Bandwidth Allocation for Wireless Multimedia Systems With Most Regular Sequences

Chung Shue Chen, *Student Member, IEEE* and Wing Shing Wong, *Fellow, IEEE*

Abstract—In integrated wireless multimedia service, isochronous traffic of different connections can be scheduled by using a most regular binary sequence (MRBS). Such a sequence schedules traffic in an evenly spaced manner to achieve any arbitrary rate asymptotically and while avoiding excessive delay or buffering requirement. Flexible slot assignment that can match requests exactly improves bandwidth efficiency in multirate operations. The most regular binary sequence provides a distributed solution for multiaccess control that is based on limited information exchange. As a generalization, the concept of a most regular code sequence (MRCS) is proposed to support variable rate transmission in wideband code division multiple access (CDMA) systems and to provide spreading factor (SF) optimization. This scheme improves channel utilization efficiency in supporting traffics of various classes and hence results in an overall capacity gain.

Index Terms—Mobile communication, multiple access, resource management, scheduling.

I. INTRODUCTION

IN THE VISION of future wireless communications, mobile multimedia applications will become a part of our daily life. Different classes of traffics are supported with their respective quality of service (QoS). Variable data rate transmission is essential for such integrated services. This can be done by means of efficient assignment of slots, frequencies, codes, or their combinations. Radio spectrum is always the most valuable resource of a wireless system. Spectrum efficiency is one of the primary concerns in system design. No matter what kind of scheme is used, there is always a high demand on efficient bandwidth allocation. Scheduling techniques for multiuser access are required for these mixed-traffic wireless networks. Resource allocation and channel assignment in the third generation (3G) and future wireless systems have been discussed in many research works such as [1], [2]. Better resource management means more efficient bandwidth utilization. Consequently, more users can be supported in the system and this leads to an overall capacity gain.

Here, we will focus on the bandwidth allocation of variable rate transmission for mixed-traffic wireless systems, especially those based on the wideband CDMA systems [3], [4]. Channel

The authors are with the Department of Information Engineering, the Chinese University of Hong Kong, Shatin, Hong Kong (e-mail: cschen9@ie.cuhk.edu.hk; wswong@ie.cuhk.edu.hk).

splitting optimization can improve resource utilization and allow a better sharing of bandwidth by different users. Resource scheduling has been suggested in the case of multiple users under different classes of traffic requirements to achieve collision-free transmission. As a result, flexible multiservice operations can be done in wireless multimedia systems. In this paper, we concentrate on the bandwidth allocation technique by the most regular sequences. It is interesting to see that these evenly distributed sequences are helpful in channel splitting and resource management.

This paper is organized as follows. In Section II, the system model is described and the problem is formulated. Analysis on multiuser, multirate operations and some related previous works are presented. The channel splitting property of most regular binary sequence (MRBS) is introduced in Section III. Multiaccess control is considered in Section IV in order to achieve collision-free channel sharing. In Section V, we focus on the novel concept of MRCS and its application in spreading factor optimization. Section VI provides a study on performance evaluation. We will present a numerical analysis on system capacity of proposed bandwidth allocations scheme under multiclass traffic based on the generalized Erlang-B model. The last section contains the concluding remarks.

II. SYSTEM MODEL AND PREVIOUS WORKS

The standardization on IMT-2000 framework has been carried out to provide seamless global wideband mobile multimedia services. Harmonization of different radio transmission technologies results in the recommendation of code-division multiple-access (CDMA)-based wideband cellular solutions as the major standards [5], [6]. The wideband CDMA (WCDMA) standardization from 3GPP [7] is named UMTS standing for Universal Mobile Telecommunications System, while 3GPP2 [8] works on cdma2000. As a result of convergence, the three basic operation modes include multicarrier (MC) CDMA, direct spread (DS) CDMA, and time division duplex (TDD) CDMA. MC-CDMA operation is based on the CDMA2000 proposal while the DS-CDMA and TDD-CDMA modes are based on UMTS Terrestrial Radio Access, Frequency Division Duplex (UTRA FDD) and UTRA TDD proposals respectively. In this paper, we have our discussions based on the UTRA models.

A. UTRA Framework

UTRA is a 3G wideband standard. UTRA FDD and TDD schemes employ wideband DS-CDMA and TD-CDMA techniques, respectively. The basic chip rate is 3.84 Mc/s and carrier bandwidth is about 5 MHz. Orthogonal variable spreading factor (OVSF) codes are used in UTRA for data spreading. The
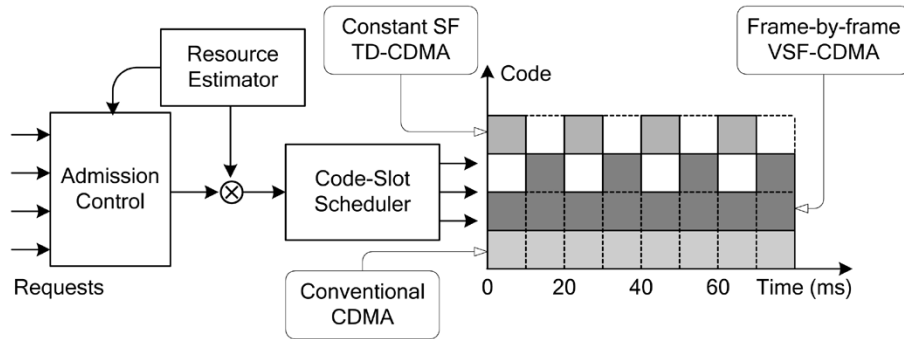
Fig. 1.   RRA for various wideband CDMA operations in the time-code space.

common frame length is 10 ms, in which spreading factor is kept constant. Variable user data rates are achievable with variable spreading factor (VSF) codes on a frame-by-frame basis. In downlink, connections from the same base station (BS) to mobile stations (MS) are separated by different channelization codes managed by the radio network controller. In uplink, channelization codes are used to separate dedicated physical data channel (DPDCH) and control channel (DPCCH) from the same terminal.

B. OVSF Code Tree

An orthogonal tree-structured code set is generated by the OVSF matrix [9]. Each BS assigns channelization codes with its own OVSF code tree. Spreading codes of different lengths are deployed. Code orthogonality is kept between various spreading factors. Code set $\{C_N(n)\}_{n=1}^N$ contains N elements of the same code level. Each has a length of N chips. According to UTRA, N starts from 4. Codes from C_4 to C_{256} are employed with the corresponding SF of 4 to 256. To keep orthogonality, parent-code and child-codes cannot be used at the same time. For example, when $C_{16}(1)$ is used, its parent-code $C_8(1)$ and child-codes $C_{32}(1)$ or $C_{32}(2)$ cannot be used. Similarly, parent-code of $C_8(1)$ and child-codes of $C_{32}(1)$ or $C_{32}(2)$ are not allowed. Different users follow this rule of orthogonal spreading. Code orthogonality is maintained in the synchronous multiuser mode. Since the number of codes is limited, code utilization efficiency is a major concern.

C. Resource Management for Multirate Operations

Optimal radio resource allocation (RRA) [2], [10] is always an important topic in cellular systems. In wideband CDMA proposals, code and time are scheduled to assign channels for different users. RRA is done to maximize the number of users supported within the limited bandwidth. Code-time plane is divided into slots. As shown in Fig. 1, different requests are first managed by admission control with the information of resource estimation. Different combinations of time and code slots are scheduled to support mixed traffics. Different schemes can be applied to offer bandwidth on demand with variable data rates, such as VSF-CDMA or hybrid TD-CDMA.

D. Related Works and Problems

Multirate operation is one of the most important features in 3G wireless systems [11]. It allows flexible data rate for different
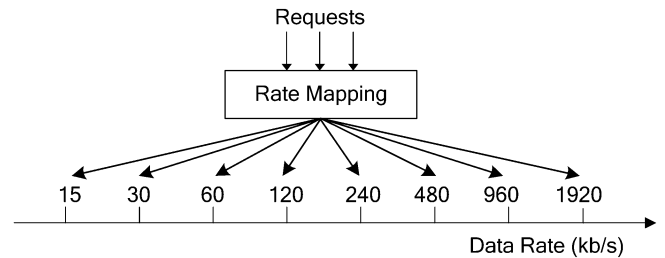

Fig. 2.   Rate mapping of multirate request to defined data rate set.

classes of multimedia service. In WCDMA, multiple rate transmissions are supported by means of variable spreading factor while in CDMA2000 this is achieved via multicode techniques. Related works and possible solutions are described in the following.

