Understanding the Trade-off between Multiuser Diversity Gain and Delay - an Analytical Approach

Roshni Srinivasan and John S. Baras
University of Maryland, College Park MD 20742
Email: {roshni,baras}@eng.umd.edu

***Abstract*— Innovative scheduling algorithms for packet-switched air-interfaces have been shown to support higher data rates by exploiting the *multiuser diversity* inherent to cellular wireless systems. While such *opportunistic* schedulers significantly improve the system throughput, they could degrade the user experience through unfair resource allocation and increased variability in the scheduled rate and delay. The growing demand for service differentiation between real-time multimedia traffic and data traffic underscores the need for these schedulers to incorporate delay constraints.**

The analytical results in this paper not only highlight the inherent trade-off between system throughput and the delay experienced by mobile users with opportunistic scheduling, but they also provide a novel way of quantifying the Quality of Service(QoS) offered by a general opportunistic scheduler. Our analysis is strongly supported by system simulations of a time-slotted cellular downlink shared by multiple mobile users with independent, time-varying channels. In addition, we use a continuous approximation to compute closed form expressions for the scheduler statistics.

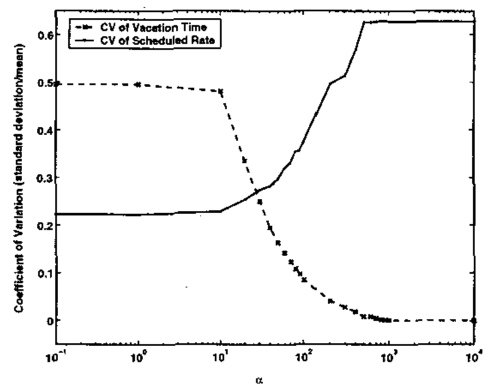

Fig. 1. CV of Scheduled Rate and CV of Vacation Time vs alpha for 16 Users at Nominal SNR of 2.5dB

I. Introduction

The utilization of multiuser diversity [1] to exploit fading in 3G and 4G cellular wireless systems has given rise to a new class of *opportunistic* schedulers. The maximum SNR scheduler best illustrates the gains acheivable by these schedulers. In a time-slotted system where mobile users constantly report channel quality to the transmitter, this scheduler maximizes system throughput by transmitting to the user with the best channel or the highest *requested rate* in every time slot. It is easy to see that the gains in system throughput come at the cost of unfair resource allocation and variability in the scheduled rate and delay. In order to illustrate this trade-off, consider the following scheduler metric, $m(t)$, that combines multiuser diversity gain with delay constraints:

$$m(t) = R(t) + \alpha\frac{v(t)}{N} = R(t) + \alpha V(t) \qquad (1)$$

In a time slotted system, $R(t)$ represents a mobile user's requested rate at the beginning of time slot $[t, t+1)$, $v(t)$ is the delay in slots since a waiting packet in the user's queue was previously served and α is a configurable control weight. The scheduling delay, normalized by the number of users is represented by $V(t)$. In the case of opportunistic schedulers, waiting packets could be delayed because the scheduler is serving other users either because their channel conditions are better or for reasons of fairness. We therefore refer to the normalized delay, $V(t)$ as scheduler *vacation time* in the rest of this paper. As the number of users increases, $v(t)$ increases proportionally. Using the normalized version of $v(t)$ ensures that the number of users does not affect the balance between multiuser diversity gain and delay implicit in the metric.

In Figure 1, we plot the coefficient of variation (standard deviation/mean) of the scheduled rate and the coefficient of variation(CV) of the delay between scheduling slots as a function of α. The scenario we consider is the downlink of a scheduled cellular wireless system similar to the 1xEV-DO [2] data system. In this system, which is described in more detail in Section IV-A, the BS serves mobile users in a cell using a time-slotted downlink with an asynchronous circuit-switched uplink. We see that for small values of α, the scheduler described above behaves like the Maximum SNR scheduler in which channel conditions dominate delay in the metric. As α is increased, the contribution of the delay towards selecting a user increases. For large α , the scheduler is channel agnostic and delay alone determines user selection. Users are served cyclically as in Longest Wait First (LWF) scheduling.

