

Joint Design of Adaptive Channel Coding and Multiple Access Control for Integrated Voice and Data Services in a Cellular Wireless Network—With Contention Queue

Vincent K. N. Lau and Yu-Kwong Kwok

Email: {knlau, ykwok}@eee.hku.hk

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
Tel: +852-2857-8483, Fax: +852-2559-8738
Corresponding Author: Vincent K. N. Lau

Abstract—A new MAC protocol¹ employing a joint design of the multiple access control (MAC) layer and the physical layer is proposed for cellular wireless systems with integrated voice and data services. This is a novel concept and is shown to achieve very significant gains (in capacity, packet loss and delay) compared to existing protocols through the exploitation of the synergy between the two protocol layers.

I. INTRODUCTION

In this paper, we focus on exploiting the synergy between the multiple access control (MAC) layer and the physical layer for the reverse-link direction of a cellular wireless network, providing integrated voice and data services. A variable throughput, channel-adaptive physical layer is employed to protect information across the hostile wireless channel. With the help of a low capacity feedback channel (carrying channel state information—CSI), the level of redundancy and hence the offered throughput is variable depending on the instantaneous CSI [6], [5]. Hence, the bit-error-rate (BER) is maintained at a targeted level over a range of channel condition. Figure 1 illustrates the design of the physical layer. Transmission modes with *normalized throughput* varying from 1/2 to 5 are available depending on the channel condition.

The MAC layer is responsible for resource allocation to a number of users contending for service in the uplink direction. Traditional MAC previous protocols are designed and analyzed based on the as-

¹This research was jointly supported by research initiation grants from the HKU CRCG under contract numbers 10202523 and 10202518, a HKU URC seed grant under contract number 10203010, and by a grant from the Hong Kong Research Grants Council under contract number HKU7124/99E.

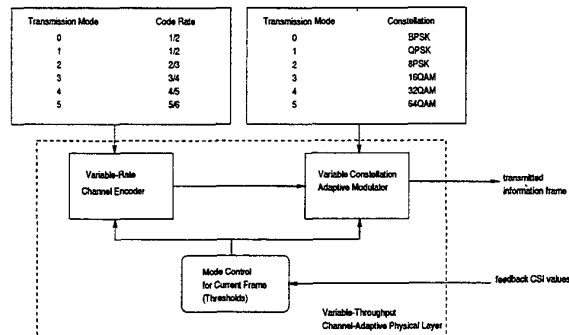


Fig. 1. A conceptual block diagram of the variable throughput channel adaptive physical layer.

sumption that packet transmission through the physical channel is error-free. This is a valid assumption in fixed-line transmission environment but obviously failed to apply in wireless environment where the error rate of the physical layer is high and the physical layer performance is time-varying. This motivates our new design by matching the MAC protocol to the physical layer in order to achieve a better resource utilization. In this paper, we consider D-TDMA based MAC [1], [2] where a frame is divided into request contention phase and traffic phase. The novelty of the proposed scheme, called CHARISMA (Channel Addaptive Reservation-based Isynchronous Multiple Access), stems on the fully adaptive design in both the physical layer and the MAC layer. Unlike classical D-TDMA protocols, the user contention requests are gathered by the base-station in the first phase of the time frame without immediately announcing the information slots assignment right after

each contention mini-slot. After all requests are received, the information slots are assigned to the users based on their respective CSI ranking. Synergy is therefore achieved by providing an additional input to the MAC protocol, namely the channel state information (CSI) from the physical layer. To demonstrate the performance gain derived from the synergy, we compare the performance of our proposed scheme with two baseline systems as shown in Figure 2. Baseline System-I is also a D-TDMA based MAC layer but a traditional fixed throughput physical layer is used. In Baseline System-II, a variable throughput channel-adaptive physical layer is incorporated with a D-TDMA based MAC layer but there is no interaction between the two layers [4]. Our extensive simulation results indicate that for the same quality of service level, the proposed CHARISMA almost doubles the system capacity.