1) Conventional CDMA: In conventional DS-CDMA systems, the spreading factor is fixed. For example, using the parameters in UTRA FDD [12] as an illustration, supported transmission rates are limited to be $2^i \times 15$ kb/s, where $i = 0, 1, \ldots, 7$. In Fig. 2, the requested rate is mapped to the limited set of supported data rates. Each connection will choose a suitable channel with a corresponding spreading factor. However, due to quantization effects, channel utilization is inefficient when the request is between two successive supported rates. In such a case, a larger rate must be offered. Bandwidth wastage is particularly serious when high data rate operations are involved. Such an utilization inefficiency can be eliminated if any arbitrary request can be perfectly matched.

2) MC Transmission: In MC-CDMA, multiple spreading codes are used in parallel to support requested data rate. In UTRA FDD, the maximum channel splitting factor is defined to be 6, that is, we may use at most six codes simultaneously to match the requested data rate R_q. As a result, the provided data rate R_p is a combination of the supported basic rates given in Fig. 2 such that

$$R_p = \sum_{i=0}^{7} N_i(15 \times 2^i)$$

where N_i is either 0 or 1. The sum of N_i cannot be larger than 6. When the request cannot be matched exactly, a larger channel bit rate is provided to fulfill the requirement. We have rate matching as $R_p \geq R_q$. This can be solved by rounding up R_q to the closest binary combination of basic rate 15 kb/s under the N_i
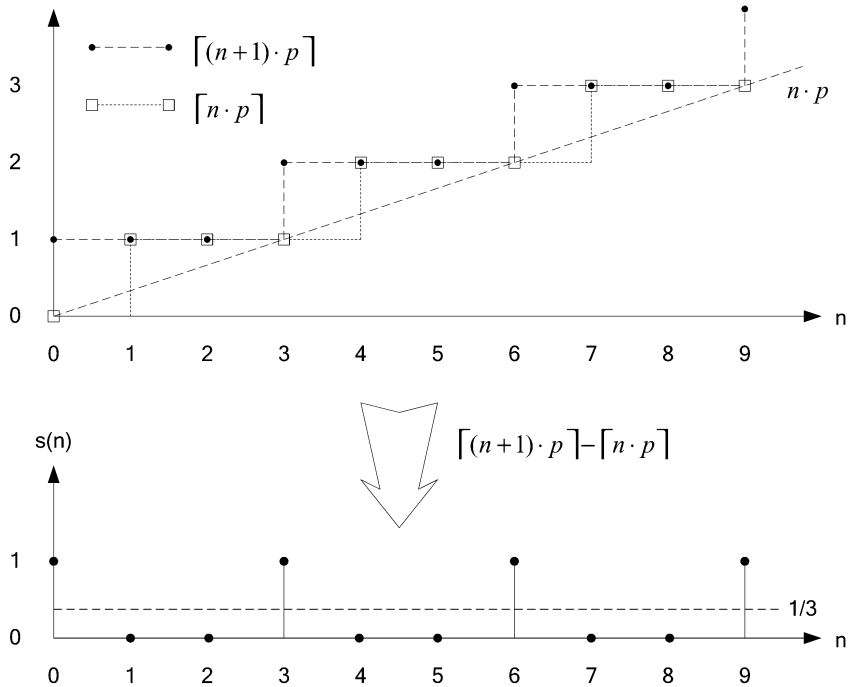
Fig. 3. Zero-one valued MRBS of asymptotic mean 1/3 is constructed by projecting values in each step n upward to the nearest integer. The difference gives the resultant sequence $\{1, 0, 0, \ldots\}$.

constraint. With finite number of simultaneous codes, there is always quantization approximation and bandwidth wastage.

Assume R_q is uniformly distributed between 15 kb/s and 1.92 Mb/s, the expected bandwidth allocation efficiency (BAE $= R_q/R_p$) is 0.75 when single code is used as given in Appendix A. Simulations [13] on multirate operations with respect to channel splitting factor shows that the expected BAE can be improved to 0.99 when the factor is larger than 4. However, this requires the receiver be able to detect such a number of codes in parallel and thus increase the receiver side complexity, which is especially hard at a thin mobile terminal. On the other hand, it causes more serious fragmentation in the OVSF tree code structure for channel assignment. The larger the number of codes used simultaneously, the higher is the system complexity.

3) Hybrid TDMA/CDMA: Hybrid time division multiple access (TDMA)/CDMA technique has been discussed in [14] to support multirate operation. Time slots of different durations are provided for different transmission rates. Low bit-rate services such as speech or voice call are operated in short time slots while high bit rate services are operated in long time slots. This provides flexible bit rate. Different slot lengths are defined in advanced for different classes of services. However, it is practically not easy to have many time slot lengths for fine enough granularity.

III. REGULAR CHANNEL SPLITTING

To support multirate operation, the proposed schedule based on the most regular sequence splits slot resource evenly among different users according to their demands. For any request with a rate expressible as a rational faction, not greater than 1, of the total available bandwidth, such a schedule allocates transmission slot to it with the properties: 1) the asymptotic rate is equal to the request rate p; 2) the scheduling sequence is periodic and deterministic; and 3) the slots are most regular distributed in the sense this evenly spaced extremal zero-one valued sequence has the most regular interarrival time among all arrival sequences of the request rate. The details are given in the following.

A. MRBS

The MRBS generating function is defined [15] as

$$s(n) = \lceil (n+1)p + \theta \rceil - \lceil np + \theta \rceil \qquad (1)$$

where n is a nonnegative integer, $0 < p \leq 1$, and θ is the initial phase. Without loss of generality, we assume the initial phase to be zero in our discussion. As given in Appendix B, the sequence has an asymptotic mean of p. For any positive rational number $p = m/l$ where m and l are relatively prime integers and $0 < m \leq l$, it is periodic with duration l bits. The generated MRBS is cyclic and deterministic. As an illustration, a periodic MRBS of asymptotic mean 1/3 is generated as in Fig. 3. Along the asymptotic line, the value of np at each step is projected up to the nearest integer. $\lceil np \rceil$ outputs $\{0, 1, 1, 1, 2, 2, 2, \ldots\}$ while $\lceil (n+1)p \rceil$ outputs $\{1, 1, 1, 2, 2, 2, 3, \ldots\}$, which is a shifted version of the $\lceil np \rceil$ sequence. Both of them increase by one every three slots. The difference of these two sequences, which is either "0" or "1," gives a resultant pattern $\{1, 0, 0\}$.

As proved in [15], [16], among all arrival processes for an exponential server queue with specified average arrival and service rates, the arrival process which minimizes the average delay and all moments of the delay, as well as other related quantities, is the process with constant interarrival times. The proof relies on the convexity property that the expected number of customers in an
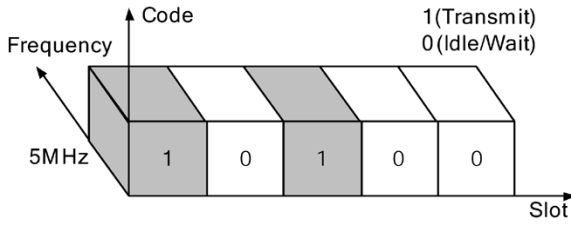
Fig. 4. MRBS time slot assignment with $p = 2/5$. The outputted periodic sequence is $\{1, 0, 1, 0, 0, \ldots\}$. The period of this sequence is 5.

exponential server queue as well as all higher moments at a given time is a convex function of the set of previous interarrival times. The MRBS is such an extremal zero-one valued sequence with specified mean that the resultant pattern has the most regular interarrival time among all arrival sequences of the request rate. This arrival process leads to the smallest long-term average queue size and corresponding delay. We are motivated by its most regular channel splitting property on resource scheduling.