Scheduling algorithms for wireless networks that are optimized to support delay guarantees [3], [4], [5] have been well studied in the literature. Feasibility and complexity limit the practical implementation of these schedulers in providing explicit QoS guarantees. To the best of our knowledge, the results in this paper are the first to completely characterize the scheduled rate and delay experienced by mobile users in terms of a configurable scheduler metric.

II. Related Work

Knopp and Humblet [1] were among the first to highlight the concept of multiuser diversity in wireless systems. The Proportional Fair scheduler used in 1xEV-DO [6] is the first application of an opportunistic scheduler in a commercial wireless system. Other schedulers that incorporate delay constraints have also been proposed. The Modified-Largest Weighted Delay First (M-LWDF) [3] rule attempts to optimally provide QoS guarantees in terms of predefined guarantees for the probability of loss and minimum long-term throughput for each user. The exponential rule [4] is optimized to share a wireless channel among multiple real-time users with deadlines. The authors in [7] introduce the notion of effective capacity to characterize the time-varying capacity of the fading channel. Using effective capacity for admission control, they design a scheduling algorithm that combines Round Robin Scheduling with Maximum SNR scheduling to provide QoS guarantees. We take an entirely new approach and model the system formed by the BS scheduler and the mobile users with their time-varying channels as a dynamical system. In the discrete case, our analysis models a time-slotted system with a finite number of users. We also approximate the behaviour of the system in continuous time and compute closed-form expressions for the scheduler statistics.

III. Analysis

Let $\mathcal{N} = \{0, \ldots, N-1\}$ denote the set of mobile users served by the base station. For simplicity, time is assumed to be slotted and the channel conditions for each user are assumed to be identically distributed and independent from one time slot to the other. Naturally, the requested rates of the users are i.i.d. random variables. Let $\mathcal{R} = \{r_0, r_1, \ldots, r_{max}\}$ denote the finite set of rates requested by the users with $f_R(r) = P(R = r), r \in \mathcal{R}$. Then, for each user $i \in \mathcal{N}$, $f_{R_i}(r) = f_R(r)$.

A. The Discrete Case : A Time-slotted System with a Finite Number of Users

For a finite number of users, N, let $R_i(t)$ denote the requested rate and $V_i(t) = v_i(t)/N$ denote the vacation time seen by each user $i \in \mathcal{N}$, at the beginning of time slot $[t, t+1)$. In every time slot, the BS transmits to the user with the highest metric as computed from equation 1. In the event of a tie, a single user is picked with uniform probability from among the users with the highest metric. The analysis of the scheduler in the state space formed by the users, their delays and time-varying channel conditions is very complex. However, the problem becomes tractable and amenable to analysis when the state space is defined based on a *permutation* of the user space. In this space, users are *rank-ordered* in every slot according to the delay they have experienced since they were last scheduled. Let the vector, $\mathbf{U}(t)$ denote the rank-ordering of users at the beginning of time slot $[t, t+1)$.

$$\mathbf{U}(t) = \{u_0(t), u_1(t), ..., u_{N-1}(t)\} \quad (2)$$

By definition, this permutation has the property that

$$V_{u_0}(t) \leq V_{u_1}(t) \leq ... \leq V_{u_{N-1}}(t) \quad (3)$$

where $V_{u_i}(t)$ is the vacation time seen by the user who is ranked in position i at the beginning of time slot $[t, t+1)$. Naturally, since $u_0(t)$ is the index of the user scheduled in the previous time slot, $V_{u_0}(t) = 1/N$. At the beginning of time slot $[t, t+1)$, the scheduler picks a user whose rank, $S^*(t)$ is

$$S^*(t) = \arg\max_i m_{u_i}(t) \quad (4)$$

The selection of a user in one slot causes the rank-ordering of the users to change at the beginning of the next slot. At the beginning of time slot $[t+1, t+2)$, the user, S^* that was selected in the previous time slot moves to position 0 in the rank-ordered space. Since users are arranged in ascending order of their vacation times, all users with rank greater than that of S^* do not change their order in any way, while all users below the rank of S^* increment their rank by one. Specifically, $\mathbf{U}$ evolves over time as:

$$u_i(t+1) = \begin{cases} u_{S^*}(t), & i = 0 \\ u_{i-1}(t), & 0 < i \leq S^*(t) \\ u_i(t), & S^*(t) < i \leq N-1 \end{cases} \quad (5)$$