Deriving synergy between protocol layers is a relatively unexplored research topic. In [4], an adaptive modulation reservation TDMA scheme is proposed. However, such scheme is essentially a simple cascade of the MAC layer with an adaptive modulator and thus, no synergy is exploited².

II. VARIABLE THROUGHPUT CHANNEL-ADAPTIVE PHYSICAL LAYER

A. Wireless Channel Model

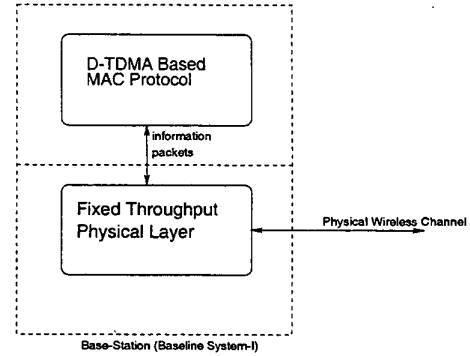
The wireless communication environment considered in this paper is the reverse-link situation of a wireless system where a number of mobile terminals contend to transmit information to a base-station. The wireless link between a mobile terminal and the base-station is characterized by two components, namely the *fast fading* component and the *long-term shadowing* component. Fast fading is caused by the superposition of multipath components and is therefore fluctuating in a very fast manner (on the order of a few msec). Long-term shadowing is caused by terrain configuration or obstacles and is fluctuating only in a relatively much slower manner (on the order of one to two seconds).

Let $c(t)$ be the combined channel fading which is given by:

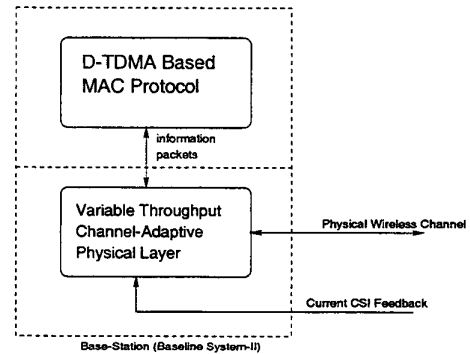
$$c(t) = c_l(t)c_s(t)$$

where $c_l(t)$ and $c_s(t)$ are the long-term and short-term fading components, respectively. Both $c_s(t)$ and $c_l(t)$ are random processes with a coherent time on

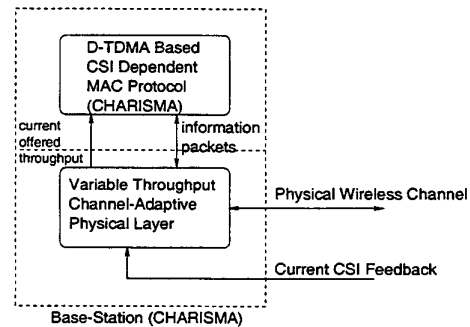
²The scheme in [4] is similar to our baseline System-II.



(a) baseline System-I (fixed-rate channel encoder)



(b) baseline System-II (adaptive channel encoder)



(c) baseline System-I (adaptive channel encoder with MAC interaction)

Fig. 2. Conceptual models of the physical and MAC layers.

the order of a few milli-seconds and seconds, respectively.

Short-Term Fading:

Without loss of generality, we assume $\mathcal{E}[c_s^2(t)] = 1$ where $\mathcal{E}[\cdot]$ denotes the expected value of a random variable. The probability distribution of $c_s(t)$ follows the Rayleigh distribution which is given by:

$$f_{c_s}(c_s) = c_s \exp\left(-\frac{c_s^2}{2}\right)$$

In this paper, we assume the maximum mobile speed is 80km/hr and hence, the Doppler spread [9], $f_d \approx 100\text{Hz}$. It follows that, the *coherent time*, T_c , is approximately given by:

$$T_c \approx \frac{1}{f_d} \quad (1)$$

which is about ten msec.