B. MRBS in TD-CDMA Time-Slot Assignment

As discussed before, hybrid TDMA/CDMA technique can support multirate operation by defining time slots of different lengths corresponding to the service data rates. In particular, a finite set of slot lengths are used to match a large variety of traffic classes. However, the granularity is still generally not fine enough. This degrades the resource utilization. Comparatively, the MRBS scheme uses a fixed length slots and assigns them according to the expanded zero-one sequence. A channel has permission to transmit if and only if it has a "1" in the given slot. Exact rate matching can be achieved. An example is given in Fig. 4.

Suppose the requested rate is R_i and the channel bandwidth is R. We have $p_i = R_i/R$. It is assumed that R_i is a rational number less or equal to R and consequently $p_i = m_i/l_i$ for some integers $0 < m_i \le l_i$. So, p_i represents the required proportion of request rate by user i in the provided capacity, which is denoted by 1. Given p_i, the sequence generator will switch between "1" and "0" with respect to the slot number n. Please note that the asymptotic rate is always equal to the request rate. We assume admission control is carried out to make sure the sum of all requests is less or equal to the available capacity. That is, $\sum_i p_i \le 1$. An alternative is if multiple channels are available and by mapping users to different channels the condition $\sum_i p_i \le 1$ will still hold for each channel. The system is in a stable state.

IV. MULTI-ACCESS CONTROL

If MRBS is applied to individual user independently without any synchronization, it is possible that contention to use the same slot by more than one user can occur. This leads to the question of how to prevent such contention from occurring.

A. Conflict and Resolution

To fix ideas, assume there are M users in the system and each has requested a rate R_i with corresponding p_i. Each user has a traffic regulator. By MRBS, the permission to use the channel at slot j is given by the sequence

$$s_i(j) = \lceil (j+1)p_i \rceil - \lceil j \times p_i \rceil. \tag{2}$$

Each user is assumed to have a circular buffer of size "1" analogously equal to one packet defined as the amount of data to be transmitted in one slot unit. The buffer can transmit and receive data at the same time.

To illustrate our contention resolution algorithm, we have the following example in a TD-CDMA system. Assume there are three users, A, B, and C, with requested rates p_i of 1/3, 1/4, and 1/6 respectively. MRBS slot patterns $\{1, 0, 0, \ldots\}$, $\{1, 0, 0, 0, \ldots\}$, and $\{1, 0, 0, 0, 0, 0, \ldots\}$ are generated, respectively, according to their p_i values. The buffering and scheduling model is shown in Fig. 5, in which gray slot indicates access request. In slot 1, all of the three users are "1" and require an access. However, the BS can provide only "1" at this moment. There is a conflict here which can be resolved by proper buffering and transmission scheduling. We may buffer B and C first and assign A for the access of first slot. In slot 2, B and C have a higher priority since they have buffered data now. Iteratively, slots are assigned to users in a conflict-free manner. In our proposed scheme, the allocated transmission slots also follows MRBS pattern. In this example, the aggregate sum of the three requests is equal to 3/4. Consequently, the slot allocation pattern is $\{1, 1, 1, 0, \ldots\}$. In the scheduling system, a buffer of size "1" for each user suffices. To show that this idea can be applied in general, we need some technical results. The details are explained in the following.

Proposition 1: Let $S(j) = \sum_{i=1}^{M} \sum_{n=0}^{j} s_i(n)$ and $S_{\tilde{M}}(j) = \sum_{n=0}^{j} s_{\tilde{M}}(n)$, where the MRBS $\{s_i(n)\}$ and $\{s_{\tilde{M}}(n)\}$ have an asymptotic mean p_i and $p_{\tilde{M}} = \sum_{i=1}^{M} p_i$, respectively. Then, $0 \le S(j) - S_{\tilde{M}}(j) \le M - 1, \forall j$.

Note that $S(j)$ is the total number of transmission slots required for all users up to time j while $S_{\tilde{M}}(j)$ is the total number of provided slots for the aggregate rate $p_{\tilde{M}}$. *Proposition 1* gives the upper and lower bounds on the difference of transmission slot demand $S(j)$ and supply $S_{\tilde{M}}(j)$. The difference is always smaller than M. The proof is presented in Appendix C. It is worth of pointing out that the number of provided slots is the minimum necessary amount for the requests. The proposed contention resolution algorithm allocates slot according to the MRBS $s_{\tilde{M}}(n)$ with a mean of the aggregate rate $p_{\tilde{M}}$ and holds the properties: 1) total number of provided transmission slots for the requests is kept at a minimum necessary level and 2) the provided slots for requests still follows the MRBS pattern. The incentive of first property is to maintain the optimal resource utilization by using a minimum necessary amount of transmission slots to satisfy all the requests. The second property is useful for multiple level slot scheduling or resource allocation as the resultant allocation pattern is always a MRBS. This could be extended to multiple-layered model or hierarchical network. In this paper, we focus on the aforementioned single-layered model, which can serve as the basis for future works.

The result of *Proposition 1* gives us a hint on how to solve the multiaccess problem. Since the difference of $S(j)$ and $S_{\tilde{M}}(j)$ is always less than M, intuitively a buffering requirement of size "1" at each user can be sufficient to accommodate all the users. On the other hand, this will keep the buffering at minimum and avoid excessive delay. To obtain a conflict-free algorithm, the idea is to ensure that the MRBS sequences are modified while
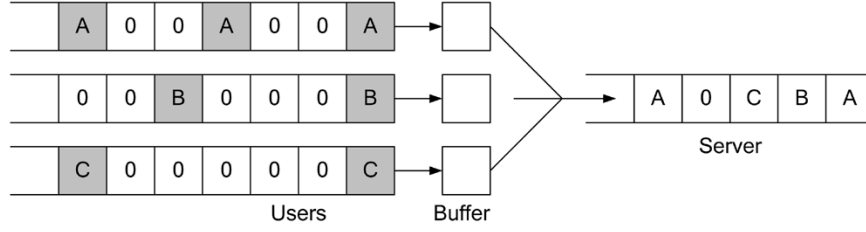
Fig. 5. Buffering and scheduling model for MRBS-SSA. A buffer of size "1" for each user suffices.

each user needs to buffer at most one packet of data. To achieve this, a generalized MRBS slot scheduling algorithm (SSA) for multiaccess control is described in the following.

B. MRBS Slot Scheduling

The MRBS-SSA assigns transmission priority to users according to the following rules. First, a user with buffered data has the highest priority. This rule aims to avoid excessive delay and buffer overflow. Second, for users with the same priority according to the first rule, the user with an earlier arrival has a higher priority to transmit. Third, if previous rules fail to resolve the priority, a user requesting a higher rate has a higher priority. Users with the same rate are assigned different predefined priorities. Each user has a circular buffer of size "1" as defined.

To initialize the algorithm, we construct $s_i(j)$ for each p_i and $s_{\tilde{M}}(j)$ for $p_{\tilde{M}}$ from $j = 0$ to $N - 1$, where $N = \mathrm{LCM}_i(l_i)$ in which LCM is the lowest common multiple operator. Let $b_i(j)$ be the amount of buffered data of user i at time j. A priority indicator ID_i from 1 (highest priority) to M (lowest priority) is assigned to each i for their p_i values. If $p_i > p_j$, then $ID_i < ID_j$. If $p_i = p_j$ and $i < j$, then $ID_i < ID_j$. So, ID_i is unique for each i. Let $d_i(j)$ be the smallest positive integer such that $s_i(j + d) = 1$. The following MRBS-SSA can be run independently among different users with known p_i values.

Procedure (MRBS-SSA)
1) Start at $j = 0$, $b_i(0) = s_i(0)$ and go Step 3.
2) (Data buffering) $b_i(j) = b_i(j - 1) + s_i(j)$.
3) If $s_{\tilde{M}}(j) = 0$, then go Step 5, else assign current service slot to the highest priority user by the following.
  3.1 (Buffered data has the highest priority) $B = \{i : b_i(j) = \max_k b_k(j)\}$. If $|B| = 1$, $H = B$ and then go Step 4.
  3.2 (B has multiple elements) $D = \{i \in B : d_i(j) = \min_{k \in B} d_k(j)\}$. If $|D| = 1$, $H = D$ and then go Step 4.
  3.3 (D has multiple elements) $H = \{i \in D : ID_i = \min_{k \in D} ID_k\}$.
4) $b_{i \in H}(j) = b_{i \in H}(j) - 1$.
5) End if $j = N - 1$, else $j = j + 1$ and go Step 2.

