

Integrating Power Control, Error Correction Coding, and Scheduling for a CDMA Downlink System

Yuming Lu and Robert W. Brodersen, *Fellow, IEEE*

Abstract—Transmitting multimedia data over a code-division multiple access (CDMA) channel presents a new set of challenges: data demands will sometimes exceed the system capacity, in which case the system must make the most efficient use of its limited resources. In this paper, we present our approach for unifying power control, variable forward error correction (VFEC), and scheduling for a downlink system by allocating the system resources. The resources we consider are: fixed bandwidth available for each user and the transmit power budget for each cell. Our objective is to maximize the overall system satisfaction, which we call “system utility.” This objective is achieved by applying a distributed algorithm which divides the overall optimization problem into a hierarchy of three levels (system, cell, and user), with each performing independent and parallel optimizations. In the second half of this paper, we perform simulation-based evaluation of the system performance with a simple cell structure and uniformly distributed users. The system is simulated using a specific utility function: the step function. Finally, several practical issues regarding the implementation are investigated.

Index Terms—Code-division multiple access (CDMA), multimedia, power control, variable forward error correction.

I. INTRODUCTION

THIS paper addresses design and control issues for multiuser multimedia indoor wireless communication systems. While this work is motivated by the Berkeley InfoPad project, the approach and the techniques are not limited to this system.

It is assumed that the downlink transmissions will use direct-sequence code-division multiple-access (DS-CDMA) scheme. The base station radios transmit a variety of multimedia data types ranging from low-data rate and error sensitive control information to high-data rate but more error tolerant multimedia data such as real-time video. In addition to this wide range of error tolerance and bandwidth requirements, multimedia data streams are bursty compared to traditional cellular voice data. Therefore, it is challenging to design a system that dynamically adapts to the time-varying channel and traffic and, at the same time, delivers the different data streams according to their required quality-of-service (QoS).

Manuscript received January 1997; revised January 1999. This work was presented in part at the Infocom Conference, San Francisco, CA, 1996, and at PACRIM, Victoria, B.C., Canada, 1997.

Y. Lu was with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 USA. She is now with Metawave Communications, Redmond, WA 98073 USA.

R. W. Brodersen is with the Department of Electrical Engineering and Computer Science, University of California, Berkeley, CA 94720 USA (e-mail: rb@eecs.berkeley.edu).

Publisher Item Identifier S 0733-8716(99)04074-3.

Several techniques can be applied for reliable transmission over a time-varying channel. Power control (i.e., adjusting the transmit power level at the base station) is used to control received signal-to-noise ratio (SNR). Forward-error-correction (FEC) introduces redundancy to combat transmission errors. Because bandwidth demands for multimedia applications vary greatly, prioritizing (i.e., scheduling) among data streams is needed when bandwidth demands exceed channel capacity.

Many solutions are feasible by various combinations of these three techniques. Therefore, designers are left to choose what they want to optimize. For example, previous work has generally been aimed at maximizing the Erlang capacity which is defined as the number of active users, minimizing transmit power while meeting certain SNR requirements, or maximizing the Shannon capacity. However, none of these design objectives consider the heterogeneities of multimedia data and users. This leads to many shortcomings, such as the difficulties of prioritizing one data type over others, of prioritizing users using these design metrics, or of distinguishing various users' preferences, etc. In this paper, we will present a “utility” based approach to the overall system control problem. The control algorithm integrates power control, variable forward error correction (VFEC), and scheduling in order to maximize the overall system utility, where utility measures the degree of user satisfaction. As shall be discussed later, we assume that a user's satisfaction with an application depends on the delivered bandwidth and bit error rate (BER).

This paper is organized as follows. First, the system structure and its constraints are presented. Next, we discuss the qualitative properties of utility functions. We argue that maximizing system utility should be the sole objective when designing a system, and we present the utility maximization algorithm. In the second half of this paper, we quantify the system characteristics with simulation results, for a particularly simple form of the utility function. Several traffic models are used in our simulation, including Poisson process and two-state continuous-time Markov chain (CTMC) in order to span a wide range of traffic characteristics. Finally, several practical issues regarding the implementation are investigated.

II. SYSTEM STRUCTURE

A. System Requirements

This system, shown in Fig. 1, is a wide-band multimedia communication network which is designed to serve a large number of users simultaneously through portable terminals [1]. This system employs asymmetric communication channels

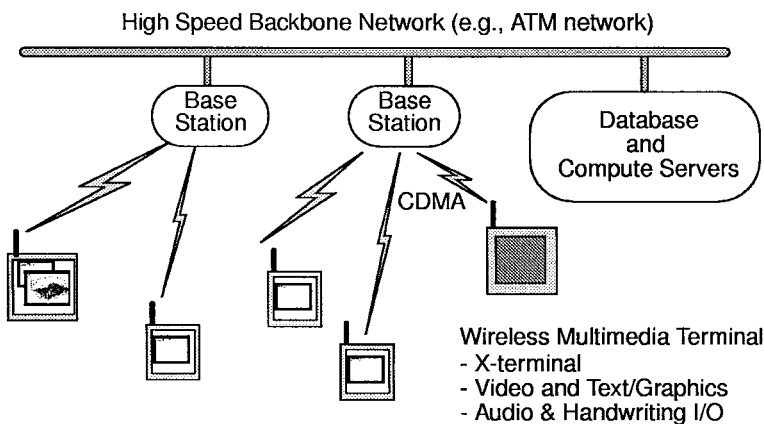


Fig. 1. Infopad system overview.

as the high speed real-time video is only available for the downlink transmission (i.e., base station to portable). We assume the downlink transmission employs DS-CDMA. For low power, cost, and usability reasons, these terminals have minimal general purpose computational power; instead, the computations are carried out at compute servers. The computational results, together with other types of multimedia data, are sent to the terminals via a high speed backbone network, base station, and wireless channel.

The downlink radio system has a limited capacity due to the limited resources such as transmit power and bandwidth. As will be shown later, the design of this system reduces to a global optimization problem. This optimization trades off the quality with the quantity of delivered data. In addition, the optimization trades resources among users. The objective of these tradeoff is to maximize the overall user satisfaction by applying power control, variable forward error correction (VFEC),¹ and scheduling.

Compared to conventional wireless communication systems, which have been designed mostly for low-rate single-data-type applications such as mobile voice and paging services, our CDMA downlink radio, on the other hand, is designed to address a different set of user imposed system performance requirements. First, our downlink supports multiple data types consisting of real-time video, audio, text/graphics and control data. Consequently, the bandwidth requirements are orders of magnitude greater. Second, the multimedia data consist of several data types with highly variable (QoS) (e.g., typical speech data can tolerate a BER around 10^{-3} , whereas control data require a BER around 10^{-9}). And finally, our multimedia system requires a scheduler because bandwidth is statistically shared among data types; the demands vary greatly and will sometimes exceed channel capacity. This scheduler dynamically allocates bandwidth for each application, leading to a priority scheme for multimedia data.

B. Practical Constraints

Our system is designed to fully utilize the available re-

¹Variable forward error correction (FEC) is a technique of using a family of FEC codes to achieve a wide range of error correction capabilities. For example, we use the family of 63-bit BCH codes with correction capability of $t = 0, 1, 2, 3, 4, 5, 5, 7, 10, 11, 13, 15, 31$ bit errors.