Long-Term Fading:

The long-term fading component, $c_l(t)$, is also referred to as the *local mean* [9], which, as shown by field test measurement, obeys the *log-normal* distribution, $f_{c_l}(c_l)$. That is,

$$f_{c_l}(c_l) = \frac{4.34}{\sqrt{2\pi}\sigma_l c_l} \exp\left(-\frac{(c_l(\text{dB}) - m_l)^2}{2\sigma_l^2}\right)$$

where m_l , σ_l are the mean (in dB) and the variance of the log-normal distribution, i.e., $c_l(\text{dB}) = 20\log(c_l)$. Since $c_l(t)$ is caused by terrain configuration and obstacles, the fluctuation is over a much longer time scale. Again, from field test results, the order of time span for $c_l(t)$ is about one second.

Since mobile terminals are scattered geographically across the cell and are moving independently of each other, we assume the channel fading experienced by each mobile terminal is independent of each other.

B. Frame Structures

Uplink frame is used for transmitting information from the mobile terminals to the base-station and as such, the *multiple access* problem has to be tackled. By contrast, downlink frame is used for transmitting information from the base-station to the mobile terminals and is, therefore, a *broadcasting* channel such that no contention is involved. Our focus in this study is on the multiple access aspect of the uplink.

In the uplink, a frame is divided into three subframes. They are the *request subframe*, *information*

subframe, and the *pilot symbol subframe*. Specifically, there are N_r mini-slots in the request subframe for voice requests reservation and data requests contention. Note that a data user is not allowed to make reservation in the sense that even if a data user successfully seizes an opportunity to transmit in the current frame, he/she has to contend again in the next frame if he/she has some more data to send. There are N_i information slots in the information subframe for the transmission of voice or data packets. Finally, there are N_b slots in the pilot symbol subframe, the function of which will be elaborated in detail later in Section III-B. The frame duration is 2.5 msec. Such a short frame duration has the advantage of shorter delay and is practicable in wideband systems [8].

On the other hand, a downlink frame is similarly partitioned into four subframes, namely the *acknowledgment subframe*, *poll-for-CSI subframe*, *information subframe*, and *announcement subframe*. The frame duration is also 2.5 msec and the number of slots in the subframes are given by N_r , N_b , N_i , and N_b , respectively. The functionality and operation of each subframe are described in Section III.

C. Variable Throughput Physical Layer

Redundancy is incorporated to the information packet for error protection. To exploit the time-varying nature of the wireless channel, a variable rate channel-adaptive physical layer is employed as illustrated in Figure 1. Channel state information (CSI), $c(t)$, which is estimated at the receiver, is fed back to the transmitter via a low-capacity *feedback channel*. Based on the CSI, the level of redundancy and the modulation constellation applied to the information packets are adjusted accordingly by choosing a suitable transmission mode. Thus, the instantaneous throughput is varied according to the instantaneous channel state. In our study, a 6-mode variable rate adaptive bit-interleaved trellis coded modulation scheme (ABICM) is employed [5]. Transmission modes with *normalized throughput*³ varying from 1/2 to 5 are available depending on the channel condition.

We assume the coherent time of the short-term fading is around ten msec which is much longer than an information slot duration. Thus, the CSI remains approximately constant within a frame and it follows that the transmission mode for the whole frame is determined only by the current CSI level. Specif-

³Normalized throughput refers to the number of information bits carried per modulation symbol.

ically, transmission mode q is chosen if the feedback CSI, \hat{c} , falls within the *adaptation thresholds*, (ζ_{q-1}, ζ_q) . Here, the operation and the performance of the ABICM scheme is determined by the set of adaptation thresholds $\{\zeta_0, \zeta_1, \dots\}$. In this paper, we assume that the ABICM scheme is operated in the *constant BER* mode [5]. That is, the adaptation thresholds are set optimally to maintain a target transmission error level over a range of CSI values. When the channel condition is good, a higher mode could be used and the system enjoys a higher throughput. On the other hand, when the channel condition is bad, a lower mode is used to maintain the target error level at the expense of a lower transmission throughput. Note that when the channel state is very bad, the adaptation range of the ABICM scheme can be exceeded such that the throughput (mode-0) becomes so low, making it impossible to maintain the targeted BER level. Given the above considerations about the channel state, the instantaneous throughput offered to the MAC layer, ρ , is also variable and is therefore a function of the CSI, $c(t)$, and the target BER, P_b , denoted by $\rho = f_\rho(c(t), P_b)$.