C. Properties of MRBS-SSA

In order to illustrate and prove the properties of MRBS-SSA, starting from here, we use $b_i(j)$ and $\tilde{b}_i(j)$ to denote the amount of buffered data of user i at time j before and after slot allocation, respectively. Some technical results from the MRBS-SSA are given in the following propositions.

Proposition 2: At any time j, by MRBS-SSA, no more than one user receives a service slot.

Proof: Since ID_i is unique for each i, in Step 3.3, $|H| = 1$. In the case $s_{\tilde{M}}(j) = 1$, a service slot is assigned to the highest priority user chosen by Step 3.1 to 3.3 in a deterministic way. $|H|$ is always equal to 1. So, no more than one user receives a service slot. ∎

Proposition 3: At any time j, by MRBS-SSA, $0 \leq \tilde{b}_i(j) \leq 1$ for all user i if each of the arrivals follows their corresponding MRBS of p_i, where $i = 1, 2, \ldots, M$.

As a result, $\tilde{b}_i(j) \leq 1$ guarantees there is no buffer overflow while $\tilde{b}_i(j) \geq 0$ shows that, if $s_{\tilde{M}}(j) = 1$, there must be a user with nonempty buffer which is assigned to transmit. So, every service slot is utilized. By *Proposition 2*, this highest priority user is uniquely defined such that exactly one user will transmit. The proof of *Proposition 3* is given in Appendix E.

Proposition 4: For rational number p_i, where $i = 1, 2, \ldots, M$, the order of transmission for all users by MRBS-SSA is deterministic and periodic.

The proof is presented in Appendix F. In the beginning of every N-slot cycle, $s_i(0) = 1$, $\forall i$. Since, $s_{\tilde{M}}(0) = 1$, $S(0) - S_{\tilde{M}}(0) = M - 1$. We need to buffer exactly $M - 1$ packets at $j = 0$ after allocating the first service slot. This is the most congested case. In the end of every N-slot cycle, $\tilde{b}_i(N-1) = 0$. All buffers are empty and this gives the lower bound on the inequality of *Proposition 1*. MRBS slot allocation provides a distributed deterministic channel splitting method in the sense that both the server and clients can individually generate the equivalent scheduled sequence as long as requested rates in the system are given. This can be run in advance or in a real time manner. The order of transmission is deterministic and periodic.

V. EXTENSION TO MRCS

The idea of MRBS has been further developed to the most regular code sequence (MRCS). We have proposed a generalized fraction-valued sequence in [17] to represent an arbitrary value in asymptotic mean. An MRCS allocates a fraction equals to $1, 1/2, \ldots$, or up to $1/2^i$ of the resource in each time slot. This type of sequence can be used in a CDMA system to satisfy a requested data rate with no fragmentation lost. For example, in a UTRA single code CDMA system, each mobile user will be provided one transmission channel with a corresponding channelization code. Variable spreading factor assignment is allowed. Defining the frame duration as one transmission slot, transmission scheduling by means of MRCS can be realized by using the OVSF code tree. In the paper, we do not deal with the issue of dynamic spreading code reassignment as discussed in [18], [19]. We assume that as long as the total rate in a given time slot is less or equal to 1, it is realizable by some dynamic code assignment schemes.

A. MRCS

The sequence generating function is defined as

$$c(n) = (p_U - p_L) \times \left(\left\lceil \frac{(n+1)p}{p_U - p_L} \right\rceil - \left\lceil \frac{np}{p_U - p_L} \right\rceil \right) \quad (3)$$

where n is a nonnegative integer, $0 \leq p_L < p \leq p_U \leq 1$, p_L and p_U are in the form of 2^{-i} or 0, $i \in \mathbb{Z}^+$. To minimize the variance of MRCS, p_L is chosen to be the closest value smaller than p. It is named as the lower-closest code (LCC) rate while p_U, the closest value greater or equal to p, is defined as the upper-closest code (UCC) rate. So, $p_L < p \leq p_U$. If $p_L > 0$, $p_U = 2 \times p_L$. When $p = 2^{-i}$, $c(n) = p_U$. Switching between the UCC and LCC values, we can achieve any requested granularity asymptotically by such an evenly spaced fractional sequence.

As shown in Appendix G, MRCS is a periodic function for any positive rational number $p = m/l$, where m and l are relatively prime integers and $0 < m \leq l$. It is cyclic and deterministic. The period is equal to l. For example, when p is $1/3$, p_L and p_U are $1/4$ and $1/2$ respectively. So, $c(n) = \{1/2, 1/4, 1/4, 1/2, 1/4, 1/4, \ldots, \}$. The periodic pattern $\{1/2, 1/4, 1/4\}$ has a mean of $1/3$. Comparing MRBS and MRCS at $p = 1/3$, it is interesting to see that the MRBS $\{1, 0, 0\}$ has been transformed into $\{1/2, 1/4, 1/4\}$ by redistributing the "1" among "0"s. The fraction-valued MRCS is more flexible to be used for various spreading factors and has a smaller variance than the zero-one valued MRBS.

B. WCDMA Spreading Factor Optimization

MRCS can be employed in code division resource allocation scheme to provide efficient channel splitting. Assume each user has a request rate R_i, which is expressible as a fraction p_i of the total available bandwidth R. So, $R/2^{i+1} < R_i \leq R/2^i$ while $2^{-i-1} < p_i \leq 2^{-i}$. In a capacity of 3.84 Mb/s, channels of data rate from 15 to 960 kb/s are supported. Table I gives the corresponding spreading factors from 4 to 256 with defined code rates 2^{-i}, respectively, based on a chip rate of 3.84 Mc/s, where $i = 2$ to 8. For example, if R_i is equal to 375 kb/s, the closest lower and upper data rates are 240 and 480 kb/s respectively. So, $p_i = 375/3840 = 25/256$ and $p_L = 1/16$. MRCS will switch between 1/16 and 1/8 regularly to achieve an average of 25/256. The sequence period is 256. The requested rate is provided in the time average of the period with a minimum rate fluctuation.

C. MRCS Slot Scheduling in Multiple Access

Assume there are M users in the system and each asks for p_i. Code rate $c_i(j)$ is provided to user i at slot j following

$$c_i(j) = (p_{U(i)} - p_{L(i)})$$
$$\times \left(\left\lceil \frac{(j+1)p_i}{p_{U(i)} - p_{L(i)}} \right\rceil - \left\lceil \frac{j \times p_i}{p_{U(i)} - p_{L(i)}} \right\rceil \right). \quad (4)$$

For simplicity, we assume p_i is a rational number in the form of m_i/l_i and the aggregate rate $p_{\tilde{M}} = \sum_i p_i \leq 1$. Similar to MRBS, multiaccess control is essential to resolve spreading code contention in multiple users. For example, there are three users and each has p_i of 1/3. So, all of them will have the sequence $\{1/2, 1/4, 1/4\}$. At slot $j = 0$, the sum of $c_i(j)$ equals

TABLE I
SPREADING FACTOR AND CODE RATE WITH RESPECT TO REQUESTED DATA RATE

Data Rate (kb/s)	Code Rate	Spreading Factor
15	1/256	256
30	1/128	128
60	1/64	64
120	1/32	32
240	1/16	16
480	1/8	8
960	1/4	4

to 3/2 while the available aggregate code rate is 1, which cannot satisfy all the users. Data buffering and scheduling should be performed to accommodate each request from all the users.

We define an algorithm to resolve the multiaccess problem based on MRCS here, called the MRCS-SSA. Similar to MRBS-SSA, a user with more buffered data has a higher priority. In the case several users have the same priority according to this rule, the one with an earlier arrival has a higher priority to transmit. If previous rules fail to resolve the priority, a user requesting a higher rate has a higher priority. Users with the same rate are assigned different predefined priorities. Following this idea, we can find a proper scheduling for the MRCS based on the approach of MRBS-SSA to have an explicit transmission order for all users. A generalized MRCS slot scheduling algorithm is described in the following.