Correspondingly, the vacation time seen by every user who was not scheduled increases, while the vacation time seen by the scheduled user is reset to the minimum possible value

$$V_{u_i}(t+1) = \begin{cases} \frac{1}{N}, & i = 0 \\ V_{u_{i-1}}(t) + \frac{1}{N}, & 0 < i \leq S^*(t) \\ V_{u_i}(t) + \frac{1}{N}, & S^*(t) < i \leq N-1 \end{cases} \quad (6)$$

We define a selection density function, $\pi_{u_i}(t)$ which represents the probability of scheduling the ith rank-ordered user, u_i at the beginning of time slot $[t, t+1)$.

$$\pi_{u_i}(t) = Pr(S^*(t) = u_i), \quad i \in \mathcal{N} \quad (7)$$

with $\sum_{i=0}^{N-1} \pi_{u_i}(t) = 1$. The dynamics of the functions, $\mathbf{V}_u(t) = \{V_{u_i}, i \in \mathcal{N}\}$ and $\boldsymbol{\pi}_u(t) = \{\pi_{u_i}, i \in \mathcal{N}\}$ are governed by the choice of α in the scheduler metric and the time-varying channels of the users. In the long term, $\mathbf{V}_u(t)$ and $\boldsymbol{\pi}_u(t)$ converge to time-invariant functions, $\mathbf{V}_u$ and $\boldsymbol{\pi}_u$ respectively, that depend only on the statistics of the users' channels as well as α. We now analytically compute $\mathbf{V}_u$ and $\boldsymbol{\pi}_u$, using which we determine the scheduler statistics.

1) Computation of the Vacation Function, $\mathbf{V}_u$*:* The vacation function, $\mathbf{V}_u$ characterizes the normalized delay experienced by the users in the system. In the analysis that follows, we assume the existence of a selection density function, $\boldsymbol{\pi}_u(t)$. Recall from equations 5 and 6 that in the set of rank-ordered users, $\mathbf{U}$, the selected user S^* at the beginning of time slot $[t, t+1)$ moves to position 0 at the beginning of time slot $[t+1, t+2)$. All users with rank greater than S^* experience an increase in delay, but do not change their rank in any way.

$$V_{u_i}(t+1) = V_{u_i}(t) + \frac{1}{N}, \quad i > S^*(t)$$

As a result of the scheduled user occupying the very first position in the next slot, all users with rank less than S^* experience an an increase in rank as well as delay, i.e.,

$$V_{u_{i+1}}(t+1) = V_{u_i}(t) + \frac{1}{N}, \quad i < S^*(t)$$

The vacation function at position i in the rank-ordered space is therefore subject to two transforming forces. The first causes its value to increase by $1/N$ whenever a user with a rank less than i is scheduled. This event occurs with probability $\sum_{j<i} \pi_{u_j}(t)$. The second transformation causes its value to decrease whenever the rank of the user scheduled is i or higher. In this case, the value of the vacation-time at position i is replaced by that at position $(i-1)$ and also increased by $1/N$. The probability of this event is $\sum_{j \geq i} \pi_{u_j}(t)$.

In an equilibrium state, the vacation function is invariant to these transforming forces with the potential increase balancing the potential decrease. Dropping the dependence on time,

$$\frac{1}{N}(\sum_{j<i} \pi_{u_j}) = (V_{u_i} - (V_{u_{i-1}} + \frac{1}{N})) \sum_{j \geq i} \pi_{u_j}$$

With the initial condition, $V_{u_0} = \frac{1}{N}$, the vacation function $\mathbf{V}_u$ at equilibrium may be computed recursively as

$$V_{u_i} = V_{u_{i-1}} + \frac{1}{N\left(1 - \sum_{j<i} \pi_{u_j}\right)}, \quad 1 \leq i \leq N-1 \quad (8)$$