III. THE CHANNEL-ADAPTIVE MAC LAYER

In the following, we illustrate the potential benefit of exploiting inter-protocol layer synergy through our proposed D-TDMA based MAC protocol as an example. The wireless communication system consists of only voice and data terminals. As in most previous studies, it is assumed that voice packets are delay sensitive while data packets are delay insensitive.

A. Source Models

The wireless system considered in this paper is aimed to support integrated voice and data services. As such, we assume that there are only two types of mobile terminals, namely the voice terminal and the data terminal in the system. Voice packets are assumed to be delay sensitive while data packets are assumed to be delay insensitive. Thus, voice packets are labeled with *deadlines*. A voice packet will be dropped by a voice terminal if the deadline expires before being transmitted. Such packet dropping has to be controlled to within a certain limit (e.g., below 1% as indicated in [3]) in order that the quality of service to the voice users is still acceptable. The source and contention models are summarized below.

- **Voice Source Model:** The voice source is assumed to be continuously toggling between the talkspurt and silence states. The duration of a talkspurt

and a silence period are assumed to be exponentially distributed with means t_t and t_s seconds, respectively (as indicated by the empirical study in [7], $t_t = 1$, and $t_s = 1.35$). We assume a talkspurt and a silence period start only at a frame boundary.

- **Data Source Model:** The arrival time of data generated by a data terminal is assumed to be exponentially distributed with mean equal to one second. The data size, in terms of number of packets, is also assumed to be exponentially distributed with mean equal to 100 packets. Again we assume that the packets arrive at a frame boundary.

- **Terminal Contention Model:** As in most previous studies, to avoid excessive collisions, even if a voice or data user has some packets awaiting to be sent, the user will attempt to send a request at a request mini-slot only with a certain *permission probability*. The permission probability for voice and data users are denoted by p_v and p_d , respectively.

B. Operations of CHARISMA

In this section, we describe the design of CHARISMA, a novel channel state dependent MAC protocol that works closely with the variable throughput channel-adaptive physical layer (described in Section II). The operation of the CHARISMA protocol is divided into two phases, namely the *request phase* and *transmission phase*. In the request phase, mobile terminals which have voice or data packets to transmit will send a request packet in one of the request slots, governed by the respective permission probability. The request packet is very short, occupying only a mini-slot, as illustrated in Figure 4(a). It contains the mobile terminal ID, request type (voice or data), packet deadline, number of information packets desired to transmit as well as some pilot symbols. If more than one mobile terminals send request packets in the same request slot, collision occurs and all the request packets are lost (i.e., capture effect is not considered). After each request slot, an acknowledgment packet will be broadcast from the base-station through the acknowledgment slot in the downlink frame as illustrated in Figure 3(a). The acknowledgment packet contains only the successful request packet ID. Mobile terminals that fail to receive an acknowledgment will retransmit the request packet in the next request slot, again governed by the permission probability. On the other hand, successfully acknowledged users will wait for announcement on the allocation schedule from the base-station.

Similar to traditional MAC protocols, the number of request slots in the CHARISMA protocol, N_r , is

