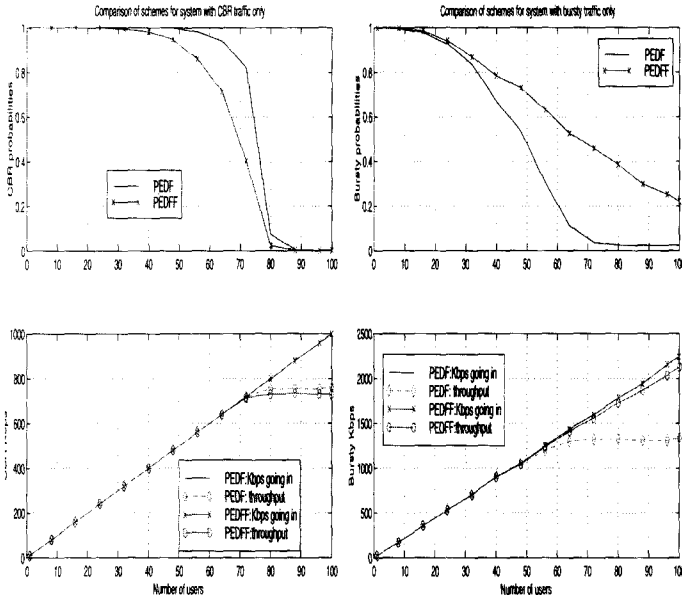computational error caused by the high complexity of the algorithm at that point.

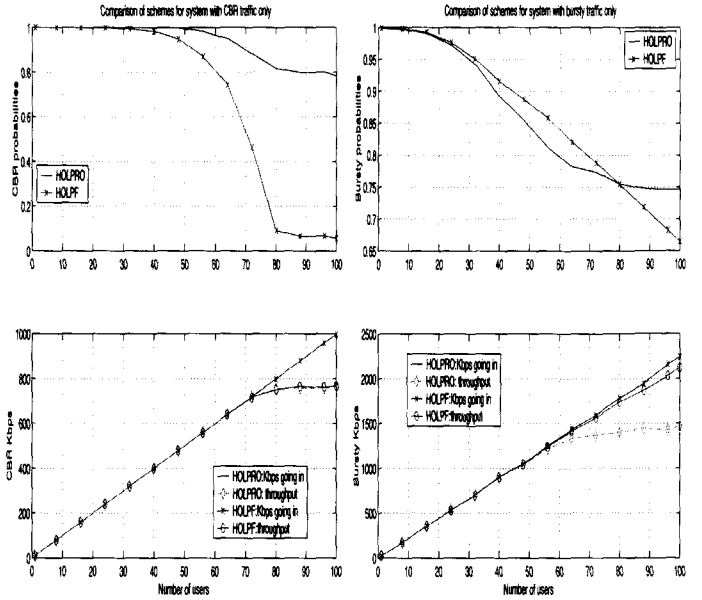Fig. 11. PEDF vs. PEDFF – (a) CBR, (b) Bursty.
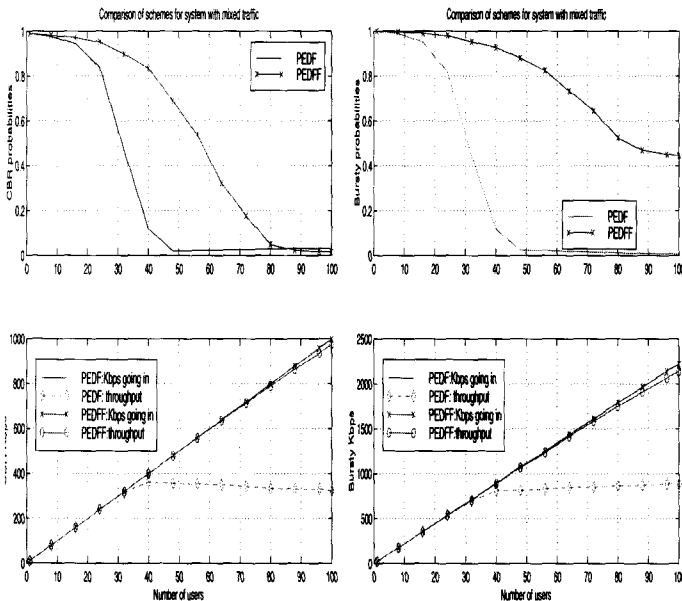

Fig. 13. HOLPRO vs. HOLPF – (a) CBR, (b) Bursty.

