Multi-Channel Packet Resource Allocation for Multimedia Support

An Mei Chen and Ramesh R. Rao

Department of Electrical & Computer Engineering
University of California, San Diego
9500 Gilman Drive, La Jolla, CA 92093-0407, USA

Abstract—In a wireless packet network that supports an integrated mix of multimedia traffic, a large variety of mobiles with different service rates will need to be accommodated efficiently in an unified bandwidth-on-demand manner. We have proposed a robust resource allocation protocol to utilize the multiple parallel packet channels to support mixes of multi-rate traffics. The protocol incorporates a spectrally efficient accessing scheme and across-cell resource sharing technique to effectively multiplex multi-rate packet transmission. The protocol simplifies the network setup and system operation and at the same time reduce the system call blocking and dropping probabilities for all different rate classes of mobiles and has very high channel utilization efficiency. In this paper, we study the effectiveness of the protocol when the channel is subject to Rayleigh fading. We study a scenario when the mobile is downloading a number of media objects and compare the delay performance with the single packet channel. Analytical results indicate that the performance of the protocol is highly dependent on the Doppler frequency and the fading margin and significant gain in the system performance can be achieved for a given set of channel conditions and Doppler frequencies.

1. INTRODUCTION

Future generations of Personal Communication Networks (PCNs) are expected to support multimedia applications; thus, these networks should be able to service integrated traffic generated by different types of sources such as voice, video, and data. It is essential that these integrated traffic be accommodated in a bandwidth efficient manner with the Quality of Service (QoS) requirements of various types of applications guaranteed.

In [2] and [3], we proposed a cellular network infrastructure and a resource allocation protocol for accommodating a large variety of mobiles with different service rates and mobility profiles. The proposal considered a pico-cell network infrastructure with distributed control for providing service coverage in an urban area with high and varied traffic demand. To facilitate network setup in such areas, a flexible base station placement scheme was employed, where the base stations are placed corresponding to expected traffic, and each base station works autonomously without the consideration of the frequency reuse pattern. It is envisioned that these base stations will be small and low-powered and can be easily mounted on a street lamp post.

Due to its size and the radiated power requirement, base station's coverage area will be micro- or pico-cell. In a high traffic demand area, many base stations will be deployed to accommodate the heavy traffic load. The high density of base stations and their irregular placement will generate cells with overlapping coverage area. Mobiles in the overlapping coverage area can be served by its current base station or adjacent base station(s) responsible for the overlapping coverage.

Within the overlapping coverage area, if no channels are available at one base station, the adjacent base station(s) may have available channels that can provide service for calls that would otherwise be blocked. Beside reducing the blocking probability of calls, the coverage area can be exploited to serve heterogeneous traffic with different data rates and QoS requirements. To accommodate integrated traffic, existing base station must be redesigned to handle data rates ranging from few Kbps as in a voice call to several hundred Mbps as in video conferencing. Hence, the hardware complexity and the size of the base stations will have to be significantly increased. However, since the types of applications to be expected in a service area are not known a priori, engineering base stations that can service all types of application with different data rates and QoS requirements will waste valuable resource and up-front revenue.

The alternative scheme presented in [2] and [3] can accommodate integrated traffic with different QoS without requiring additional modification of those existing base stations that currently support only low data rate calls. In our proposal, the base station can be engineered with only a single air interface for a basic data rate, i.e. voice. Mobile terminals running applications with data rates that are multiples of basic rate will utilize multiple frequency channels to meet their data rate requirements. Since the cell sizes are small and the mobiles are within communication range of more than one base stations, mobiles can dynamically select which base station and how many base stations to establish multiple channels with. The selection process is based on the mobiles’ QoS requirements and the overall system traffic load balancing objective. One advantage of this across-cell resource sharing scheme is to prevent few mobiles with high data rate from hogging all the available channels in a cell and thereby block all future new and handoff calls. Mobile users with high mobility profile can utilize the multiple channels to provide a gradual handoff of high data rate calls, since not all channel connections need to be handed off at once when the mobile users move. For loss sensitive applications, redundant data can be sent on multiple channels to provide diversity against poor channel condition and handoff failure. For mobiles that are outside the range of communication of multiple base stations, multiple channels can be established to the same base station to meet the mobiles high data rate requirements. Studies in [3] showed that integrated data traffic can be accommodated efficiently.