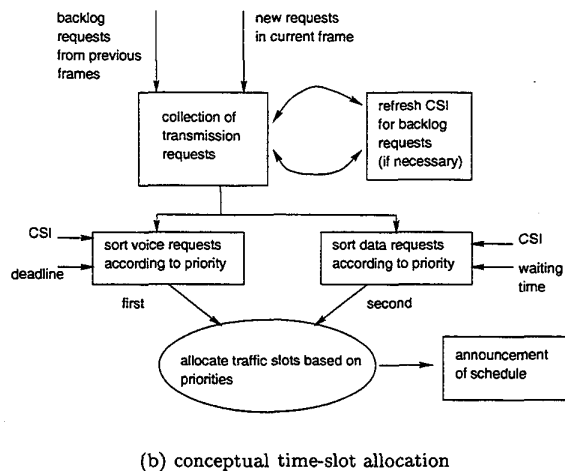
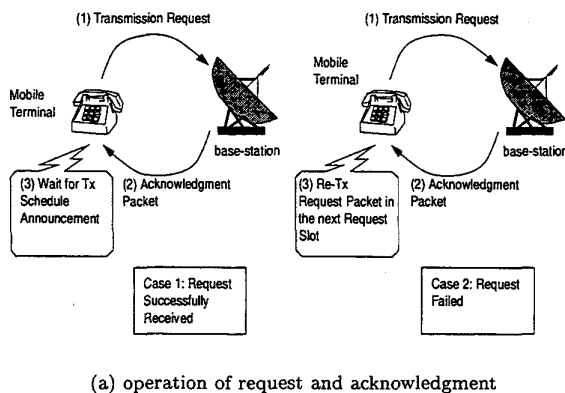


Fig. 3. Operations of the proposed CSI-dependent MAC protocol (CHARISMA).

slightly larger than the number of information slots, N_i . This is useful for gathering more mobile terminal requests as candidates for information time slot allocation. The base-station will collect all requests in the current request phase as well as the *backlog requests* from the previous frames. All the requests will be assigned priorities which are computed according to the deadline, CSI, service type (voice or data), as well as the waiting time of the request (i.e., the number of elapsed frames since the request is acknowledged). The time slot allocation algorithm is conceptually depicted in Figure 3(b). Since the physical layer offers a variable throughput which is dependent on the CSI, the rationale behind the CHARISMA MAC protocol is to give higher priority to the mobile terminals that are in better channel condition in the bandwidth allocation process. The motivation of this

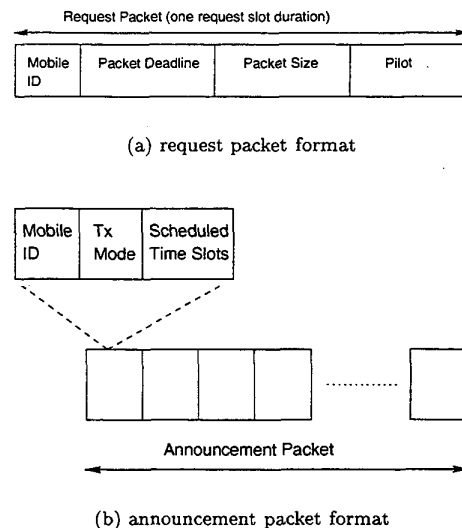


Fig. 4. Formats of the request and announcement packets.

strategy is that a user with better channel condition, with the support of the variable rate channel encoder, can enjoy a larger throughput and therefore, can use the system bandwidth more effectively. Nevertheless, for fairness's sake, information slots should also be allocated to mobile terminals that are approaching their deadlines, despite their possibly worse channel states; otherwise, the queued information packets will be dropped. A higher priority will be assigned to requests with a higher throughput (indicated in the first term) as well as requests with a urgent deadline or long waiting time (indicated in the second term). Thus, in the allocation phase, information time slots are allocated to all service requests according to the priority metrics. If there are not sufficient time slots to service all requests, remaining requests are queued and re-considered in the next frame⁴. After the request phase, the results of information time slot allocation will be broadcast in the announcement subframe of the downlink frame. The announcement packet contains the time slot allocation schedule as well as the transmission mode as illustrated in Figure 4(b). Mobile terminals will then start to transmit information packets on the allocated time slots.

⁴If the deadline for a remaining request has expired, this request will not be queued anymore. The information packet at the mobile terminal will be dropped.

IV. SIMULATION RESULTS

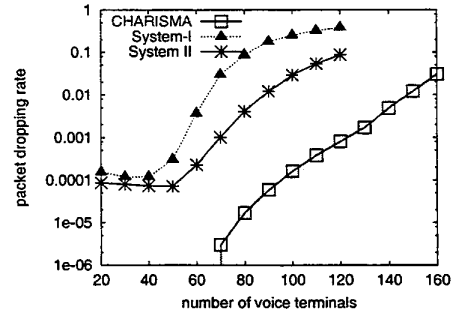
For fair comparison, all the three schemes considered are operating at the same average error level at the same power. They have voice reservation as well as contention request queue. The physical layer offers a fixed throughput of 1 bit per symbol to the MAC layer of system-I while an average throughput of 2 bits per symbol to the MAC layer of system-II and CHARISMA. Table I summarizes the parameters we used.