Observe that V_{max}, the upper bound on the normalized delay seen by any user can be computed from the above equation as

$$V_{max} = V_{u_{N-1}} = \frac{1}{N} + \sum_{j=1}^{N-1} \frac{1}{N\left(1 - \sum_{j<i} \pi_{u_j}\right)} \quad (9)$$

2) Computation of the Selection Density function $\boldsymbol{\pi}_u$: At the beginning of time slot $[t, t+1)$, the probability of selecting the i^{th} user is given by

$$\pi_{u_i}(t) = P(m_{u_i}(t) > m_{u_j}(t), \forall j \neq i) + P(u_i \text{ sel. in tie}) \quad (10)$$

The computation of the probability of selecting the i^{th} user in the event of a tie is outlined in the Appendix in [8]. Since the channel rates are i.i.d. random variables with distribution $f_R(r)$, the time-invariant selection density function π_{u_i} for each user $i \in \mathcal{N}$ at equilibrium is given by

$$\pi_{u_i} = \sum_{r=r_0}^{r_{max}} \prod_{j \neq i} F_R(r + \alpha(V_{u_i} - V_{u_j})) f_R(r) + P(u_i \text{ sel. in tie}), \quad (11)$$

where r_{max} is the maximum rate that can be supported by the mobile user.

3) Distributions for Scheduled Rate and Vacation Time: Let $\mathcal{V}_{S^*}$ denote the random variable representing the vacation time seen by the *scheduled* user. We see that, for some non-negative number γ,

$$\begin{aligned} P[\mathcal{V}_{S^*} \leq \gamma] &= \sum_{j=0}^{i(\gamma)} \pi_{u_j}, \quad \text{where} \quad (12) \\ i(\gamma) &= \arg\max_k (V_{u_k} \leq \gamma) \end{aligned}$$

The pdf of the scheduled rate as a function of α is

$$\begin{aligned} f_{R_{S^*}}(r) &= \sum_{i=0}^{N-1} f_R(r) (\prod_{j \neq i} F_R(r + \alpha(V_{u_i} - V_{u_j})) + \ldots \\ &\quad Pr(u_i \text{ sel. in tie})) \quad (13) \end{aligned}$$

B. Continuous Approximation for the Vacation Function

In this section, we derive an approximation for the vacation function in continuous time as the number of users, N becomes large, i.e., $N \to \infty$. The normalized, rank-ordered user space is now defined over the continuum $[0, 1]$ and the duration of a slot is infinitesimally small. Let the function $V(u, t)$ represent the vacation time experienced by a user with rank-ordered index, $u \in [0, 1]$ at time t. Naturally, the equivalent in the discrete case, $\mathbf{V}_u(t)$ may be derived by sampling $V(u, t)$ uniformly in $u \in [0, 1]$ at N points.

We also define $\pi(u, t)$, the continuous analog of $\boldsymbol{\pi}_u(t)$, which represents the time-varying probability of a user being scheduled, with the property that $\int_u \pi(u, t) du = 1$. Now, assuming the existence of a selection density function, $\pi(u, t)$, consider the time-evolution of the vacation function, $V(u, t)$. The process of scheduling users constantly subjects the function $V(u, t)$ to transformations. To understand this process, let us focus attention on the partial derivative $V_u'(u, t) = \frac{\partial V(u,t)}{\partial u}$. This derivative is subject to two transforming forces.

As in the discrete case, the local neighborhood of any point which has a non-zero probability of being scheduled experiences an increase in slope. This corresponds to the fact that the local user mass in any interval du around the point u is reduced due to scheduling. Neglecting higher order terms,

$$V_u'(u, t + dt) = \frac{V_u'(u, t)}{(1 - \pi(u, t)dt)} = V_u'(u, t)(1 + \pi(u, t)dt)$$