To initialize the algorithm, we generate $c_i(j)$ and $c_{\tilde{M}}(j)$ based on p_i and $p_{\tilde{M}}$, respectively. Let $b_i(j)$ be the amount of buffered data of user i at time j. A priority indicator ID_i from 1 (highest priority) to M (lowest priority) is assigned to each i for their p_i values. If $p_i > p_j$, then $ID_i < ID_j$. If $p_i = p_j$ and $i < j$, then $ID_i < ID_j$. So, ID_i is unique for each i. Under the single code operation constraint, each user can transmit at a rate in the form of 2^{-k}, $k \in \mathbb{Z}^+$. Service rate assignment is based on the aforementioned service disciplines with priority order PO_i defined in the following algorithm. ($PO_i = 1$ is the highest.) In Step 3.1 to 3.3, the allocation priority order of users is explicitly determined and we assign the aggregate rate to them from the highest priority one to the lowest. Scheduled MRCS is formulated in a deterministic way.

Procedure (MRCS-SSA)
1) Start at $j = 0$, $b_i(0) = c_i(j)$ and go Step 3.
2) (Data buffering) $b_i(j) = b_i(j - 1) + c_i(j)$.
3) Assign priority order for users.
   3.1 (Buffered data has the highest priority) If $b_i(j) \geq b_k(j)$, $PO_i(j) \leq PO_k(j)$.
   3.2 For those have same priority by 3.1, $PO_i(j) \leq PO_k(j)$ if $c_i(j + 1) \geq c_k(j + 1)$.
   3.3 For those have same priority by 3.2, $PO_i(j) < PO_k(j)$ if $ID_i < ID_j$.
4. Allocate service rate $\gamma_i(j)$ to i in the ascending order of PO_i, $c_{\tilde{M}}(j) = c_{\tilde{M}}(j) - \gamma_i(j)$, $\gamma_i(j) = \max_k 2^{-k}$, $k \in \mathbb{Z}^+$, under constraints (i) $\gamma_i(j) \leq c_{\tilde{M}}(j)$ and (ii) $\gamma_i(j) \leq b_i(j)$.
5) $b_i(j) = b_i(j) - \gamma_i(j)$
6) $j = j + 1$ and go Step 2.

TABLE II
SCHEDULING DELAY OF MRBS ARRIVALS VS. UNIFORM DISTRIBUTION (IN THE UNIT OF SLOT LENGTH)

Number of users	Average scheduling delay		Scheduling delay variance	
	MRBS arrivals	Uniform distribution	MRBS arrivals	Uniform distribution
2	0.9490	3.7410	3.1969	17.8341
3	1.4192	4.5569	7.9188	32.3883
4	1.8795	5.0402	14.3584	43.8101
5	2.3416	5.2152	20.0825	50.9110
6	2.7926	5.3225	30.8625	57.1646
7	3.2602	5.5592	38.6005	66.4987
8	3.7087	5.6455	48.5594	74.0398
9	4.1565	5.8023	59.8018	83.3085
10	4.5840	5.8766	71.2828	87.6026

VI. PERFORMANCE EVALUATION

A. Simulation on Slot Scheduling Algorithm

We have implemented the slot scheduling algorithms MRBS-SSA and MRCS-SSA. In the simulations, the number of users is random and between 2 and 10. Each user has a uniformly distributed $p_i \in (0, 1)$. The generated p_i is normalized to make sure that the sum of all p_i is less or equal to 1 and then they are used to construct the most regular sequences for slot scheduling. In the simulations, there is no unsolvable contention with the provided buffering and scheduling disciplines. In a long run, the original and scheduled sequences are compared for their consistence on the asymptotic mean. The provided service rates are fully shared by the users. The amount of data buffered in each user is finite and bounded. A deterministic transmission order of all users is formulated for such an explicit bandwidth allocation.

B. Resource Utilization and Delay Constraint

Delay constraint is highly emphasized in wireless multimedia traffics. In most regular sequences, transmission request is evenly distributed to avoid excessive buffering requirement and scheduling delay. As mentioned previously, among all arrival sequences of a specific rate, the one which yields the smallest mean delay as well as its higher moments in a queue is the one with most regular interarrival times [15]. This encourages us to employ these regular sequences in bandwidth allocation and at the same time to have effective resource utilization. With appropriate slot scheduling, flexible multirate transmission can be achieved in a conflict-free manner for multiple users.

Table II gives the delay performance of MRBS arrivals in simulations. In our analysis, the number of users is from two to ten. Under a specific number of users, each of them has a uniformly distributed $p_i \in (0, 1)$. The generated p_i is normalized to make sure the sum of p_i is less or equal to 1. Each simulation set has been run for 1000 times and the average is taken. Generally, the modified transmission slot sequence has a scheduling delay in comparison with the original MRBS due to the service disciplines of multi-access control. The delay is in fact equivalent to the time packets buffered in the queueing model as shown in Fig. 5. The average delay and corresponding variance

are obtained. As a reference, the scheduling delay of arrival processes with the same set of request rates in which "0" and "1" are generated in a uniform distribution are measured. A comparison is provided to investigate the difference due to different arrival (access) patterns. The results in Table II show that the MRBS arrivals have experienced smaller average scheduling delay and delay variance than those from uniform distribution. It is more important to look at the variance than the average value especially in the case of high-speed networks. There can be a high tolerance on the average delay by keeping a sufficient buffering before the start of service. It does not matter so much on the latency but is quite sensitive to the variance in transit time. Due to the property of most regular interarrival times, MRBS results in the minimized average delay and its higher moments in access contention under multiple users.

In addition, the bound on MRBS scheduling delay is deterministic. By *Proposition 3*, a buffer of size "1" for each user is sufficient in contention resolution. For any arrivals, buffered packet will be transmitted before the accumulation of the coming one. Consequently, we can derive the upper bound on the delay experienced by each packet in the corresponding buffer. The worst case delay is equivalent to the durations between two successive arrivals in the sequence. That is, $\lceil 1/p_i \rceil$ slot lengths. The proof is given in Appendix H.

C. System Capacity and Blocking Model

We have employed the generalized Erlang B model [20], [21] to investigate the blocking probabilities of different traffic classes and the system capacity. Suppose there are k Poisson arrival streams each with mean rate λ_i. User from stream i is defined as class i customer. Each class i customer has a request of average rate b_i kb/s. The service time for class i customer has an exponential distribution of mean $1/\mu_i$. The traffic intensity of class i is equal to λ_i/μ_i. If a class i customer arrives and cannot find a channel b_i kb/s available, it is blocked. For a finite bandwidth of C kb/s, the blocking probability of class i user is given by

$$\mathrm{P}_{Block}^{b_i} = 1 - \mathrm{P}\left(\mathbf{n} : \sum_{j=1}^{k} n_j b_j \leq C - b_i\right)$$
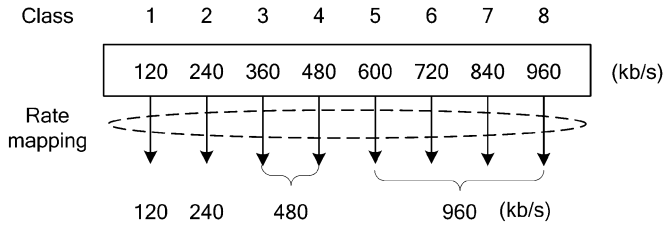
Fig. 6. Mapping of request $b_i = \{120, 240, \ldots, 960 \text{ kb/s}\}$ to the limited set of supported channel $\{120, 240, 480, 960 \text{ kb/s}\}$ due to quantization effect in conventional CDMA scheme.

where $\mathbf{n}$ is in the set of allowable states

$$\Omega = \left\{ \mathbf{n} : \sum_{i=1}^{k} n_i b_i \leq C \right\}, \quad n_i > 0.$$

In the numerical study, we assume $C = 3.84$ Mb/s according to UTRA and the total arrival of intensity λ is distributed by equal probability into k streams such that $\lambda_i = \lambda/k$. The service time μ_i is the same for all i. System capacity and blocking probability in multiclass arrivals is evaluated among the proposed MRBS bandwidth allocation scheme with fully flexible requested rates and conventional single-code CDMA or hybrid TDMA/CDMA operation in which the channel granularity is not fine enough. Due to quantization effect, we assume the latter can support a requested channel in the quantized level 2^i only. A higher rate channel is provided in the case if the request is in between two supported values as shown in Fig. 6. We have the numerical analysis on the set of requests from 120 to 960 kb/s.