TABLE I
SIMULATION PARAMETERS.

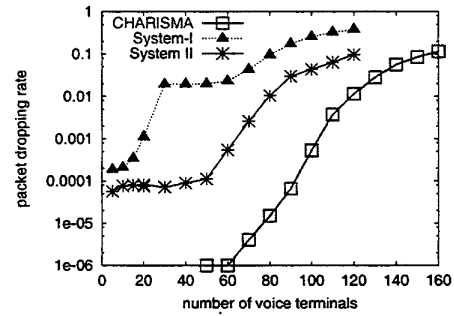
Parameter	Value
N_r	13
N_i	8
N_b	5
t_i	1000 msec
t_s	1350 msec
p_v	0.3
p_d	0.2
data rate (data users)	19.2 Kbps
number of frames	2×10^6

The quality of voice communication is determined by the average packet loss probability and is illustrated in Figure 5 for the CHARISMA, System-I, and System-II. For data users, the performance is specified by the average delay as illustrated in Figure 6.

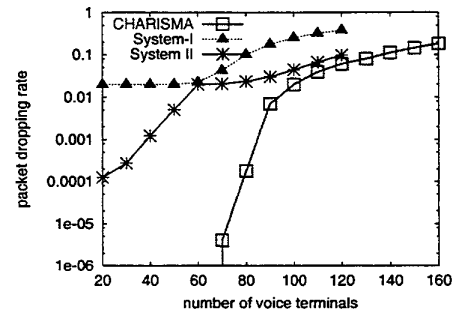
In all test cases System-II is superior than System-I in terms of the packet loss probability and delay because it enjoys twice the offered throughput relative to System-I. This illustrates the merit of using a variable throughput physical layer. However, the potential of synergistic gain is demonstrated by the significant improvement of CHARISMA relative to System-II. For example, at the 1% voice packet dropping rate threshold, the number of voice terminals that can be accommodated by CHARISMA is 2.5 times and 1.87 times greater than those of System-I and System-II, respectively, for $N_d = 0$ (see Figure 5(a)). For data performance, the number of data users that can be accommodated by CHARISMA is 14 times and 2 times greater than those of System-I and System-II, respectively, at 1 sec delay and $N_v = 0$. The performance results for other cases show a similar trend. Hence, the benefit of a joint design of MAC and physical layer is illustrated by the significant performance gains for voice and data users.



(a) Number of data user $N_d = 0$

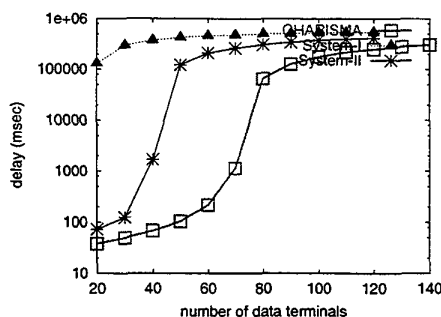


(b) Number of data user $N_d = 10$

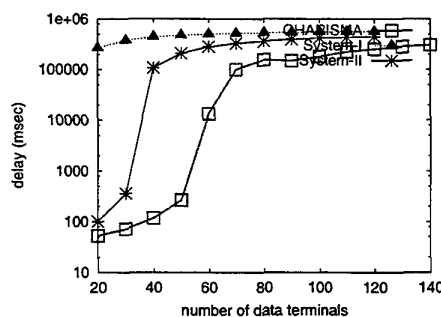


(c) Number of data user $N_d = 20$

Fig. 5. Illustration of synergistic effect—voice terminals performance comparison.



(a) Number of voice user $N_v = 0$



(b) Number of voice user $N_v = 20$

Fig. 6. Data terminals performance—delay against traffic load.