This research was supported in part by the DDR&E Focused Research Initiatives (FRI) on Wireless Multimedia Networks under ARO grant DAA H04-95-1-0248.

0-7803-4320-4/98/$5.00 © 1998 IEEE

1770 VTC '98

Authorized licensed use limited to: Univ of Calif San Diego. Downloaded on April 16,2010 at 20:38:20 UTC from IEEE Xplore. Restrictions apply.
dated with low blocking and handoff failure probabilities and at the same time with high system utilization efficiency. In this paper, we compare the system performance of a single high data rate channel versus multiple low data rate channels in terms of delay. We employ a number of parallel channels model to analyze the performance of the multi-channel packet resource allocation protocol assuming that individual channel is in a Rayleigh multipath fading environment.

The paper is organized as follows. In Section II, we describe the system architecture and resource sharing strategy for multirate wireless packet networks. In Section III, we present some analytical performance evaluation of the multi-channel packet resource allocation protocol. The numerical results are shown in Section IV along with the discussion of system design issues and on-going work.

II. SYSTEM DESCRIPTION

The multi-channel packet resource allocation protocol and network architecture utilize the medium access technique called Time-Frequency Slicing [6]. This scheme is a spectrally efficient and cost effective technique that allow multiple users with different date rate requirements to share a communication medium. With a Time-Frequency Slicing technique, time is slotted and users can modulate a signal in one or more of the available frequency bands over one or more time slots. Unlike [6], we consider a system where the users are assigned no more than one time slot. Low data rate users can fill one time slot and one frequency band, while high data rate users can fill one or more frequency bands in the same time interval. A Basic Rate Unit (BRU), which consists of one frequency band allocation for one time slot, is the minimum amount of resource available to a user. We assume that a BRU is designed to transmit one packet (ATM cell).

Let us assume that there are \( K_{\text{max}} \) different classes of users in the network. The grouping of the total user population to classes is based on the type of applications supported by the mobile terminals, where applications belonging to a group have the same average service rate. Here we use the term average service rate to denote the average number of packets that the source generates per time slot. On the other hand, the term transmitter rate represents the number of BRUs can be supported by a transmitter. For example, a source that generates an average of 5 packets per time slot can employ a transmitter with 5 BRUs transmitter rate. A mobile terminal that can support 5 BRUs implies that the transmitter can modulate 5 or more frequency bands in a time slot.

Different type of transmitter and receiver structures can be used to realize the need. In Fig. 1 and 2, we show an example of transceivers using Code Division Multiple Access (CDMA) technology, since CDMA is known for its robustness to several serious impairments of wireless communication channel and system such as multipath fading, interference, frequency selectivity, and synchronization. Figure 1 and 2 show a multicarrier DS-CDMA transmitter and receiver structures that can be used to implement the Time Frequency Slicing technique, where the high data rate source with serially incoming bit stream is converted to \( k \) parallel bit streams. Each parallel bit stream is spread with an unique direct sequence and modulated on a different carrier. The converter effectively converts a single high rate bit stream to multiple parallel low rate bit streams. The advantage of multicarrier signalling is that it does not require a continuous spectrum for operation.

We recently noticed that the CDMA Development Group (CDG) is incorporating similar ideas to ours in the next generationITU 3G Standards [5]. The main objective of the CDG is to have the next generation communication networks to be fully compatible with the existing IS-95 Standard, where no change of infrastructure hardware is required. In the proposal, both the current IS-95 users and the high data rate users share the same spectrum. The concept of both fundamental and supplemental channels was introduced to support the high data rate users. In this scheme, the high data rate user will use the fundamental channel as the same way as the current IS-95 users with the additional information is sent via supplemental channels to support its high data rate. Multicarrier CDMA forward link signalling is used, where these signals are orthogonal to IS-95 forward link signals. Wideband information is transmitted on multiple 1.25 MHz carriers. Wideband data employs QPSK modulation and QPSK spreading, where higher data rate users are supported with shorter Walsh codes. These data rates will be made available to users depending on the user’s mobility and location.