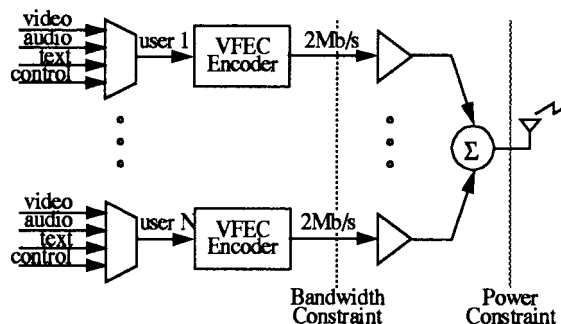


Fig. 2. System architecture and constraints.

sources while meeting the various QoS requirements of multimedia data. This system has a limited capacity due to the limited resources derived from a DS-CDMA radio. The resources we consider in this study are: 1) each user is allocated a fixed link bandwidth; the total bandwidth of a user's applications after FEC encoding must be less than this link bandwidth. In the case of the Infopad, it was determined that a 2-Mbit/s link bandwidth for each user was desired, and 2) the total transmit power from a basestation (which serves all users in a cell) must be less than an allowable power level set by implementation constraint. Fig. 2 shows the downlink radio structure together with these practical constraints.

C. Three "Control Knobs"

In order to achieve the optimal system performance, three "control knobs" are available for fulfilling various bandwidth or BER requirements. These knobs are VFEC, power control, and scheduling. Among them, VFEC selects a particular FEC code; this FEC then introduces redundancy to combat transmission errors. Downlink power control varies the transmit power to adjust the received signal quality. Together VFEC and power control are used to support applications with widely varying QoS and to mitigate excessive interference. Finally, prioritizing (i.e., scheduling) allocates bandwidth among data types; this is especially needed when application demands exceed channel capacity.

The goal of our design is to unify VFEC, power control, and scheduling into one system. We claim that the system design

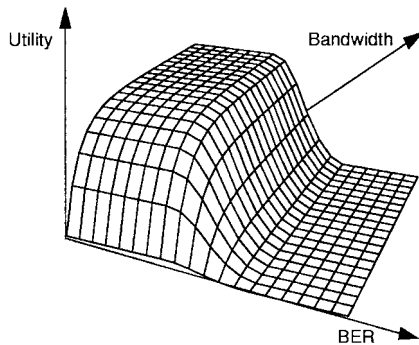


Fig. 3. A typical application utility function.

is incomplete if the methods of VFEC, power control, and scheduling are not considered simultaneously. For example, consider a design that only controls power; if all active applications consume less than the available bandwidth, then the remaining bandwidth is wasted. Compare this to a system that, instead of wasting this bandwidth, uses it for FEC coding *and* reduces the transmit power while maintaining the same QoS. This example illustrates the interaction between power control and FEC. Finally, the result from scheduling affects both the VFEC selection and the transmit power level in order to meet the specific QoS of that data type.

III. SYSTEM OBJECTIVE AND CONTROL STRATEGIES

A. Concept of Utility

A large set of solutions are feasible by various combinations of VFEC, power control, and scheduling. For example, suppose a user can choose receiving low resolution video together with text graphics *or* just high resolution video; or the system may either support six users with high QoS *or* ten users with medium QoS, how do we choose between these alternatives? A multiuser multimedia system is ultimately designed to satisfy users. Therefore, any design decisions should be based on user satisfaction. For instance, if users prefer receiving more data at cost of having a higher BER, the system design should reflect that accordingly. Therefore, our objective is to maximize total user satisfaction, which we call “system utility.”

In general, the concept of system utility is somewhat vague. However, if we assume that utility is additive, then system utility becomes the sum of user utilities, and a user’s utility is the sum of that user’s application utilities. For each application, the performance clearly depends on the quality and quantity of data delivered. In this study, we express application utility in terms of delivered BER and bandwidth. An example on application utility function is shown in Fig. 3.

At this point, let us discuss the qualitative properties of the application utility functions.² We proceed by first holding BER constant so that utility only depends on delivered application bandwidth.

For all applications, the application utility is a monotone nondecreasing function with respect to the bandwidth. We can categorize applications into many classes; nevertheless,

²The actual utility function for any particular application can be determined through either simulations or experiments with users.

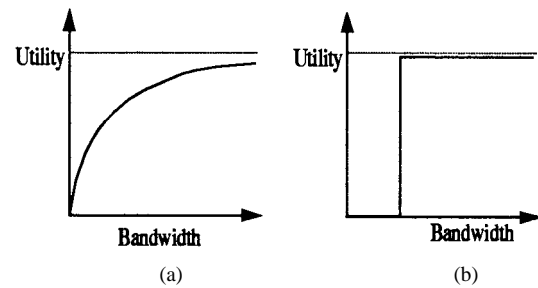


Fig. 4. (a) and (b) Two classes of utility functions with respect to the bandwidth.

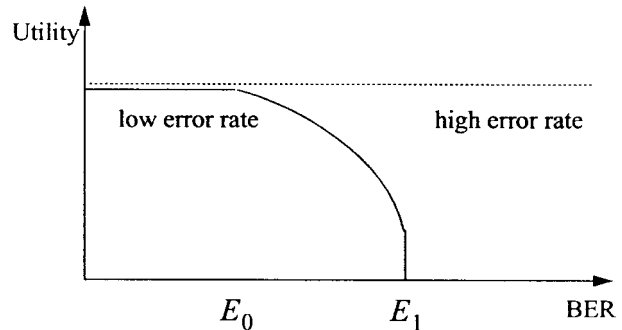


Fig. 5. Utility function with respect to the error rate.

we will discuss and contrast only two such classes. One class includes applications for which performance gradually improves as their allocated bandwidth increases however, with a decreasing marginal utility (e.g., video and text/graphics). The utility functions for this class of applications are therefore concave everywhere, as shown in Fig. 4(a). Another class includes applications such as control information for which the received data are of no value to users if only partial information is delivered; however, once the necessary amount of data is delivered, there is no extra benefit for receiving more data. Fig. 4(b) shows the utility function for this class of applications.

Let us now turn our attention to the other parameter of the utility function, that is, the error rate. When the received BER is high, users are generally unsatisfied with application performance. As the error rate improves, their satisfactions rise as well. However, once the BER improves beyond a certain level, very little additional satisfaction is achieved. For instance, the reception quality of video is nearly identical between BER’s of 10^{-4} and 10^{-8} . Fig. 5 illustrates the utility function with respect to BER. As a final remark, the application utility as a function of bandwidth and BER shown in Fig. 3 is obtained by multiplying utilities from Fig. 4(a) and Fig. 5.