A. Interpretations

Synergy is clearly demonstrated by the experiment results described above. The substantial performance improvements are explicable by the following arguments.

- A voice terminal may experience a deep fading for a long time when it is affected by shadowing. In System-I or System-II, bandwidth allocation in the MAC layer is carried out regardless of the current channel condition as detected in the physical layer. Thus, information slots could also be allocated to such a user and the transmitted packets will be very likely lost due to the poor channel condition. In other words, assigned slots are simply wasted. This kind of wasteful allocation is avoided in our proposed CHARISMA protocol.
- Selection diversity is implicitly incorporated in the CHARISMA protocol. Through the priority based assignment process, every frame is packed with a se-

lected group of information packets with good channel states. Thus, the effective delivered throughput per frame achieved in CHARISMA can be much higher than that in System-II. In CHARISMA, a large number of transmission requests are collected first before allocation of information slots. From the collection of requests, there is a high likelihood that a sufficient number of requests with good channel states can be selected to fully utilize the information slots in an effective manner (i.e., high throughput). For those requests with poor instantaneous channel states, their transmissions are deferred until when the CSI improves or the deadlines are approaching. By contrast, in System-II, requests are served in a first-come-first-serve manner due to the traditional strategy of immediately assigning slots upon successful receipt of requests. Thus, the channel states of such requests are highly diverse and, most importantly, some requests with bad channel states (hence very low throughput) are also served, whereby causing inefficient the bandwidth utilization. For example, a voice terminal may experience a very good CSI for a long time (out of shadowing). In System-I or System-II, this user, however, may fail to successfully transmitted a request to the base-station, probably because of excessive collisions in the request phase. In comparison, our proposed scheme gathers a large number of requests through successive frames, and allocate time slots to the users that can use the system bandwidth more effectively. Thus, the likelihood of “missing” a user with good channel state is much lower and the utilization of bandwidth is therefore higher.

REFERENCES

- [1] N. Amitay and L. J. Greenstein, “Resource Auction Multiple Access (RAMA) in the Cellular Environment,” *IEEE Transactions on Vehicular Technology*, vol. 43, no. 4, pp. 1101–1111, Nov. 1994.
- [2] D. J. Goodman, R. A. Valenzuela, K. T. Gayliard, and B. Ramamurthi, “Packet Reservation Multiple Access for Local Wireless Communications,” *IEEE Transactions on Communications*, vol. 37, no. 8, pp. 885–890, Aug. 1989.
- [3] J. Gruber and L. Strawczynski, “Subjective Effects of Variable Delay and Speech Clipping in Dynamically Managed Voice Systems,” *IEEE Transactions on Communications*, vol. COM-33, no. 8, pp. 801–808, Aug. 1985.
- [4] M. Kawagishi, S. Sampei, and N. Morinaga, “A Novel Reservation TDMA Based Multiple Access Scheme Using Adaptive Modulation for Multimedia Wireless Communication Systems,” *Proceedings of VTC’98*, pp. 112–116, 1998.
- [5] V. K. N. Lau, “Performance of Variable Rate Bit-Interleaved Coding for High Bandwidth Efficiency,” *Proceedings of VTC’2000*, Tokyo, May 2000, to appear.
- [6] V. K. N. Lau and S. V. Maric, “Variable Rate Adaptive Modulation for DS-CDMA,” *IEEE Transactions on Communications*, vol. 47, no. 4, pp. 577–589, April 1999.
- [7] X. Qiu and V. O. K. Li, “Dynamic Reservation Multiple (DRMA): A New Multiple Access Scheme for Personal

- Communication System (PCS)," *ACM/Baltzer Wireless Networks*, vol. 2, pp. 117-128, 1996.
- [8] A. Urie, M. Streeton, and C. Mourot, "An Advanced TDMA Mobile Access System for UMTS," *Proceedings of the IEEE PIMRC'94*, 1994.
- [9] M. D. Yocoub, *Foundations of Mobile Radio Engineering*, McGraw Hill, 2nd edition, 1986.