A. Dynamic Base Stations Selection

The proposed flexible placement scheme allows new base stations to be placed according to the expected traffic profile. In the hot traffic area, the high density of base stations and the irregular placement of them generate many cells with large area of overlapping coverage. Figure 3 shows an overlapping coverage of base stations in a heavy traffic area where base stations are concentrated to accommodate the traffic load. The coverage area of the base station is the area in which the base station can provide and receive sufficient signal strength with acceptable fidelity for both the uplink and the downlink. In the overlapping area, more than one base stations can provide acceptable quality strength for the mobile. In [3], we proposed a scheme in which a mobile autonomously and dynamically selects a set of base stations to serve its call, where the selection procedure is based on the mobile’s QoS requirement and the overall system load balancing objective.

Each base station in the system is assumed to periodically broadcast a Base Station Status (BSS) signal indicating the amount of its available BRUs. To make a call, a mobile in the overlap area would lock on a set of base stations that have adequate signal quality to meet its application QoS requirement. From the set of selected base stations, the mobile would rank these base stations according to the number of available BRUs starting with the base station that has the most available resource. A mobile would establish a single channel to each base station in the set or up to its service data rate requirement. If the number of base stations in the list is less than the mobile service rate requirement, the mobile will establish more than one connection to one or more base stations. Using this procedure, each mobile not only optimizes its signal quality and reduces the frequency of handoff by connecting to those base stations with strong signal that are close to it but also performs system traffic load balancing. Mobiles can dynamically configure these channels to meet QoS needs.

B. Multi-Channel Packet with n-Retransmission

We now analyze the effectiveness of the link-level diversity by determining the optimal number of parallel channels should be used in downloading a fixed number of media objects when the channel is subjected to multipath fading. We assume that a mobile user needs to download a number of media objects that totaled up to \( M \) packets. A packet that is lost or received in error must be re-transmitted until it is successfully received. It is considered that the transmitter receives feedback from receiver of lost or
error packets. Once the packet is lost, it is assumed that n previous packets must be retransmitted. If n = 1, then only the lost packet needs to resend. During poor channel conditions, many errored packets. Once the packet is lost, it is assumed that the delay exceeds twice the reset and transmission will start over again. If the mobile terminal utilizes n parallel channels, then the number of packets transmit on each channel is \( M/n \), where each low data rate channel applies the same retransmission protocol as the single high data rate channel.

### III. Performance Analysis

In this section, we analyze and compare the delay performance of multi-channel and single channel packet link protocols in the presence of multipath fading. A cellular system with no inter-cell interference is considered. As we know, mobile radio channels are severely affected by time-varying losses due to distance, shadowing, and multipath fading. While the variation in the losses due to distance and shadowing is relatively slow, the variation due to multipath fading is quite rapid. In this paper, we assume that the slow varying signal losses due to distance and shadowing are perfectly compensated through power control, whereas the rapid variations due to multipath fading remain uncompensated. The fading envelope due to multipath often follows a Rayleigh distribution. We will employ channel model with Rayleigh fading to analyze the performance of our system. Our analysis concentrates only one the first hop, from the mobile to the base station (or base to mobile), since the wireline transmission medium is relatively error-free. We also assume that the propagation delay from mobile to base is negligible since the size of a pico-cell is on the order of few meters.

#### A. A Burst-Error Channel Model

In [8], a model was derived to describe the binary process of packet success/failure, where it was shown that, for most situations typically encountered in personal communications and mobile computing, a first-order binary Markov process adequately characterizes the error process at the packet level for transmission on a fading channel with additive Gaussian noise.

The pattern of the packet errors on a wireless channel is characterized by a first-order Markov process with the following probability transition matrix

\[
P(x) = \begin{pmatrix} p(x) & q(x) \\ r(x) & s(x) \end{pmatrix},
\]

\[
P(1) = \begin{pmatrix} p & q \\ r & s \end{pmatrix}^n,
\]

where \( p(x) = 1 - q(x) \), \( r(x) = 1 - s(x) \) is the probability that the packet is successful given that the packet was successful (unsuccessful). Thus, \( 1/r \) is the average length of the geometrically distributed burst of block errors. We denote 0 and 1 for good and bad states respectively.