Fig. 12. PEDF vs. PEDFF in mixed traffic – (a) CBR, (b) Bursty.
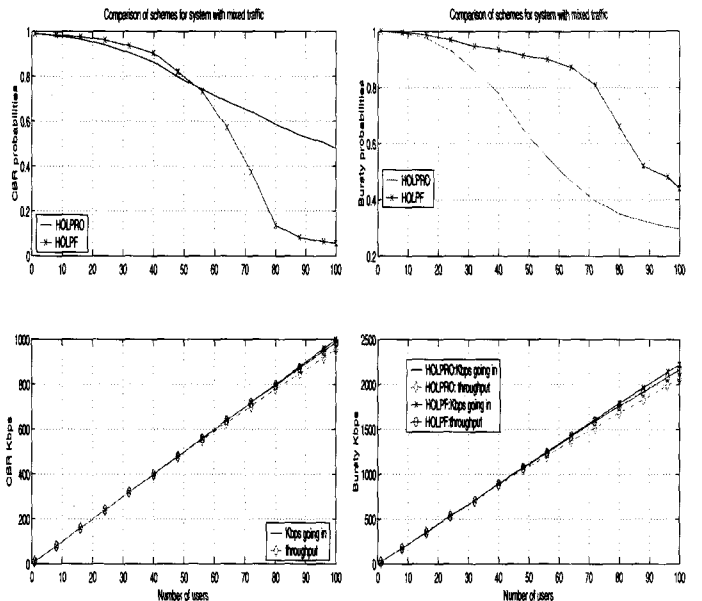

Fig. 14. HOLPRO vs. HOLPF in mixed traffic – (a) CBR, (b) Bursty.

results for HOLPF are shown in Figs. 9 and 10. In Figs. 11 and 12 PEDFF is compared to PEDF, while in Figs. 13 and 14 HOL-PRO is compared to HOLPF. The comparison between PEDFF and HOLPF is found in Figs. 15 and 16.

From these results, we can see that HOLPRO and PEDF outperform HOLPF and PEDFF in the case of CBR-only traffic. We infer that, in this case, the serving criteria influence the performance significantly. It seems that it is more efficient for the remaining users to be served one by one according to the earliest deadline or maximum normalized pseudoprobability criteria rather than by using small equal portions of the remaining power, especially since the number of coexisting CBR users is large and therefore equally-divided power portions are small. The same phenomenon appears also to be behind the CBR-in-

mixed-traffic graph for HOLPF for more than 56 users in the system. It is important to notice, though, that HOLPF and PEDFF outperform their peers in all other cases. Thus, for the cases considered here, we can see that if we want more than 90% of users to satisfy QoS, PEDF supports only 64 users with CBR-only traffic, 24 with bursty-only traffic and 16 users with mixed traffic, while HOLPRO gives around 70 users for CBR, 38 for bursty and 30 for the mixed case. On the other hand, PEDFF can support 52 users for CBR, 30 for bursty and approximately 30 for mixed traffic, while HOLPF can hold 52 for CBR, 44 for bursty, 40 for the mixed case. Also we can clearly see that again HOLPF outperforms PEDFF.