The blocking probabilities of different classes in conventional CDMA and proposed MRBS scheme are calculated recursively by the generalized Erlang B formula based on the algorithm suggested in [20] and plotted in Figs. 7 and 8, respectively. The results are obtained for two cases in which arrivals are from Class 1 to 4 (120–480 kb/s) and Class 1 to 8 (120–960 kb/s), respectively. Given the same capacity, it is intuitively reasonable the blocking probability of the same class in the case of (Class 1–4) is larger than that in the case of (Class 1–8) since we need to allocate more resource for some higher rate requests. This fact can be found from both Figs. 7 and 8. In the case (Class 1–4) of Fig. 7, the blocking probabilities of Class 3 (360 kb/s) and 4 (480 kb/s) are the same because a channel of 480 kb/s will be given for a request of 360 kb/s in conventional scheme to satisfy the service requirement. Similarly, in the case (Class 1–8), Class 3 and 4 have the same blocking probability curve while Class 5 to 8 have the same curve. In comparison, our proposed scheme outperforms the conventional one as it gives a lower blocking probability generally on various classes under the same total arrival traffic intensity. From another point of view, the proposed scheme allows a higher traffic intensity than the conventional one under the same blocking probability.

The overall performance of the conventional CDMA and proposed scheme is compared in Fig. 9. Average blocking probability is defined as the mean of blocking probabilities of different classes with weightings corresponding to their arrival intensities. On the same blocking level, the proposed bandwidth allocation scheme can support a higher total traffic intensity than the conventional one. There is more than 10% capacity
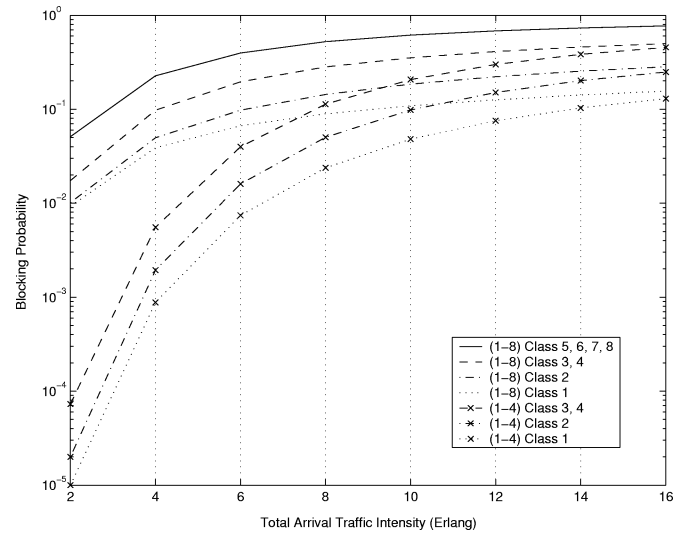

Fig. 7. Blocking probabilities of arrivals from different stream classes in conventional CDMA.
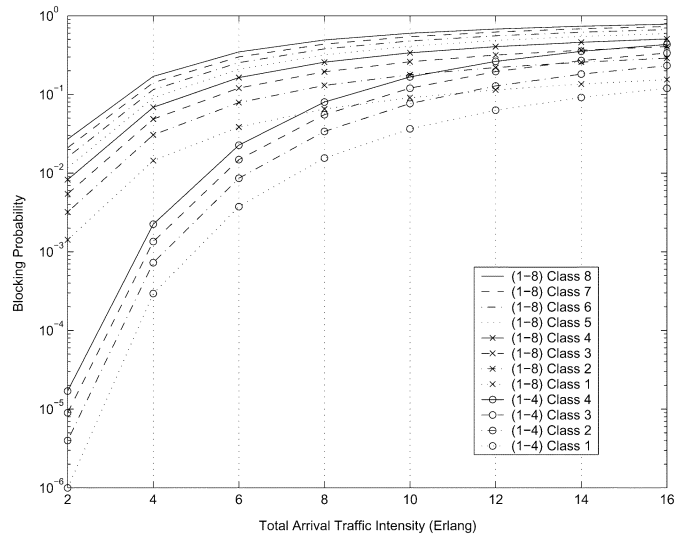

Fig. 8. Blocking probabilities of arrivals from different stream classes in the proposed MRBS bandwidth allocation scheme with fully flexible requested rates.

gain under each blocking threshold. The flexible rate scheme achieves a higher resource utilization efficiency and hence can support more traffic in integrated multiclass services. This result gives an illustration on the effectiveness of prefect rate match in multi-rate operations.

VII. CONCLUSION

In this paper, we focus on the channel assignment efficiency for integrated services in wideband CDMA systems. Channel splitting optimization in multirate operation is proposed with the most regular sequences. Time varying assignment of slot resource can provide fully flexible transmission rate. MRBS has been suggested in TD-CDMA bandwidth splitting to provide collision free transmission even under strict buffering requirement and delay constraint. Every user can individually generate the corresponding scheduled sequence with known p_i. This gives a relatively distributed solution for the whole system.
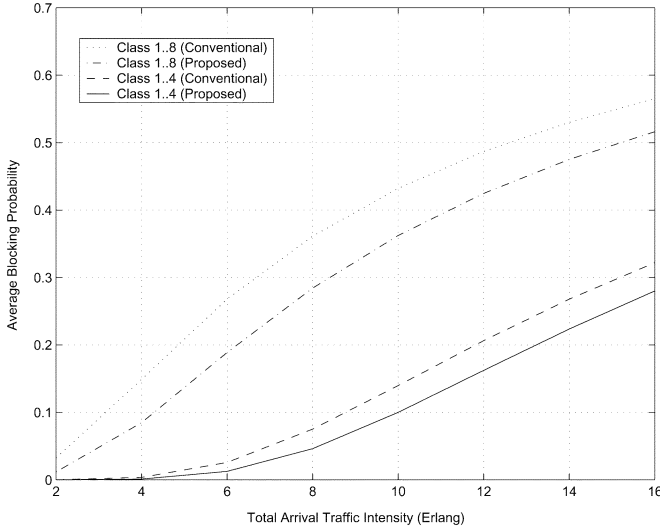
Fig. 9. Overall comparison on the proposed MRBS scheme and conventional CDMA.

In addition, we have explored fraction-valued MRCS for spreading factor optimization in single-code CDMA systems with OVSF technique. From our numerical analysis, resource utilization efficiency under the proposed flexible rate scheme is improved. It can accommodate a higher system traffic intensity. It is interesting and challenging to have further investigation on the MRCS.

APPENDIX A
EXPECTED BAE IN SINGLE CODE CASE

Let $a_i = 15 \times 2^i$ kb/s, where $i = 0, 1, \ldots, 7$. Since R_q is uniformly distributed between a_1 and a_7, the expect BAE is given in the following:

$$\sum_{i=0}^{6} \left(\Pr(R_q \in [a_i, a_{i+1}]) \times \frac{\mathrm{E}[R_q | R_q \in [a_i, a_{i+1}]]}{R_p} \right)$$
$$= \sum_{i=0}^{6} \left(\frac{a_{i+1} - a_i}{a_7 - a_0} \times \frac{0.5(a_i + a_{i+1})}{a_{i+1}} \right)$$
$$= \frac{3}{4} \sum_{i=0}^{6} \left(\frac{a_{i+1} - a_i}{a_7 - a_0} \right)$$
$$= \frac{3}{4}.$$

APPENDIX B
MRBS: ASYMPTOTIC MEAN AND PERIODICITY

For any real number p, the asymptotic mean of the MRBS $s(n) = \lceil (n+1)p \rceil - \lceil np \rceil$ is given by the following:

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} s(k) = \lim_{n \to \infty} \frac{\lceil np \rceil}{n} = p.$$

For any positive rational number $p = m/l$, where m and l are relatively prime integers and $0 < m \le l$, let $n = il + j$, where $i, j \ge 0$ and $j < l$. Since $s(n) = \lceil (il + j + 1)m/l \rceil -$

$\lceil (il + j)m/l \rceil = \lceil (j+1)m/l \rceil - \lceil jm/l \rceil$, $s(n) = s(j)$ for $0 \le j \le l - 1$. So, the sequence has a period of l.