The *potential increase* in the local slope is therefore,

$$V_u'(u, t + dt) - V_u'(u, t) = V_u'(u, t)\pi(u, t)dt \quad (14)$$

The curve $V(u, t)$ constantly experiences a transforming force pushing it to the right i.e., increasing u, as a result of the total user mass in the interval $[u, 1]$ which gets scheduled. The amount by which the curve is shifted in a time dt is

$$\epsilon = [\int_{x=u}^{1} \pi(x) dx dt] = I(u)dt \quad (15)$$

The new slope at position u and time $t + dt$ is

$$V_u'(u, t + dt) = V_u'((u - I(u)dt), t)$$

Expanding $V_u'((u - I(u)dt), t)$ around the point u in a Taylor series and neglecting the higher order terms, we get

$$V_u'((u - I(u)dt), t) = V'(u, t) - V_u''(u, t)I(u)dt$$

where $V_u''(u, t)$ is $\frac{\partial^2 V(u,t)}{\partial u^2}$. The *potential decrease* in the local slope due to this transformation is given by

$$V_u'(u, t) - V_u'((u - I(u)dt), t) = V_u''(u, t)I(u)dt \quad (16)$$

In an equilibrium state, when the curve is invariant with respect to time, these two transforming forces are equalized. Hence, from Equations 14 and 16, we get

$$V_u''(u, t)I(u)dt = V_u'(u, t)\pi(u, t)dt$$

TABLE I

TRANSMISSION RATE PER SLOT AS A FUNCTION OF SNR

SNR (in dB)	Rate (Kb/s)	SNR (in dB)	Rate (Kb/s)
-12.5	38.4	-1.0	614.4
-9.5	76.8	1.3	921.6
-6.5	153.6	3.0	1228.8
-5.7	204.8	7.2	1843.2
-4	307.2	9.5	2457.6

which, by dropping the dependence on time gives us the vacation function at the fixed-point, $V(u)$ in terms of $\pi(u)$.

$$V(u) = \int_u exp(\int_u \frac{\pi(u)du}{\int_{x=u}^{1} \pi(x)dx})du \tag{17}$$

IV. SYSTEM MODEL AND IMPLEMENTATION

A. System Implementation

In our simulation experiments, we use a system architecture that is similar to the 3G CDMA wireless data systems such as 1xEV-DO and 1xEV-DV. We combine a time-slotted downlink with an asynchronous, cicrcuit-switched uplink. Packet streams for individual users are assigned separate queues by the BS. Every user is always assumed to have data in the queue. Fixed length packets of 512 bytes are segmented into link-layer (LL) segments of 8 bytes for transmission over the air link. At the beginning of each time slot, the scheduler at the BS computes the metric as in equation 1 and selects the data user with the highest metric. The number of segments transmitted in a slot depends on the current SNR of the selected data user; this correspondence is enumerated in Table I. The slot duration of 1.667ms and peak rate of 2.45Mbps achievable in this model are similar to the 1xEV-DO system. When all the LL segments corresponding to the packet at the head of the queue for a particular user have been transmitted over the airlink, the packet is dequeued. Transmission errors can be simulated by probabilistically delaying packet transmission. Since we assume that the channel state is known to a high degree of accuracy, we assume a negligible loss probability.

B. Wireless Channel Model

If $x_i(t)$ and $y_i(t)$ denote the vectors of transmitted and received symbols in a slot for user i, then

$$y_i(t) = h_i(t)x_i(t) + z_i(t), \qquad i = 0, 1, \ldots, N-1 \tag{18}$$

Using a flat fading model, we assume that the time-varying channel response $h_i(t)$, from the BS to the mobile is constant over the duration of the slot. $z_i(t)$ is a noise vector with i.i.d., zero mean Gaussian components with variance σ_i^2. Assuming unit-energy signals, the nominal SNR for user i is $C_{NOM,i} = \frac{1}{\sigma_i^2}$, with the instantaneous SNR for this user, $C_i(t) = \frac{h_i(t)}{\sigma_i^2}$. We use the well-known Jakes [9] model to approximate a single-path Rayleigh fading channel. In this model,

$$h_i(t) = \sum_{j=0}^{K-1} h_{i,j} exp(j2\pi f_d^i t cos(2\pi\phi_j)) \tag{19}$$

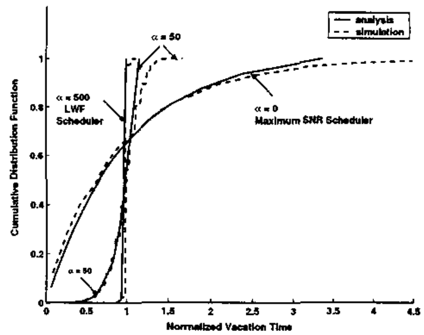

Fig. 2. Vacation Time distribution

$h_{i,j}, j = 0, .., K-1$ are complex gaussian random variables, $\mathcal{N}(0,1)$ representing the magnitudes of the subpaths, each with a phase delay, ϕ_j, that is uniformly distributed in $[0, 2\pi]$.