B. Layered Approach to Utility Maximization

Now that we have discussed the qualitative behavior of the application utility functions, let us return to the problem of maximizing the system utility subject to constraints. Suppose a user makes a request for several applications, each with utility $u_i(B_i, E_i)$, where B_i is the application bandwidth, and E_i is the received BER (after FEC decoding). Recall that the

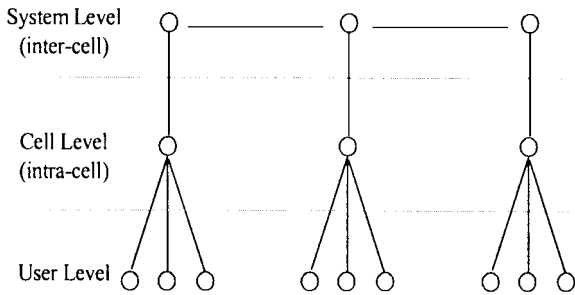


Fig. 6. Three levels of optimization hierarchy.

utilities are assumed to be additive, therefore, the user, cell, and system utilities can be expressed as

$$\text{user util} = \sum_{i=1}^L (\text{appl. util})_i = \sum_{i=1}^L u_i(B_i, E_i) \quad (1)$$

$$\text{cell util} = \sum_{j=1}^M (\text{user util})_j = \sum_{j=1}^M \sum_{i=1}^L (u_i(B_i, E_i)) \quad (2)$$

$$\text{system util} = \sum_{k=1}^N (\text{cell util})_k = \sum_{k=1}^N \sum_{j=1}^M \sum_{i=1}^L (u_i(B_i, E_i))_{jk}. \quad (3)$$

With this formulation of the utility functions, we are able to overcome the shortcomings from the previous design objectives (from Section I). First, a user may prioritize various applications according to (1). Second, by introducing a weight factor, users in the system can be prioritized as well, as in (3).

Our objective is to maximize the system utility subject to constraints, where parameters B_i and E_i are controlled by power control, VFEC, and scheduling. One approach to this global resource allocation problem is to apply a centralized algorithm which considers all users' statistics simultaneously. The advantage of this approach is that it does not require any iterative steps for achieving the optimum; however, the computational complexity and communication requirements are impractical for any reasonably sized system. Since our system is designed to be scalable, we propose a distributed algorithm that divides the overall control problem into three separate levels: user level, cell level, and system level, as shown in Fig. 6.

The nodes within each level are independent and can be optimized in parallel. The optimization results, expressed in terms of the resource requirements, are passed up to the higher level. For example, the optimal user utility, expressed in terms of user's channel SNR, is directly proportional to the user transmit power level; this transmit power level is then determined by the intracell power allocation. In addition, at the system level, a cell communicates with its interfering neighbor cells to negotiate its cell power budget so as to maximize the neighborhood utility. This layering approach yields a distributed algorithm, and the details are explained in the following section using a bottom-up strategy: from user level to cell level, to system level.

We assume the feedback loop from a portable terminal to a basestation gives the basestation an estimate of the channel

SNR. The basestation then calculates the propagation loss and the interference factors for each user, based on the user's channel SNR and the transmit power levels of the neighbor cells.

1) *User Level Optimization:* At the user level, we optimize user utility. Two techniques are applied: first, the system performs scheduling by allocating bandwidth to each application. Second, the system optimizes error rate by applying VFEC. VFEC enables the system to trade-off the quality with quantity of delivered data for each application.

Recall that the user utility is the sum of the application utilities

$$U_{\text{user}}(\bar{B}, \bar{E}) = \sum_{i=1}^L u_i(B_i, E_i). \quad (4)$$

Our objective is to maximize (4), over $\bar{B} = (B_1, B_2, \dots, B_L)$ and $\bar{E} = (E_1, E_2, \dots, E_L)$, subject to the link bandwidth constraint

$$\sum_{i=1}^L B_i \cdot \eta(\text{SNR}, E_i) \geq \text{Link Bandwidth} \quad (5)$$

where

$$\eta(\text{SNR}, E_i) = \frac{\text{link bandwidth}}{\text{data bandwidth}}. \quad (6)$$

To maximize the user utility, we do not directly choose the received BER E_i ; instead, we select the level of VFEC, together with the known link SNR, to determine the received BER. The choice of FEC code is reflected by the "bandwidth expansion function," $\eta(\text{SNR}, E_i)$, which corresponds to the FEC code that achieves E_i (after decoding) for a given channel SNR. $\eta(\text{SNR}, E_i)$, expressed in (6), is the ratio of the link bandwidth to the data bandwidth, thus $B_i \cdot \eta(\text{SNR}, E_i)$ is the actual channel bandwidth consumed by application i .

Maximizing (4) under the constraint (5) is an optimization problem over $2L$ variables (\bar{B} and \bar{E}). Observe that the channel SNR is the only undetermined variable during the optimization; as a result, both the optimal \bar{B} and \bar{E} , and thus the optimal user utility, are functions of the channel SNR. We denote the maximum user utility by $U_{\text{user}}^*(\text{SNR})$.

To summarize, two results are achieved through this user level utility optimization. First, for any SNR, we can apply scheduling and FEC selection (i.e., choosing \bar{B} and \bar{E}) such that the user utility is maximized. Second, the maximum user utility can be expressed as a function of SNR (presuming an optimal choice of \bar{B} and \bar{E}). As will be shown in the next section, this maximal user utility function becomes the cornerstone of the cell level power allocation.

2) *Cell Level Optimization:* An indoor wireless communication environment is interference limited. In such an environment, users are subject to three sources of interference: intercell interference, intracell interference, and background noise. The channel SNR for a user, say user j located in cell #0, is the ratio of the received signal power to the noise power:

$$\text{SNR}_j = \frac{\text{Received Signal Power}}{\text{Total Noise Power}} = \frac{g_j \cdot \phi_j \cdot P_0}{\delta_j + \Delta_j + (\sigma_B^2)_j} \quad (7)$$

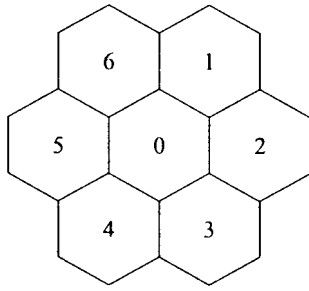


Fig. 7. Cell topology.

where g_j is the aggregate gain factor for user j , ϕ_j is the fraction of the total cell transmit power allocated to user j , P_0 is the total transmit power for cell #0, and $\delta_j, \Delta_j, (\sigma_B^2)_j$ are the intracell interference, intercell interference, and background noise, respectively.

Assume a cell topology as shown in Fig. 7, where a center cell has six neighboring cells.³ Suppose cell # i is in the center, and let $\bar{P}_i = (P_i, P_i^1, \dots, P_i^6)$ be the total power budget for each of the seven cells, where (P_i^1, \dots, P_i^6) are neighboring cell power levels. Using this notation, \bar{P}_0 represents the power budget for a cell neighborhood where #0 is in the center.

Without loss of generality, our focus is on user j in cell #0 with six adjacent neighbor cells. The intracell and intercell interference are modeled as fractions of cell power levels, which are expressed in (8) and (9), respectively.⁴ The α_j and β_j are intracell and intercell interference coefficients which depend on the CDMA spreading factor and propagation environment. These coefficients can be determined through measurements⁵

$$\delta_j = \alpha_j \cdot P_0 \quad (8)$$

$$\Delta_j = \beta_{j1} \cdot P_0^1 + \beta_{j2} \cdot P_0^2 + \dots + \beta_{j6} \cdot P_0^6. \quad (9)$$

We can combine (8) and (9) using a vector notation to represent the overall interference noise power

$$\delta_j + \Delta_j = \bar{\gamma}_j \bullet \bar{P}_0 \quad (10)$$

where $\bar{\gamma}_j = (\alpha_j, \beta_{j1}, \dots, \beta_{j6})$, and \bullet is the vector dot product. Consequently, substituting the expression of interference noise power (10) into (7), the channel SNR for user j is

$$\text{SNR}_j = \frac{g_j \cdot \phi_j \cdot P_0}{\bar{\gamma}_j \bullet \bar{P}_0 + (\sigma_B^2)_j}. \quad (11)$$

During the downlink transmission, a CDMA radio transmits to all users within a cell simultaneously and is subject to a power budget which is determined through the higher layer optimization. The goal of performing cell level optimization, for cell #0 in our case, is to distribute the power budget P_0 to

³This assumption of having hexagonal cells is not necessary for our algorithm; it is only for illustrative purposes.