APPENDIX C
UPPER AND LOWER BOUNDS ON $S(j) - S_{\tilde{M}}(j)$

As defined in *Proposition 1*, we have the following expressions on $S(j)$ and $S_{\tilde{M}}(j)$, respectively:

$$S(j) = \sum_{i=1}^{M} \sum_{n=0}^{j} (\lceil (n+1)p_i \rceil - \lceil n \times p_i \rceil) = \sum_{i=1}^{M} \lceil (j+1)p_i \rceil$$

$$S_{\tilde{M}}(j) = \sum_{n=0}^{j} \left(\left\lceil (n+1) \sum_{i=1}^{M} p_i \right\rceil - \left\lceil n \sum_{i=1}^{M} p_i \right\rceil \right)$$
$$= \left\lceil (j+1) \sum_{i=1}^{M} p_i \right\rceil.$$

Let $(j+1)p_i = a_{i,j} + b_{i,j}$, where $a_{i,j}$ is a positive integer and $0 \le b_{i,j} < 1$.

$$S(j) - S_{\tilde{M}}(j) = \sum_{i=1}^{M} \lceil a_{i,j} + b_{i,j} \rceil - \left\lceil \sum_{i=1}^{M} (a_{i,j} + b_{i,j}) \right\rceil$$
$$= \sum_{i=1}^{M} \lceil b_{i,j} \rceil - \left\lceil \sum_{i=1}^{M} b_{i,j} \right\rceil.$$

If $b_{i,j} = 0, \forall i, \sum_{i=1}^{M} \lceil b_{i,j} \rceil - \left\lceil \sum_{i=1}^{M} b_{i,j} \right\rceil = 0$. Then, $S(j) - S_{\tilde{M}}(j) = 0$.

Otherwise, assume there are l of $b_{i,j} \ne 0$, where $1 \le l \le M$. So, $\sum_{i=1}^{M} \lceil b_{i,j} \rceil = l$. Since $0 \le b_{i,j} < 1$, $0 < \sum_{i=1}^{M} b_{i,j} < l$. Consequently, $1 \le \left\lceil \sum_{i=1}^{M} b_{i,j} \right\rceil \le l$. We have $0 \le \sum_{i=1}^{M} \lceil b_{i,j} \rceil - \left\lceil \sum_{i=1}^{M} b_{i,j} \right\rceil \le l - 1$, where $1 \le l \le M$.

Hence, $0 \le S(j) - S_{\tilde{M}}(j) \le M - 1$.

APPENDIX D
UPPER AND LOWER BOUNDS ON THE MEAN OF MRBS

Proposition 5: If there are totally j "1"s within k slots in a portion of MRBS, where $k \ge 1$ and $j \le k$, then the corresponding mean p obeys: $(j-1)/k < p < (j+1)/k$.

Proof: By the definition of MRBS, given any phase θ, $j = \lceil (k+n)p + \theta \rceil - \lceil np + \theta \rceil$. Let $np + \theta = l + b$, where l is the integer part and $0 \le b < 1$. So, $j = \lceil kp + b \rceil - \lceil b \rceil$.

If $b = 0$, $j = \lceil kp \rceil \Rightarrow j - 1 < kp \le j \Rightarrow ((j-1)/k) < p \le (j/k)$.

If $b \ne 0$, $j = \lceil kp + b \rceil - 1 \Rightarrow \lceil kp + b \rceil = j + 1 \Rightarrow j < kp + b \le j + 1 \Rightarrow ((j-1)/k) < p < ((j+1)/k)$.

Hence, $(j-1)/k < p < (j+1)k$. Since, by definition, $0 < p \le 1$, we can have a tighter bound on p such that $\max\{(j-1)/k, 0\} < p \le \min\{(j+1)/k, 1\}$. ∎

APPENDIX E
PROOF ON PROPOSITION 3 BY INDUCTION

Let $P(j)$ be the statement $0 \le \tilde{b}_i(j) \le 1, \forall i$, for any nonnegative integer j. If $\max_i b_i(j) = 2$, there exists a unique i such

that $b_i(j) = 2$ and a slot is allocated for this user. Buffer overflows if $\tilde{b}_i(j) > 1$. On the other hand, since a user with buffered data has the highest priority to get a service slot, $\tilde{b}_i(j) \geq 0$ if there exists $b_i(j) > 0$ when $s_{\tilde{M}}(j) = 1$.

When $j = 0$, by the property of MRBS, $s_i(0) = \lceil p_i \rceil = 1$ as $p_i > 0$. So, $b_i(0) = s_i(0) = 1, \forall i$. Similarly, $s_{\tilde{M}}(0) = 1$. This service slot will be assigned to the user with buffered data and the highest priority. So, $0 \leq \tilde{b}_i(0) \leq 1$ and $P(0)$ is true.

Assume $P(j)$ holds up to $k - 1$. At time instant k, let the number of user(s) with nonempty buffer, $2 \geq b_i(k) > 0$, be f_k. The number of user(s) with empty buffer, $b_i(k) = 0$, is then $M - f_k$. Among the f_k user(s) with nonempty buffer, assume a_k of them have $b_i(k) = 2$. (For these a_k users, they have $\tilde{b}_i(k-1) = 1$ and arrivals at k.)

The sum of buffered data is equal to $f_k + a_k$. By *Proposition 1*, $f_k + a_k - s_{\tilde{M}}(k) \geq 0$. So, if $s_{\tilde{M}}(k) = 1$, $f_k + a_k \geq 1$. In this case, there is at least one user with nonempty buffer. So, there exists $b_i(k) > 0$ and the statement $\tilde{b}_i(k) \geq 0$ is true.

On the other hand, the statement $\tilde{b}_i(k) \leq 1$ is true if $a_k = 0$. So, assume $f_k \geq a_k \geq 1$. Let $c_k = a_k - s_{\tilde{M}}(k)$, where the service slot $s_{\tilde{M}}(k)$ is either 0 or 1. Claim $c_k = 0$. (In this case, the statement holds.) Suppose $c_k > 0$. Note in this case $f_k < M - 1$ since, by *Proposition 1*, $f_k + c_k \leq M - 1$. So, there exists more than one user with empty buffer. Label these users as $E_1, \ldots, E_{M-f_k}$. Let t_i be the latest time before k that E_i is allocated a service slot and $t_{\max} = \max_i t_i$. Note $t_{\max} < k$. Label those with $b_i(k) = 2$ among the f_k users, which have $\tilde{b}_i(k-1) = 1$ and an arrival at k, by $F_1, \ldots, F_{a_k}$. Label the left users, which have $b_i(k) = 1$, by $F_{a_k+1}, \ldots, F_{f_k}$.

Claim at time $t_{\max}$, all users $E_1, \ldots, E_{M-f_k}$ have empty buffer. This is also true for $F_1, \ldots, F_{a_k}$ because there is an arrival at k, so if the buffers for these users are nonempty, one of them would be served instead.

For any user in the group $F_{a_k+1}, \ldots, F_{f_k}$, if the user receives a service slot in $[t_{\max} + 1, k - 1]$, there must be an arrival at or after the service time excluding the served one since $b_i(k) = 1$. If the user receives a service slot, it must have empty buffer at $t_{\max}$. Otherwise, the user would be served instead of users $E_1, \ldots, E_{M-f_k}$ because there is an arrival at or after the service time. That is, among the users $F_{a_k+1}, \ldots, F_{f_k}$, either the user has empty buffer at time $t_{\max}$, (label as $F_{a_k+1}, \ldots, F_{a_k+r}$) or receives no service slot in $[t_{\max} + 1, k - 1]$ (label as $F_{a_k+r+1}, \ldots, F_{f_k}$).

Note in the interval $[t_{\max} + 1, k - 1]$, assume that there are n service slots. These service slots are assigned to users among $F_1, \ldots, F_{a_k+r}$ only. Let the number of slots assigned to F_i be u_i. So, $\sum_{i=1}^{a_k+r} u_i = n$. Note all these users have empty slot at time $t_{\max}$ and nonempty buffer at k, so users $F_1, \ldots, F_{a_k}$ (those with $b_i(k) = 2$) must have $u_i + 2$ arrivals in $[t_{\max} + 1, k]$ while users $F_{a_k+1}, \ldots, F_{a_k+r}$ (those with $b_i(k) = 1$) must have u_i+1 arrivals. By *Proposition 5*, $p_i > ((u_i + 2 - 1)/(k - t_{\max}))$ for $i = 1, \ldots, a_k$ while $p_i > ((u_i + 1 - 1)/(k - t_{\max}))$ for $i = a_k+1, \ldots, a_k+r$. So, $\sum_{i=1}^{a_k+r} p_i > ((n+a_k)/(k-t_{\max}))$.