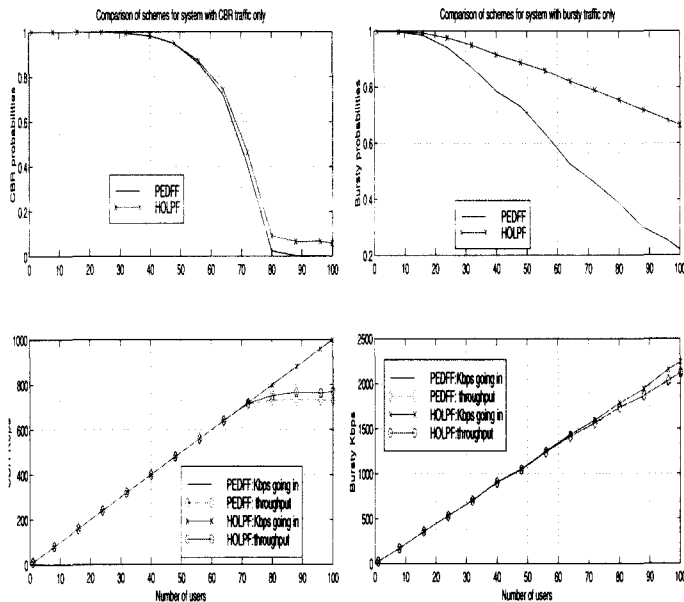These comparisons for 90% success probability are summarized in Table 2.

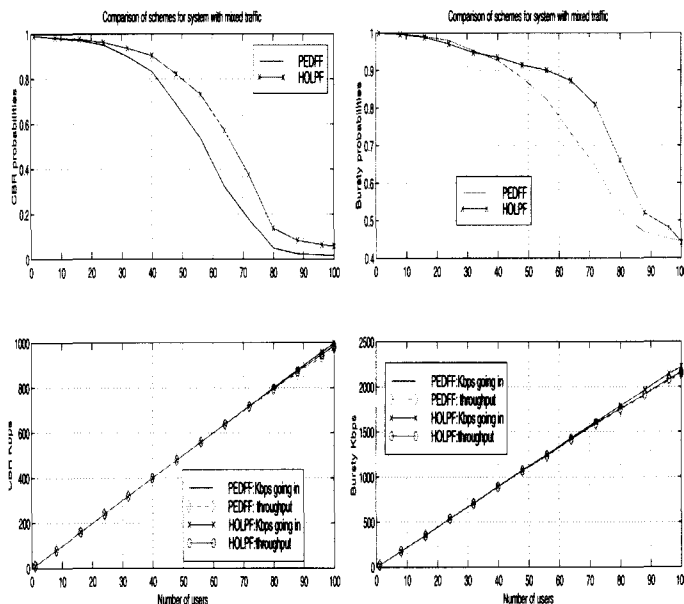Fig. 15. PEDFF vs. HOLPF – (a) CBR, (b) Bursty.

Fig. 16. PEDFF vs. HOLPF in mixed traffic – (a) CBR, (b) Bursty.

VI. CONCLUSIONS

In this paper, we have developed three new rate allocation algorithms for the downlink of CDMA multiuser systems with heterogeneous traffic. We have simulated these new schemes and compared them to each other and to existing algorithms, with the conclusion that there are many cases in which they can significantly improve the system performance. In particular, HOLPRO outperforms PEDF under all types of traffic. PEDFF improves PEDF's performance significantly for the cases of bursty-traffic and mixed-traffic. HOLPF outperforms HOLPRO for up to 80 users in a bursty-only traffic system. In the case of mixed traffic, the HOLPF bursty portion of traffic outperforms its HOLPRO peer for as high as 100 users, while the HOLPF

Table 2. Numerical comparison of the number of users supportable within 90% probability of QoS for the different allocation schemes.

Scheme	Traffic type		
	CBR only	Bursty only	Mixed
EDF	1	24	1
PEDF	64	24	16
HOLPRO	70	38	30
PEDFF	52	30	30
HOLPF	52	44	40

CBR portion outperforms its HOLPRO peer for as many as 56 users. Finally PEDFF and HOLPF give similar results for CBR-only traffic, while HOLPF outperforms PEDFF for all other cases.

This work can be extended in many interesting directions. An issue which is worth investigating is how the results are affected by the use of interference suppression in the receivers. Finally, the general case of having more than two source classes with different quality of service requirements needs to be considered.

ACKNOWLEDGMENTS

REFERENCES

[1] M. Andrews et al., "Data rate scheduling algorithms and capacity estimates for the CDMA forward link," Bell Labs Technol. Memo, Sept. 1999.

[2] A. C. Varsou, H. C. Huang, and L. Mailaender, "Rate scheduling for the downlink of CDMA mixed traffic networks," in Proc. 2000 IEEE Wireless Commun. Networking Conf., Chicago, IL, Sept. 23–27, 2000, vol. 1, pp. 370–375.