⁴Throughout this study, we only consider the first order intercell interference, i.e., interference coming from the adjacent cells.

⁵The intracell and intercell interference coefficients for a user depend on the location of the user and the indoor environment, both of which are time varying. These interference coefficients can be estimated by correlating total noise power (which can be measured) with this and neighboring cell power levels.

each user so that the total cell utility is maximized. This goal is achieved by computing the optimal ϕ_j for all users j .

Suppose cell #0 has M users, the total cell utility is therefore the sum of maximal user utilities

$$\begin{aligned} U_{\text{cell } 0}(\bar{\phi}, \bar{P}_0) &= \sum_{j=1}^M U_j^*(\text{SNR}_j) \\ &= \sum_{j=1}^M U_j^* \left(\frac{g_j \cdot \phi_j \cdot P_0}{\bar{\gamma}_j \bullet \bar{P}_0 + (\sigma_B^2)_j} \right). \end{aligned} \quad (12)$$

Our objective is to maximize (12) subject to the power budget constraint

$$\sum_{j=1}^M \phi_j \cdot P_0 \leq P_0. \quad (13)$$

Maximizing cell utility is an optimization problem of M variables $(\phi_1, \phi_2, \dots, \phi_M)$. Since the cell power budget for the entire neighborhood \bar{P}_0 remains unknown during the optimization; therefore, the optimal user power allocation $(\phi_1, \phi_2, \dots, \phi_M)$, and thus the maximal cell utility is a function of \bar{P}_0 . We denote the maximum utility for cell #0 as $U_{\text{cell } 0}^*(\bar{P}_0)$.

It is not difficult to prove that the two-step optimization process (i.e., the user level and the cell level optimization) for maximizing the cell utility is equivalent to a one-step optimization at the cell level. This one-step optimization determines the B_i , the E_i for each application among all users, and ϕ_j for each user in the cell. However, the layered approach introduces parallelism thus reducing the intercell communication.

To summarize, we have achieved the optimal user power allocation for multiple users within a cell. The optimal cell utility depends on the cell power levels for the entire neighborhood. This result provides us with a platform to perform intercell power allocation, which is to determine the total cell power budget for each cell.

3) System Level Optimization: Our ultimate goal is to maximize the overall system utility which is achieved by setting the cell power budget for each cell. Since a large number of users and cells are in the system, we seek a distributed algorithm that is scalable in both computation and communication.

As intercell interference is localized to a finite region, changing the power level for a cell only affects its nearby cells; we call this region that cell's "neighborhood." This observation suggests that we are able to simultaneously change power levels for several cells, provided that their neighborhoods do not overlap. For this discussion, we assume only first order interference.⁶ Fig. 8 shows several nonoverlapping neighborhoods (indicated by the "stars" superimposed on the cell topology).

Several key properties result if we restrict ourselves to changing power levels only for the center cells of these

⁶First-order interference means that a base station only interferes with its six immediate neighbor cells. This assumption is only for illustrative purposes, and is not necessary for our distributed algorithm. If the assumption does not hold, we can increase the neighborhood size, which will decrease the rate of convergence.

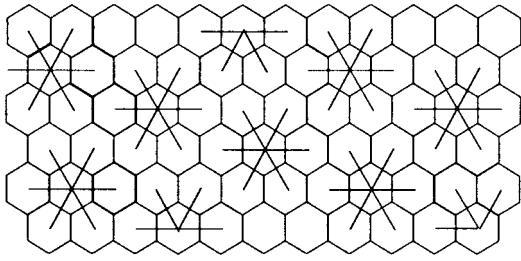


Fig. 8. Nonoverlapping interference neighborhoods.

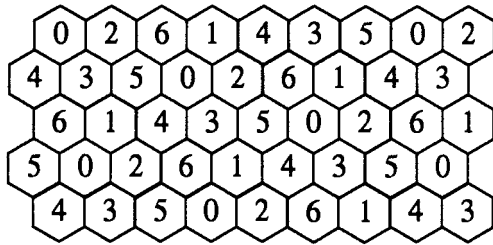


Fig. 9. Iteration step assignments.

nonoverlapping neighborhoods. First, the effects from changing a center cell's power level are limited to the neighborhood boundary; therefore, cells only need to communicate within the neighborhood. Second, a center cell faces a fixed interference environment; therefore, calculating the power budget is simplified. Finally, the remaining cells within the neighborhood have exactly one interfering cell which is changing its power; therefore, estimating intracell and intercell interference coefficients ($\bar{\gamma}_j$ from Section III-B2) is simplified.

So far, we have updated cell power budgets for only a fraction of cells. Our ultimate goal is to update power budgets for all cells in the system. This is achieved by iterating according to a periodic schedule,⁷ where the system adjusts a different subset of cells at each iteration; after one period, all cells in the system are updated at least once.

We will demonstrate this algorithm using our example of first-order interference and hexagonal cell topology. Fig. 9 shows the assignment of cells to iteration steps, and the iteration period equals seven in this case. Notice at every time step, the subset of cells which change power levels have nonoverlapping neighborhoods.

We next describe how we determine the cell power budgets. The procedure is the same for all cells; however, we will focus on cell #0, with neighboring cells 1–6.

When updating the power budget for cell #0, its power level is chosen so as to maximize its overall neighborhood utility. We know from Section III-B2 that the maximal cell utility for each cell in the neighborhood of cell #0 depends on P_0 . Therefore, the total neighborhood utility, which is the sum of the maximal cell utilities, is also a function of P_0 , as shown below

$$U_{\text{Neighborhood}}(\bar{P}_0, \bar{P}_1, \dots, \bar{P}_6) = U_{\text{cell } 0}^*(\bar{P}_0) + \sum_{i=1}^6 U_{\text{cell } i}^*(\bar{P}_i) \quad (14)$$

⁷This schedule is fixed and determined by the system designer.

While P_0 is upper bounded by the implementation limit, the goal of adjusting P_0 is to maximize the total neighborhood utility, P_0 thus has to satisfy

$$\frac{\partial}{\partial P_0} [U_{\text{Neighborhood}}(\bar{P}_0, \bar{P}_1, \dots, \bar{P}_6)] = 0.$$

This is equivalent to

$$\frac{\partial}{\partial P_0} U_{\text{cell } 0}^*(\bar{P}_0) = -\frac{\partial}{\partial P_0} \left(\sum_{i=1}^6 U_{\text{cell } i}^*(\bar{P}_i) \right). \quad (15)$$

Notice that on the left side of (15) is cell #0's marginal utility as a function of P_0 , and on the right is the total marginal utility of the neighbor cells. At the optimal point, with respect to P_0 , the marginal utility of cell #0 offsets the marginal utility of the neighbor cells.