We study the performance of 16 users ($N = 16$), all with $C_{NOM,i} = 2.5dB$ and a doppler frequency f_d^i of 10Hz. A scenario with identical channel statistics for all users was selected to enable comparison between analysis and simulation. For the case of i.i.d. channel fading, $f_R(r)$ is chosen to be identical to the marginal distribution obtained with correlated channel fading at the same nominal SNR.

V. SIMULATION RESULTS

A. Numerical Computation of $\mathbf{V}_u$ and π_u

$\mathbf{V}_u$ and π_u can be computed analytically, assuming knowledge of each other, as outlined in Sections III-A.1 and III-A.2 respectively. Since neither function is known at the outset, we use the following iterative approach. We start with the Maximum SNR scheduler. If the N users have identical channel statistics, this results in a uniform selection function, $\pi_u^{(0)} = \frac{1}{N}$. $\mathbf{V}_u^{(1)}$ can be therefore be computed using the expression derived in Equation 8. $\pi_u^{(1)}$ is then computed from $\mathbf{V}_u^{(1)}$ as outlined in Section III-A.2. In subsequent iterations, $\mathbf{V}_u^{(k)}$ is computed from $\pi_u^{(k-1)}$, which in turn facilitates computation of $\pi_u^{(k)}$. The convergence of this process has been observed empirically [8]. We are currently working on a formal proof.

B. Distributions for Scheduled Rates and Vacation Time

For clarity, we plot the Cumulative Density Functions (CDF) of the vacation time and scheduled rates in Figures 2 & 3 resp. for three values of α. The graphs illustrate the close correspondence of the Monte Carlo simulation results with the analysis in Section III. In Figure 2, we see both from simulation and analysis that the LWF scheduler ($\alpha = 500$), concentrates the mass of the CDF at the normalized delay of 1, scheduling users in a Round Robin manner. Since the scheduler is channel agnostic, as in equation 10, we see that the probability of a user being selected is zero for all but the user with the highest delay. The Maximum SNR scheduler ($\alpha = 0$), on the other hand, ignores delay, always favoring

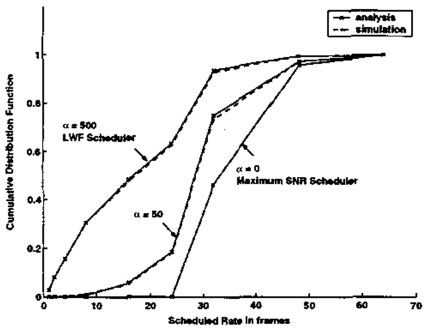

Fig. 3. Scheduled Rate distribution

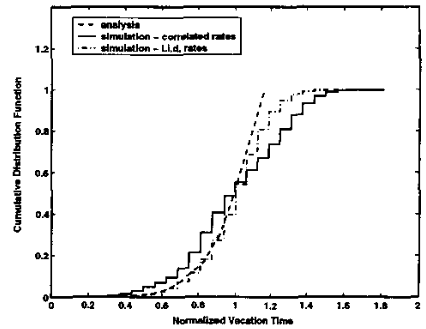

Fig. 4. Comparison of Vacaton Time CDFs for $\alpha = 50$ when channel rates are i.i.d. and correlated across time slots

users with higher SNR. The highest normalized vacation time is almost 5 times higher than that of the LWF scheduler.

Figure 3 illustrates the increase in throughput obtained by relaxing the delay constraint (smaller α) which manifests as the density concentrating in the higher end of the range. The CDF of the scheduled rates for the Maximum SNR scheduler is the product of the CDFs of the 16 users. Since there is a uniform probability of picking a user in any given slot, it follows that the CDF of the LWF scheduler is simply the CDF of the requested rates, $f_R(r)$ for any user as obtained from the Jakes model. In both figures, the CDFs for $\alpha = 50$ lie between the CDFs of the Maximum SNR and the LWF schedulers.