[3] A. C. Varsou and H. V. Poor, "HOLPRO: A new rate scheduling algorithm for CDMA downlink networks," in Proc. 2000 IEEE Fall Veh. Techol. Conf., Boston, MA, Sept. 24–28, 2000, vol. 2, pp. 948–954.

[4] L. Kalampoukas and A. Varma, "Design of a rate-allocation algorithm in an ATM switch for support of Available-Bit-Rate (ABR) service," in Proc. Design SuperCon 1997, Santa Clara, CA, Jan. 21–23, 1997.

[5] K. Gilhousen et al., "On the capacity of a cellular CDMA system," IEEE Trans. Veh. Technol., vol. 40, no. 2, pp. 303–312, 1991.

[6] L. Mailaender, "CDMA Downlink capacity with high-speed packet data," Bell Labs Engineer's Notes, May 1999. See also L. Mailaender et al., " Simple inter-cell coordination schemes for a high speed CDMA packet downlink," in Proc. 2000 IEEE Spring Veh. Technol. Conf., Tokyo, Japan, 2000, vol. 3, pp. 1845–1848.

[7] D. Stiliadis and A. Varma "A general methodology for designing efficient traffic scheduling and shaping algorithms," in Proc. IEEE INFOCOM 1997, Kobe, Japan, Apr. 7–11, 1997, pp. 326–335.

[8] CDMA/HDR: A Constructive (Backward Compatible) Approach for Migration to Wider Band Wireless Services, 3rd Generation Wider Band CDMA Technol. Conf., Atlanta, GA, Feb 1998.
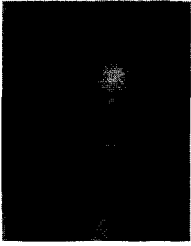
Aikaterini C. Varsou received the Ph.D. degree in Electrical Engineering in 2001 from Princeton University and a joint M.S.-B.S. in Computer Science in 1996 from National Technical University of Athens, Greece. She is currently working at the IP/Optical Division of the Computer and Communications Research Laboratories at NEC USA, Inc. In the past, she has worked for Bell Labs, Lucent Technologies, in the area of rate scheduling for CDMA systems. Her research areas include rate scheduling with applications in wireless communications, Network Calculus, packet forwarding for high-speed routers, while she is currently working on routing and protection algorithms for hierarchical optical networks. Among her honors are the Gordon Wu Fellowship of Princeton University and several fellowships from the National Science Foundation (IKY) and the Technical Chamber of Commerce in Greece. She is a member of the IEEE, the INFORMS society and the Technical Chamber of Greece.
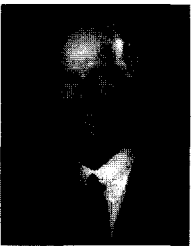
H. Vincent Poor received the Ph.D. degree in EECS in 1977 from Princeton University, where he is currently a professor of Electrical Engineering. He is also affiliated with Princeton's Department of Operations Research and Financial Engineering, and with its Program in Applied and Computational Mathematics. From 1977 until he joined the Princeton faculty in 1990, he was a faculty member at the University of Illinois at Urbana-Champaign. He has also held visiting and summer appointments at several universities and research organizations in the United States, Britain, and Australia. His research interests are primarily in the area of statistical signal processing, with applications in wireless communications and related areas. Among his publications in this area is the forthcoming book, Wireless Communication Systems: Advanced Techniques for Signal Reception. He is a member of the U. S. National Academy of Engineering, and is a Fellow of the Acoustical Society of America, the American Association for the Advancement of Science, the IEEE, the Institute of Mathematical Statistics, and the Optical Society of America. His IEEE activities include serving as the President of the IEEE Information Theory Society in 1990, and as a member of the IEEE Board of Directors in 1991-1992. Among his recent honors are an IEEE Third Millennium Medal (2000), the IEEE Graduate Teaching Award (2001), the Joint Paper Award of the IEEE Communications Society and IEEE Information Theory Society (2001), the NSF Director's Award for Distinguished Teaching Scholars (2002), and a Guggenheim Fellowship (2002-03).