The system utility optimization algorithm we described above is indeed an iterative, one-dimensional (1-D) search algorithm, in which we optimize along each P_i iteratively until reaching the maximum. There are known numerical algorithm to meet this purpose [6].

IV. SIMULATION

We have presented the "utility"-based approach to the overall control problem. Because the policy issues of allocating resources are separated from the design process, the main advantage of this framework is that it offers a great deal of flexibility to the system control. In other words, we can always choose a utility function that reflects a specific design objective. For example, if we let the application utility function be a step function (with respect to both bandwidth and error rate), then this would coincide with conventional design objectives for many wireless or cellular systems, where a BER (or SNR) requirement is imposed. However, this advantage of flexibility makes it difficult to quantify and compare system performance with other approaches. Therefore, for the rest of this paper, we will simulate the system using a specific utility function, the step function. This restriction will allow us to evaluate and quantify system performance with numerical results. In addition, these results will uncover some interesting dependencies which differentiate the multimedia optimization problem we are investigating from a single media application.

A. Simulation Setup and Parameters

The simulation assumes that each user is assigned to a 2-Mbit/s channel. Each data type has a fixed BER requirement, so data are only transmitted when the estimated received BER (after decoding) equals or exceeds this predetermined level. Since it is difficult to compare users when they have different traffic profiles, we assume all users request four types of data. The data types have static priorities, so the scheduling is well understood and will not play a major role when we evaluate power control and VFEC. As a consequence, the lower priority data can only get ahead of higher priority data when the network resource is insufficient to transmit the higher priority data. Table I shows the parameters modeled by Poisson arrivals for all data types along with their relative priorities.

TABLE I
TRAFFIC PROFILE

priority	BER requirement	packet size (bits)	arrival rate (msec)	data rate (kbps)
1 (control)	10^{-9}	100	8	12.5
2 (text/graphics)	10^{-4}	200	1	200
3 (audio)	10^{-5}	100	2	50
4 (video)	5×10^{-4}	100	0.125	800

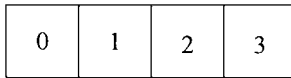


Fig. 10. Cell topology.

For simulation purposes, we assume a linear cell topology as shown in Fig. 10. This layout is chosen because it is the simplest topology which contains all aspects of a multicell environment. Since the system performance is a function of user population, location, and traffic, the simplicity of this layout makes the dependencies and cell interactions more transparent for analysis. We believe the general results and trends from this case carry over to more realistic topologies.

Four cells are simulated; they are labeled as 0, 1, 2, 3, and each has the population 4, 6, 5, 5 users, respectively. The users are assumed to be uniformly distributed in each cell. Each cell is $5 \times 5 \text{ m}^2$ in dimension, and the ceiling is 4 m high. Moreover, since these are indoor users, we assume change in user locations are negligible compared to the data rate and the frequencies which intercell and intracell algorithms are performed.

Experimentally, we found it is sufficient to simulate the system for 50 000 iterations, with each iteration corresponding to 1 ms. Both user level and cell level algorithms are executed at each iteration, whereas the system level algorithm is carried out once every three iterations. Later in Section IV-B3, we will present the system performance using various traffic models, including Poisson arrival, constant bit rate, and two-state continuous time Markov chain (CTMC). However, for the most part of this study, the packets for each data type are modeled by Poisson arrivals, unless it is specified otherwise.

B. Results

The system performance is evaluated using average aggregate throughput of delivered data for each user. This average aggregate throughput is calculated with the data that are received at the required BER. The maximum throughput at receiver for any user, is around 1.1 Mbit/s, according to Table I.

1) *Dependencies on User Distribution:* First let us explore the dependencies of throughput on user distribution. Ten independent and random sets of user locations are simulated while the population in each cell remains the same. Fig. 11 shows the average aggregate throughput for all users in the system. Users are ranked according to their throughput, from highest to lowest.

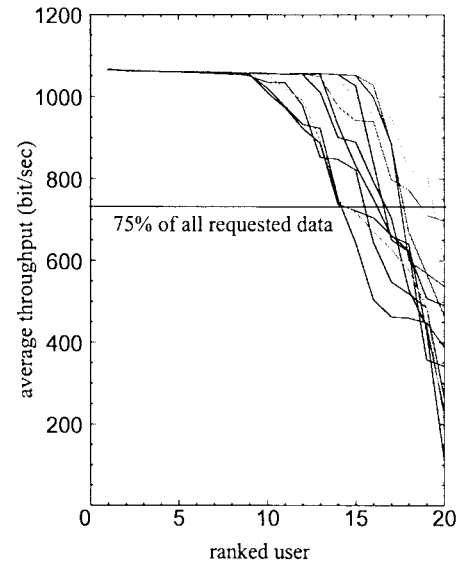


Fig. 11. User throughput for ten independent sets of user locations.

Clearly, the performance depends very much on the user distribution: the set of location which yields the best performance is shifted to the right by nearly five users comparing the location yields the worst performance. However, despite the variations from one location to another, all throughput distributions are nevertheless skewed with more users residing on the left side, indicating majority of users have very little data loss.

2) *Fairness Constraint:* “Fairness” is imposed during the intracell power allocation, because some users need more transmit power to compensate for a poor channel and/or heavy traffic demand. In the case when there is *not* enough power to meet everyone’s demand, it is difficult to justify why one user should receive more network resource than another. In the first half of this paper, we presented a utility driven optimization framework for allocating resources, in which case no fairness was considered. In this section, we will implement a “fairness” policy (defined below) and will compare the results against the algorithm that maximizes the number of users.

The “fair” power allocation consists of two stages: fairness-based power allocation followed by demand-based allocation. During the fairness-based power allocation, the total cell transmit power is equally divided among all users. The purpose of this step is to guarantee each user with a fair share of the available resource, independent of the channel quality or traffic demand. If the allocated resource is greater than a

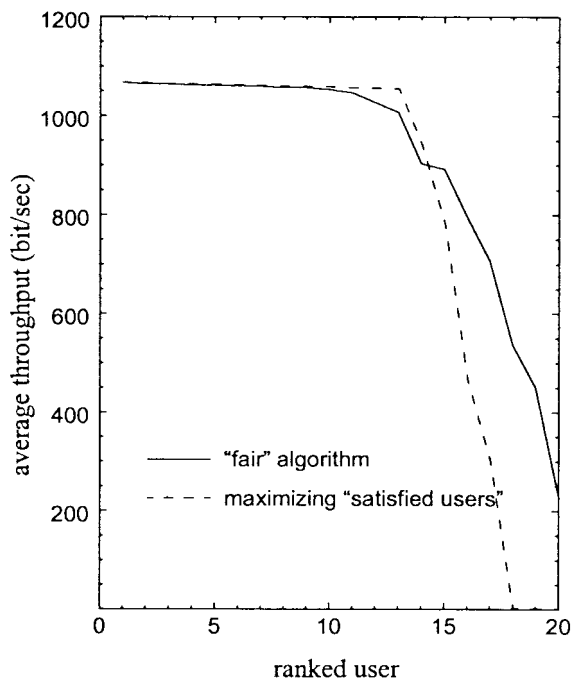


Fig. 12. Performance comparison for systems with and without fairness constraint.

user’s demand, the excess resource is then collected and then redistributed among the users (within the same cell) whose demands are greater than the allocated power. We call this redistribution the demand-based power allocation.