C. Correlated Rates

In our simulations, we also study the performance of the system with correlated fades, when the channel remains in a *good* or *bad* state across consecutive slots. As can be seen from Figure 4, which shows the CDF of the vacation time for $\alpha = 50$ and $f_d = 10$Hz, the probability of being scheduled at lower delays is higher when the requested rates are correlated. In this case the good channel conditions dominate delay in the metric in consecutive time slots. When the channel rates are correlated, there is also a higher probability of being scheduled when the vacation time is large. This may happen because (a) the user remains in a fade for a long duration, or (b) the user is pre-empted by other users with better channels for a sustained duration. This is apparent from the tail of the vacation time CDF in Figure 4.

VI. Conclusions

The time-varying wireless channel capacity adds a new dimension to the problem of supporting broadband data services in cellular networks. Implicit in the use of channel-state dependent scheduling algorithms are the questions of how these algorithms will address fairness and the provision of QoS guarantees for a mix of real-time traffic and data traffic. The analytical results in this paper address the important issue of quantifying the QoS provided by a cellular wireless system. We completely characterize the distributions of the scheduled rates and delay in a general scheduler which realizes multiuser diversity gain with constraints on scheduling delay. In order to study this trade-off, we use a general metric, $m = R + \alpha V$. The scheduler can be tuned to achieve the desired performance by varying the control parameter, α to balance the role of the channel rate, R or the normalized scheduling delay, V. In order to focus on the trade-off between throughput and delay alone, we assume that all users experience identical channel statistics. The metric can be modified to ensure *fairness* even when channels statistics vary across users as:

$$m_i(t) = (R_i(t) + \beta_i) + \alpha V_i(t) \tag{20}$$

where β_i can be chosen optimally to maximize the total scheduled rate while ensuring resource fairness for $\alpha = 0$. This result is proved in [5] and discussed further in [8].

The objective of this paper is not so much the design of an optimal scheduler as it is an analysis of the trade-off between throughput and delay in a general opportunistic scheduler. The metric decsribed in equation 1 lends itself well to analysis, while being sufficiently simple and versatile to be implemented in a real system. Our statistical analysis is validated by extensive simulations of a system architecture similar to a 1xEV-DO base station serving mobile users.

References

[1] R. Knopp and P. Humblet, *Information Capacity and Power Control in Single Cell Multiuser Communications*, IEEE ICC '95, June 1995.

[2] TIA/EIA IS-856 , *CDMA 2000 : High Rate Packet Data Air Interface Specification*, Standard, November 2000.

[3] M. Andrews, K. Kumaran, K. Ramanan, A.L. Stoylar, R. Vijayakumar, P. Whiting *CDMA Data QoS Scheduling on the Forward Link with Variable Channel Conditions*, Bell Labs Technical Report, April 2000.

[4] S. Shakkottai and A.L. Stoylar, *Scheduling for Multiple Flows Sharing Time-varying Channel : The Exponential Rule*, American Mathematical Society Translations, Series 2, Vol. 207, 2002.

[5] X.Liu, K.P.Chong, N.B.Shroff, *A Framework for Opportunistic Scheduling in Wireless Networks*, Comp. Networks, vol.41, no.4, 451-474, 2003.

[6] D. N. C. Tse, *Opportunistic Communications: Smart Scheduling and Dumb Antennas*, presentation available at http://www.eecs.berkeley.edu/ dtse/intel128.pdf, January 2002.

[7] D. Wu and R. Negi, *Effective Capacity - A Wireless Channel Model for Support of Quality of Service*, To appear in IEEE Transactions on Wireless Networking 2003.

[8] R. Srinivasan and J. Baras. *An Analysis of Delay-constrained Opportunistic Scheduling for Cellular Wireless Systems*, Technical Report, TR_2004-3, I.S.R. University of Maryland, Jan 2004.

[9] W. C. Jakes, *Microwave Mobile Communications*, IEEE Press, 1993.