The above parameters were computed in closed form for the case of varying Rayleigh fading channel in [8]. Let \( F \) be the fading margin. The parameters can be computed using the following relations

\[
e = 1 - e^{-1/F}
\]

### B. Markov Evolution

The multiple-packet channels with n-Retransmission described in Section II can be represented by a Markov chain. Let \( X^i(n) = (l^i(n), j^i(n)) \), with state space \( \{(i, j^i), 0 \leq l^i \leq 1, 1, 2, \ldots, M^i\} \), for \( i = 1, 2, \ldots, N \) and where \( l^i(0) = 0 \) and \( l^i(1) = 1 \) denote a bad and good state of channel i in time slot n respectively, and \( j^i(n) \) is the remaining number of packets to be transmitted on link i, and \( M^i \) is the total number of packets needed to be transmitted on link i, and \( N \) is the total number of parallel packet links. If \( X^i(n) = (0, j^i) \), \( j^i \geq 0 \), a packet is transmitted and received successfully. Link i is said to complete the transmission of all of its packets if \( X^i(n) = (l^i(n), 0) \) for \( i = 0, 1 \). On the other hand, if \( X^i(n) = (1, j^i) \), \( j^i \geq 0 \), no packet can be transmitted or received correctly.

Consider a channel i, the transition probabilities for multi-channel packet with n-Retransmission are follows:

\[
\begin{array}{ccc}
(1, j^i + 1) & (0, j^i - 1) & (1, j^i + n) \\
0 & p & q \\
1 & r & 0
\end{array}
\]

where \( j^i \geq 1 \), \( I(n) = 1 \) is an indicator function at \( n = 1 \), and \( P[X^i(n+1) = (1, 2M^i)]X^i(n) = (1, 2M^i)] = s \). The state \( X^i(l^i, 0) \) is the absorbing state with \( P[X^i(n+1) = (l^i, 0)]X^i(n) = (l^i, 0)] = 1 \), for \( l^i = 0, 1 \) and \( n > 0 \). This state denotes that the link i has completed transmission.

### C. Delay Analysis

In this section, we compute the average delay in terms of the number of time slots taken to transmit \( M^i \) packets on channel i in the presence of Rayleigh fading. The performance of a single packet link versus multiple parallel packet links is compared. To simplify the notation used in the analysis, we renumber all states using an integer index, starting with 0 for the absorbing state. Let \( p_{ij} \) denote the corresponding transition probability in going from state j to state i. The average time taken to transmit all \( M^i \) packets on link i is equivalent to the average time until absorption. At the start of transmission, the transmitter will find the channel in the good or bad state with equal probability, i.e. \( P[X^i(0) = (0, M^i)] = P[X^i(0) = (1, M^i)] = 1/2 \). Let \( T^i \) be a random variable that represents the total time taken to transmit all packets.
$M^i$ packets on link $i$. Computing $T^i$ is equivalent to computing the first passage time to state 0, where

$$T^i = \inf\{n \geq 0 : X^i(n) = 0\}. \quad (8)$$

The probability distribution of $T^i$ can be computed by invoking the following theorem:

**Theorem 7** Let $\alpha_j^i(n) = P[T^i = n|X^i(0) = j]$ and $\alpha_j^i(n) = P[T^i < n|X^i(0) = j]$. Then $\alpha_j^i(n)$ can be computed using the following recursive relation:

$$\alpha_j^i(n) = p_j0 + \sum_{i=1}^{\infty} p_{ji} \alpha_j^i(n-1) \quad j \geq 1, n \geq 1, i = 1, 2, \ldots, N \quad (9)$$

Since each packet link is established on a different frequency channel to the same or different base stations, it is reasonable to assume that the fading process and hence the total retrieval time of each link is independent from each other, i.e. $T^i$s are mutually independent. We further assume that each channel is subjected to similar fading condition. Hence, $T_1, T_2, \ldots, T_N$ are mutually independent and identically distributed random variables. Let $T_{1:N} \leq T_{2:N} \leq \ldots \leq T_{N:N}$ denote the order statistics obtained from the relations of the random variables $T_1, T_2, \ldots, T_N$. The maximum average delay incurred using $N$ parallel packet links can be computed as follows:

$$E[T_{N:N}] = N \sum_{n=0}^{\infty} n[u(n)]^{(n-1)} \alpha(n). \quad (10)$$

To compare the improvement of the multiple parallel packet links to that of the single link in term of reducing the average retrieval delay time, we define a delay speed-up ratio $S_d(N)$ as follows:

$$S_d(N) = \frac{E[T_{1:1}]}{E[T_{N:N}]} \quad (11)$$

where $E[T_{1:1}]$ is the delay experienced using a single packet link. (Note that the number of packets needs to be retrieved on a single link is equal to the total number of packets of all $N$ parallel links.) The speed-up ratio measures the gain in using more than one packet link if the $S_d(N) > 1$. The value of $S_d(N)$ will guide the selection of the number of channels to be used in a particular fading environment.