The result is compared against the algorithm that maximizes the number of “satisfied users,” where a satisfied user is defined as a user who receives all requested data. Note this is *not* equivalent as maximizing throughput as in the voice-only system. When a system has multiple data types and thus multiple QoS requirements, maximizing throughput may allow lower priority data ahead of higher priority data, if lower priority data require less resources to transmit. Since some high priority data such as control may carry crucial information for decoding lower priority data, we want to maintain a strict priority among data types.

The system incorporating all three control techniques is simulated for both “fair” and “competitive” cases. The results are shown in Fig. 12. The “competitive” algorithm slightly outperforms the “fair” algorithm for up to 14 users, but then becomes notably inferior to the “fair” system at the tail of the distribution. This is because the objective for the “competitive” algorithm is to maximize the number of active users; therefore, it tends to favor users who have better channels and thus allocates all resources to these users. From the result, we believe fairness benefits the overall system, especially users who have poor channels. Fairness is implemented as a part of the power control unless it is specified otherwise. As a final remark, the user’s distribution used in this simulation is the one that yields the median throughput. This distribution will be used for the remainder of the results.

3) *Dependencies on Traffic Profile:* This system is designed to support traffic sources having diverse statistics. In order to design a robust system that supports data types for

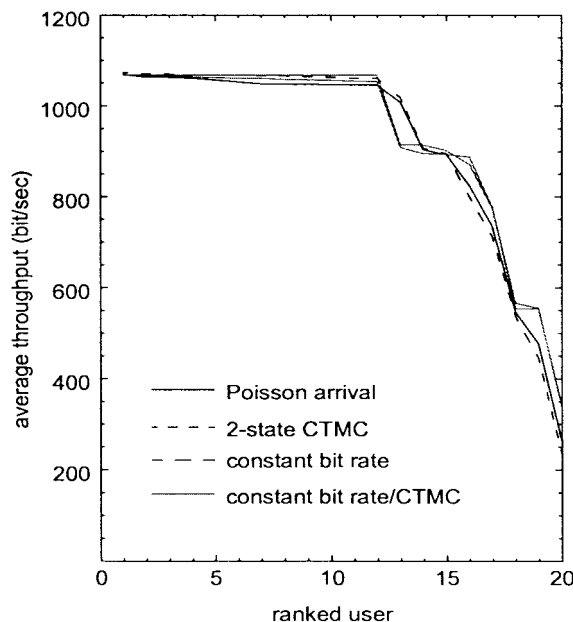


Fig. 13. System performance with four traffic models.

a wide range of arrival statistics, we test the system through several traffic models. In this study, the traffic sources are:

- 1) Poisson arrival (same as the one used in the previous section);
- 2) CTMC for modeling highly bursty traffic;
- 3) constant bit rate;
- 4) a combination of two-state CTMC and constant bit rate.

These patterns were chosen to be diverse, spanning very bursty to constant rate transmission, with only the mean arrival rate and packet size for each data type held constant. The parameters for these four models are listed in Table II.

Fig. 13 shows user throughput under these four scenarios, where no major performance difference was present. We therefore conclude the mean data rate and packet size are good parameters for characterizing the overall received data.

4) *Power Control or VFEC?:* The distributed algorithm we have described involves a number of control variables and thus has very complicated dependencies. To determine the necessity of this complexity we investigate if a simplified algorithm can offer comparable performance. For this reason, only two out of three techniques are applied: power control combined with scheduling, and VFEC combined with scheduling (scheduling is always included since it is necessary to prioritize the data).

From the traffic model, we know the average total traffic demand for each user is greater than half of the available link bandwidth, which makes us expect that power control would play a significant role in any optimization. The simulation results are compared against the original algorithm which consists all three techniques. The results, however, as shown in Fig. 14, indicates power control/scheduling, with or without the fairness constraint, is inferior to VFEC/scheduling. The power control is not so important here as it is in single media case, mostly because we support multiple kinds of data. In

TABLE II
PARAMETERS FOR VARIOUS TRAFFIC MODELS

	type 1 (control)	type 2 (text/graphics)	type 3 (audio)	type 4 (video)
packet size (bit)	100	200	100	100
avg data rate (kbps)	12.5	200	50	800
BER requirement	10^{-9}	10^{-4}	10^{-5}	5×10^{-4}
Poisson Arrival	refer Table 1			
2-state continuous time Markov chain				
p(on→off)	0.04	0.1	0.02	0.2
p(off→on)	0.01	0.2	0.01	0.8
constant-bit-rate/CTMC				
	2-state CTMC	2-state CTMC	constant bit rate	constant bit rate

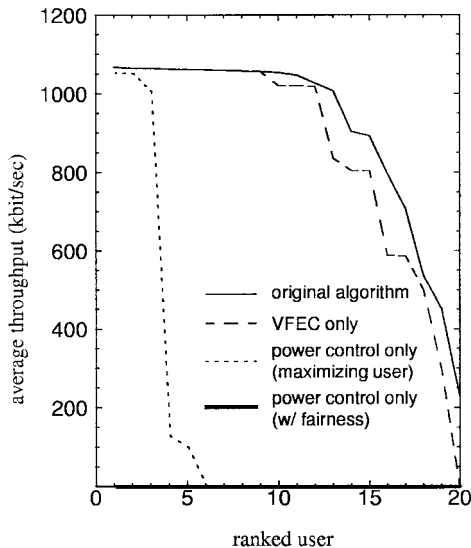


Fig. 14. Performance comparison when applying only two out of three techniques.

our system, the VFEC adapts to the time-varying traffic at packet level by encoding each data segment with the most suitable FEC; as a result, it is able to take full advantage of the total available bandwidth. The power control, on the other hand, operates at a higher level, and thus does not enjoy as much flexibility as the VFEC. The results here indicate the importance of providing direct error correction.

5) *Power Control with Fixed FEC*: The question then arises as to the importance of including the complexity of variable rate error correction, since existing cellular standards, such as IS-95, implement power control with a fixed-rate convolutional code for both uplink and downlink communications. We will apply this to our multimedia system using block codes.

To conduct fair comparisons with the algorithm that incorporates VFEC, five 63-bit BCH codes are chosen from the same family of codes used for VFEC. These codes range from double-error-correction code to eleven-error-correction code; the parameters are listed in Table III. Recall that the total link

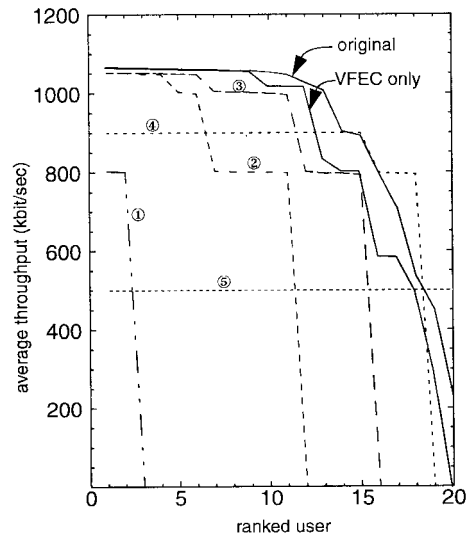


Fig. 15. Power control combined with fixed FEC (see Table III for description of fixed error correction used in each curve) with comparisons.

bandwidth for each user is 2000 kbit/s, the net data bandwidth after applying FEC is then $(2000/\text{FEC expansion factor})$ kbit/s.

The simulation results are shown in Fig. 15, from which we observe a significant overall improvement in throughput as we increase the error correction capability from 2-bit error (shown by ①) to 4-bit error (shown by ②), and finally to 5-bit-error (shown by ③) where the throughput reaches the maximum. As we move further in error correction capability, the system begins to shift from interference limited to bandwidth limited. This is clearly illustrated by ⑤ when the extremely powerful 11-bit error correction code is applied. In this case, users are immune to interference and noise so that the throughput for all users is nearly constant. However, the cost of this high reliability is the low transmission rate, which is near 500 kbit/s for each user. The original algorithm that has power control/VFEC/scheduling is also shown in the figure for comparison purposes; this is indicated by "original." The algorithm that has just VFEC, is indicated "VFEC only." Clearly, when power control is combined properly with an

TABLE III
FEC PARAMETERS

code type (n,k,t)	bandwidth expansion factor	data bandwidth (kbit/sec)	index used in the Figure 15
(63, 51, 2)	1.24	1610	①
(63, 39, 4)	1.62	1238	②
(63, 36, 5)	1.75	1143	③
(63, 30, 6)	2.1	952	④
(63, 16, 11)	3.9	507	⑤

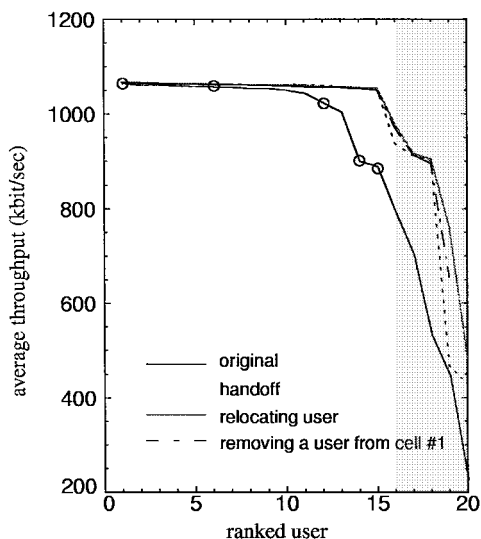


Fig. 16. Effectiveness of two handoff schemes.

FEC code, there can be a significant improvement over the power control only system.

6) *Handoff Strategies*: The system has now been investigated as whole where users were not distinguished from one another. In other words, we have not looked closely at the parameters that affect individual user’s performance, since aggregate throughput was used as the metric. In this section, we explore how a given user’s throughput can be influenced by location, cell population, and interference. Moreover, we want to know if a technique such as handoff is effective for alleviating users from excessive interference.

Let us go back to the original system where power control, VFEC, and scheduling are all applied; the overall system performance is indicated by “original” in Fig. 16. The users in the shaded region are clearly the ones with the poorest performance. It is not surprising that all these users reside in the most crowded cell, i.e., cell #1. However, the throughput for users in the second most crowded cell, cell #2 (which is also the cell experiencing the most interference) are marked by the circles. As we can see, cell #1 users suffer a significant quality degradation even it has just one more user than the next crowded cell.

One strategy to relieve the hot spot is to initiate handoff. The handoff here is slightly different from the conventional one, where it is only initiated when a user is at the boundary of two adjacent cells. In our situation, handoff is used as a method for alleviating radio network congestion. This goal is

achieved by letting a user in the most congested cell, cell #1 in our case, to communicate directly with its least crowded neighbor cell, cell #0. The user is chosen to be the one who is closest to cell #0, who is also the one furthest from cell #1.

The result of this handoff is compared with the case where the same user from cell #1 is removed and then randomly placed into cell #0. Physically relocating a user would definitely yield better performance than merely reassigning basestation. However, the result shown in Fig. 16 indicates very small difference. Notice both techniques yield a significant performance gain (that can be as large as 350 kbit/s for some users) over the “original” algorithm where a user only communicates with the closest basestation.

7) *Study of FEC Family*: The decision of using a single family of BCH code (i.e., codes that share the same block length) for VFEC purposes comes purely from implementation practicality [9]. Compared to convolutional code or punctured code, block codes offer a much wider range of error correction capabilities at reasonable implementation cost. For codes having the same block length, many building blocks such as syndrome calculation and polynomial manipulation can be shared because they all operate on the same finite field. In this section, we approach the BCH code from the system level by deciding the most suitable family of BCH codes for VFEC purpose.

From a purely performance standpoint, longer codes have better code efficiencies than shorter ones for the same percentage of bandwidth redundancy. In addition, longer codes offer greater selections for obtaining the right code to adapt to the time-varying channel and traffic. However, as the block length increases, more decoders have to be implemented; furthermore, as a code corrects more bit errors, the implementation complexity increases as well. Just as an example, one step in the decoding procedure (proposed by W. Peterson) involves computing the inverse of a $t \times t$ matrix, where t is the number of correctable errors; the complexity for this step is in the order of t^3 .

From these considerations, we study four families of BCH codes with block lengths equal to 15, 31, 63, 127 bits; these are all primitive codes. The code redundancy and error correction capability for each are listed in Table IV. The algorithms that use power control, VFEC and scheduling is simulated for each one of these four families of codes. The results are shown in Fig. 17.

It is expected that there will be a performance gap between the 127-bit family and the 15-bit family, since in the later case only three nontrivial error correction codes are available

TABLE IV
ERROR CORRECTION CAPABILITIES AND REDUNDANCIES FOR FOUR FAMILIES OF BCH CODE

block length	# of information bits	# of correctable bit errors
15	15, 11, 7, 5, 1	0, 1, 2, 3, 7
31	31, 26, 21, 15, 11, 6, 1	0, 1, 2, 3, 5, 7, 15
63	63, 57, 51, 45, 39, 36, 30, 24, 18, 16, 10, 7, 1	0, 1, 2, 3, 4, 5, 6, 7, 10, 11, 13, 15, 31
127	127, 120, 113, 106, 99, 92, 85, 78, 71, 64, 57, 50, 43, 36, 29, 22, 15, 8, 1	0, 1, 2, 3, 4, 5, 6, 7, 9, 10, 11, 13, 14, 15, 21, 23, 27, 31, 63

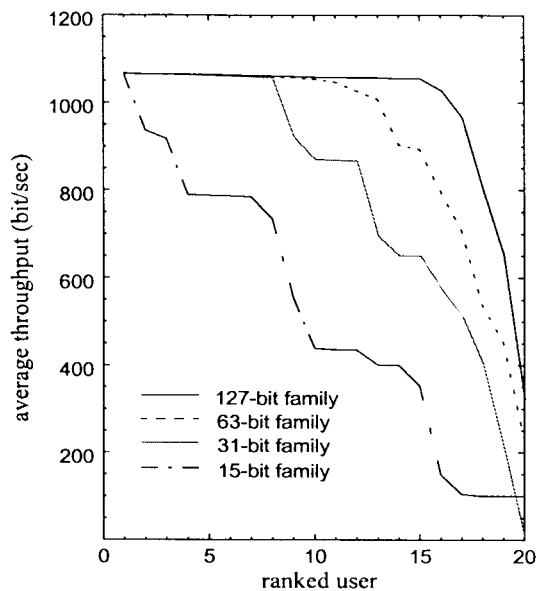


Fig. 17. System performance with respect to four families of BCH codes.

as opposed to 17 in the 127-bit family. However, the net performance gain only decreases by a factor of two as we move up to larger codes, implying extremely fine control of bandwidth redundancy may not be necessary, especially when the practical implementation constraints are considered.