**IV. Numerical Results**

In this section, we present some numerical results for the Rayleigh fading channel, in terms of the fading margin $F$, and of the normalized Doppler frequency, $f_D T$ (where $f_D$ is the Doppler frequency and $T$ is the slot duration) with 1-Retransmission.

The sensitivities of the speed-up factors to the fading margins and the normalized Doppler frequencies is studied in Figure 4 and 5. We assume that the packet is an ATM cell (424 bits) and the single channel data rate can support 256Kbps. Figure 4 shows the delay speedup factor $S_d$ vs. the fading margin, $F$ (in dB), for various number of parallel channels with the normalized Doppler frequency of $f_D T = 0.0001$, which corresponds, for example, a transmission of an ATM cells over a single 256Kbps channel at 900MHz with the mobile user speed approximately zero. For $n$ parallel channels, the normalized Doppler frequency is $n$ times the single high data rate channel’s Doppler frequency to reflect that the data rate of each of the parallel channel is $n$ times slower than the single high data rate channel (or the time slot is $n$ times longer). For a small value of fading margin, the gain in speed up is less than those with higher values of fading margin. The curves indicate that at the high values of fading margin, having three or more parallel connections can reduce the delay time by half. Fig. 5 shows the speedup factor for the case when the mobile user is traveling at the speed of 65Km/h. The opposite behavior is observed as compared to the case when the mobile is stationary. At the low values of fading margin, the gain factor ranges from 2 to 4 times of the single channel for 2 to 5 parallel packet channels. At the high values of fading margin, we see that the law of diminishing return takes effect. Having more parallel channels provides relatively little gain in term of performance.

From these results, we can conclude that the system performance is sensitive to the channel quality (i.e. the fading margin) and the channel memory (i.e. the Doppler frequency). We derive some useful insight regarding the usage of many low data rate channels for providing link diversity. We observed that having more than one low data rate channels out performs a single high data rate channel. We found that having many narrowband channels is not always better than few narrowband channels. In particular, we saw that the little gain was obtained when the user is traveling at average speed of 65Km/h and channel’s fading margin is high. One explanation for this is that when the channel condition is relatively good, almost all transmitted packets will be successfully received and hence diversity offers relatively little benefit. Since additional narrowband channel will cost extra hardware, it is necessary that the optimal number of channels is used. Hence, results obtained in this paper can be used to optimally configure the size of BRU. For instance, different data rate applications and their operation environment should be studied to derive an optimal size of BRU where all applications QoS needs are best met.

The results presented in this paper can be applied to web browsing applications. Due to the user friendly graphical user interface and the ease at which information can be accessed, it is no doubt that the Web has become one of the most popular service in the Internet today. It is expected that the future man-machine interface will be done exclusively via some web browser, i.e. **Netscape**. The Hyper-Text Transfer Protocol or known as HTTP is an application that is used in the web browser to request and retrieve information from distributed servers. The HTTP model is extremely simple; the client establishes a connection to the remote server, then issues a request. The server then processes the requests, returns a response and then closes the connection. In practice, accessing a single web page include numerous requests and responses. Request and responses could be from a single server or multiple servers. Retrieval of a complete web page requires separate requests for text and each embedded image. Each requested object constitutes a separate connection; hence, downloading a web page would represent multiple connections setup.

Currently both versions HTTP1.0 and HTTP1.1 employ the Transmission Control Protocol (TCP) as a transport layer protocol [1] and [4]. For each of the embedded images in the web page, a TCP connection will be established to the same or different servers depending on where the image resides. In the current version of web browser **Netscape**, the configuration of the number of simultaneous connections is left to the user. If the number of TCP connection is configured to be less than the number of media objects, then multiple objects are pipelined and send via one TCP connection. Since the HTTP opens more than one connections in downloading a single web page, the web browsing appli-
cation naturally fits into the multi-channel packet protocol. The complexities such as managing multiple connections, segmentation, re-assembling, and ordering of data packets introduced by the multi-channel packets protocol are transparently handled by both TCP and HTTP. Hence, existing application software could benefit from the results obtained in this work, by optimally configured the number of TCP connections for a particular channel condition.

REFERENCES