V. SUMMARY AND CONCLUSIONS

The first half of the paper investigated design and control issues faced by multimedia indoor CDMA systems. We have proposed a distributed algorithm that fully utilizes the system resources such as bandwidth and transmit power. Our algorithm integrates three techniques for multimedia downlink transmission; these techniques are power control, VFEC, and scheduling. As a multimedia system is designed to satisfy users, our objective is to maximize the overall user satisfaction, which we call the “system utility.”

The system optimization problem is divided into a hierarchy of three levels: user level, cell level, and system level. This partitioning allows us to localize the problem, so that we can perform independent and local optimizations for each user and then each cell. The optimization results are then presented to the next level in the hierarchy. At the topmost level, the system level, we allocated cell power to one set of cells while keeping their neighboring cell power levels constant. The cell power level is determined based on its contribution to the overall utility for its entire neighborhood of seven cells. This layering

approach yields a distributed algorithm, which is essential given any practical wireless system has to be scalable.

The second half of the paper is devoted to simulation based evaluation of the system performance with a simple cell structure and uniformly distributed users. In the simulations, we focus on a special case where each data type has a BER requirement: data are only transmitted when the expected received BER meet this requirement. This constraint is adopted by many existing wireless systems, expressed in terms of received SNR. Notice this is equivalent to having a step utility function with respect to BER.

The results showed the performance for an individual user is highly dependent on the physical location, which agrees with our intuition. However, the system performance showed very little difference (in terms of the aggregate throughput) as we varied the traffic model for each data type from a highly bursty Markov chain to constant bit-rate traffic, as long as these traffic sources have the same average data rate and packet size.

Later we studied the effectiveness of VFEC verse power control. The technique of VFEC showed to be significantly superior than power control. This is true even when the average traffic demand is heavy, i.e., greater than half of the available link bandwidth. In the study of handoff, our results showed handoff improves the system significantly over the situation when users communicate with the nearest basestation. Finally, we examined four families of BCH codes (block length = 15, 31, 63, 127) to study the tradeoff between implementation complexity and variable error correction capability finding that the shorter codes provide sufficient variability to be effective.

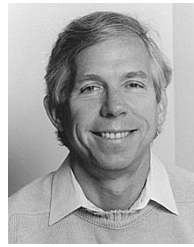
REFERENCES

- [1] S. Sheng, A. Chandrakasan, and R. Brodersen, “A portable multimedia terminal,” *IEEE Commun. Mag.*, pp. 64–75, Dec. 1992.
- [2] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver, and C. E. Wheatley, III, “On the capacity of a cellular CDMA system,” *IEEE Trans. Veh. Technol.*, vol. 40, pp. 303–312, May 1991.
- [3] R. Yates, “A framework for uplink power control in cellular radio systems,” *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1341–1347, Sept. 1995.
- [4] A. Goldsmith, “Design and performance of high-speed communication systems over time-varying radio channels,” Berkeley: Electronic Research Laboratory, Berkeley, CA, no. UCB/ERL M94/75.
- [5] A. Viterbi, *Principle of Spread Spectrum Communication*. Reading, MA: Addison-Wesley, 1995.
- [6] E. Polak, *Computational Methods in Optimization*. New York: Academic, 1971.
- [7] E. Berlekamp, *Algebraic Coding Theory*. Laguna Hills, CA: Aegean Park, 1984.
- [8] W. W. Peterson, *Error Control Coding*. Cambridge, MA: MIT Press, 1972.
- [9] Y. Lu and B. Brodersen, “Implementation architecture of a variable forward error correction decoder,” in *Proc. IEEE Globecom 98*, Nov. 1998.



Yuming Lu received the B.S. degree (with honors) in mathematics and physics from the State University of New York, Stony Brook, in 1991. From 1991 to 1998, she was a graduate student in the Department of Electrical Engineering and Computer Science at the University of California, Berkeley, where she received the M.S. and Ph.D. degrees.

She was a Member of Technical Staff in the Radio Technology Group at Bellcore, NJ, for six months during 1994, where she worked on the first-generation PACS. She is currently working on smart antenna technology for Metawave Communications, Redmond, WA.



Robert W. Brodersen (M'76–SM'81–F'82) received the B.S. degrees in electrical engineering and mathematics from California State Polytechnic University, Pomona, in 1966, and the M.Eng., M.S., and Ph.D. degrees from Massachusetts Institute of Technology, Cambridge, in 1968 and 1972, respectively.

He is a Professor in the Department of Electrical Engineering and Computer Sciences (EECS) at the University of California, Berkeley. He joined the EECS Department faculty in 1976. From 1972 to 1976, he was a member of Technical Staff, Central Research Laboratory, Texas Instruments. In addition to teaching, his present research focus is the application of integrated circuits as applied to personal communication systems, with emphasis on wireless communications and low-power design. He is the author or coauthor of over 60 journal publications, 120 published conference papers and author, coauthor, editor, or contributor to 14 books, including *Anatomy of a Silicon Compiler* (Norwell, MA: Kluwer, 1992) and *Low Power Digital CMOS Design* (Norwell, MA: Kluwer, 1995). He is the holder of three patents.

Dr. Brodersen is a member of the National Academy of Engineering. He won conference best paper awards at Eascon (1973), the international Solid-States Circuits Conference (1975), and the European Solid-States Circuits Conference (1978). In 1979, he received the W. G. Baker Award for the outstanding paper in the IEEE Journals and Transactions. He was coreipient of the Morris Libermann Award of the IEEE in 1983 for "outstanding contributions to an Emerging Technology." He received the best paper award in the IEEE TRANSACTIONS ON COMPUTER-AIDED DESIGN OF INTEGRATED CIRCUITS AND SYSTEMS in 1985 and the best tutorial paper of the IEEE Communications Society in 1992. He was appointed the first holder of the John R. Whinnery Chair in the EECS Department, University of California, Berkeley, in September 1995. In 1997, he received the distinguished IEEE Solid-State Circuits Award "for contributions to the design of integrated circuits for signal processing systems." He was the national chair of Information Science and Technology (ISAT) Study Group, sponsored by the Institute of Defense Analysis, Washington, DC, from 1992 to 1994. He currently serves on several committees with the National Academy of Sciences, Washington, DC. He has served on the editorial board or as reviewer for numerous scholarly publications, including the IEEE JOURNAL OF SOLID-STATE CIRCUITS, IEEE TRANSACTIONS ON VERY LARGE SCALE INTEGRATED (VLSI) SYSTEMS, IEEE PERSONAL COMMUNICATIONS MAGAZINE, and WIRELESS PERSONAL COMMUNICATIONS (Kluwer).