Since the number of service slots in $[t_{\max} + 1, k]$ is equal to $n + s_{\tilde{M}}(k)$, by *Proposition 5*, $p_{\tilde{M}} < ((n + s_{\tilde{M}}(k) + 1)/(k - t_{\max}))$. So, if $c_k > 0$, $\sum_{i=1}^{a_k+r} p_i > p_{\tilde{M}}$ since $(n + a_k) - (n + s_{\tilde{M}}(k) + 1) = c_k - 1 > 0$. This is a contradiction. Therefore,

$c_k = 0$. Consequently, $\tilde{b}_i(k) \leq 1$. Since $0 \leq \tilde{b}_i(k) \leq 1$, $P(k)$ is true.

Hence, the statement $0 \leq \tilde{b}_i(j) \leq 1, \forall i$, is always true for any nonnegative integer j.

APPENDIX F
ORDER OF TRANSMISSION BY MRBS-SSA: DETERMINISTIC AND PERIODIC

Each MRBS is deterministic and has a period l_i with respect to $p_i = m_i/l_i$. Let $N = \text{LCM}_i(l_i)$. We can generate MRBS for each i in N slots. Considering all i in an N-slot-block, this block is also deterministic and periodic of duration N. By the property of MRBS, $\forall i, s_i(N) = s_i(0) = 1$ and, similarly, $s_{\tilde{M}}(N) = s_{\tilde{M}}(0) = 1$. By *Proposition 1*, we have $0 \leq S(N) - S_{\tilde{M}}(N) \leq M-1$. So, $S(N-1) - S_{\tilde{M}}(N-1) = 0$. As defined, $\sum_{i=1}^{M} \tilde{b}_i(N-1) = S(N-1) - S_{\tilde{M}}(N-1)$. By *Proposition 3*, $0 \leq \tilde{b}_i(j) \leq 1$ for any j. So, $\tilde{b}_i(N-1) = 0$. This implies $b_i(N) = b_i(0)$. By MRBS-SSA, each service slot is assigned to the highest priority user in a deterministic way. Since each N-slot-block has the same pattern, initial condition $b_i(N) = b_i(0)$, and service slot sequence, MRBS-SSA produces the same scheduled block, which repeats the pattern in an N-slot-long. Hence, we can conclude that the order of transmission for all users is deterministic and periodic.

APPENDIX G
MRCS: ASYMPTOTIC MEAN AND PERIODICITY

For any real number p, the asymptotic mean of the MRCS is given by the following.

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} c(k) = \lim_{n \to \infty} \frac{(p_U - p_L) \times \left\lceil \frac{np}{(p_U - p_L)} \right\rceil}{n} = p.$$

For any positive rational number $p = m/l$, where m and l are relatively prime integers and $0 < m \leq l$, let $n = il + j$, where $i, j \geq 0$ and $j < l$. Since $1/(p_U - p_L)$ is an integer, $\lceil ((np)/(p_U - p_L)) \rceil = \lceil ((il + j)/(p_U - p_L)) \times (m/l) \rceil = ((im)/(p_U - p_L)) + \lceil ((jm)/(l(p_U - p_L))) \rceil$. So, $c(n) = (p_U - p_L) \times (\lceil (((j+1)m)/(l(p_U - p_L))) \rceil - \lceil ((jm)/(l(p_U - p_L))) \rceil)$. We have $c(n) = c(j)$ for $0 \leq j \leq l - 1$. As a result, $c(n)$ is independent of i and has a period l.

APPENDIX H
DELAY BOUND ON MRBS

From the result of *Proposition 3*, since $0 \leq \tilde{b}_i(k) \leq 1$, a buffer of size "1" at each user i is sufficient. This means that, for each MRBS traffic, buffered packet will be transmitted before the accumulation of coming one. So, the scheduling delay is at most the duration between two successive "1"s. Including the packet arriving time in delay measurement, this is equal to $k+1$ slot durations, where k is the maximum number of successive "0"s between two "1"s. By *Proposition 5*, if there are totally j "1"s within k slots in a portion of MRBS, the corresponding mean p obeys $(j - 1)/k < p < (j + 1)/k$. We can derive the upper bound on $k + 1$ by the substitution $j = 0$. Therefore,

$p < 1/k$. This implies $k < 1/p$. Since $0 < p \leq 1$, we can let $1/p = a + b$, where a is the integer part and $0 \leq b < 1$.

When $b = 0$, $1/p = a \neq 0$ and $1/p = \lceil 1/p \rceil$. So, $k < 1/p \Rightarrow k \leq 1/p - 1 \Rightarrow k + 1 \leq 1/p$. We have $k + 1 \leq \lceil 1/p \rceil$.

When $b \neq 0$, $k < 1/p \Rightarrow k \leq \lfloor 1/p \rfloor \Rightarrow k \leq a$. Since $\lceil 1/p \rceil = a + 1$, $k + 1 \leq a + 1 = \lceil 1/p \rceil$.

Therefore, the scheduling delay experienced by MRBS of p_i in multiaccess control is always less or equal to $\lceil 1/p_i \rceil$ slot lengths.

REFERENCES

[1] L. Zhu and S. B. Slimane, "A multicarrier allocation (MCA) scheme for variable-rate 3Gwireless systems," *IEEE Commun. Mag.*, vol. 38, no. 10, pp. 86–91, Oct. 2000.

[2] L. Jorguseki, E. Fledderus, J. Farserotu, and R. Prasad, "Radio resource allocation in third generation mobile communication systems," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 117–123, Feb. 2001.

[3] R. Prasad and T. Ojanpera, "An overview of CDMA evolution toward wideband CDMA," *IEEE Commun. Surveys*, vol. 1, no. 1, pp. 2–29, Fourth Quarter 1998.

[4] E. Dahlman, B. Gudmundson, M. Nilsson, and A. Skold, "UMTS/IMT-2000 based on wideband CDMA," *IEEE Commun. Mag.*, vol. 36, no. 9, pp. 78–80, Sep. 1998.

[5] H. Holma and A. Toskala, *WCDMA for UMTS: Radio Access for Third Generation Mobile Communications*. New York: Wiley, 2000.

[6] Q. Bi, G. I. Zysman, and H. Menkes, "Wireless mobile communications at the start of the 21st century," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 110–116, Jan. 2001.

[7] Third Generation Partnership Project (3GPP). [Online]. Available: http://www.3gpp.org

[8] Third Generation Partnership Project 2 (3GPP2). [Online]. Available: http://www.3gpp2.org

[9] F. Adachi, M. Sawahashi, and K. Okawa, "Tree-structured generation of orthogonal spreading codes with different lengths for forward link of DS-CDMA mobile radio," *Electron. Lett.*, vol. 33, no. 1, pp. 27–28, Jan. 1997.

[10] J. B. Kim and M. L. Honig, "Resource allocation for multiple classes of DS-CDMA traffic," *IEEE Trans. Veh. Technol.*, vol. 49, no. 2, pp. 506–519, Mar. 2000.

[11] J. Groe and L. E. Larson, *CDMA Mobile Radio Design*. Reading, MA: Artech House, 2000, pp. 252–257.

[12] R. Steele and L. Hanzo, *Mobile Radio Communications: Second and Third Generation Cellular and WATM Systems*. New York: Wiley, 1999, pp. 897–954.

[13] H. C. Fong, "Spreading factor optimization and random access stability control for IMT-2000," M.Phil. thesis , The Chinese Univ. Hong Kong, Hong Kong, 2000.

[14] H. Azad and A. H. Aghvami, "Multirate spread spectrum direct sequence CDMA techniques," in *Proc. IEE Colloquium on Spread Spectrum Technology Radio Communications Systems*, Apr. 1994, pp. 4/1–4/5.

[15] B. Hajek, "Extremal splitting of point process," *Math. Oper. Res.*, vol. 10, no. 4, Nov. 1985.

[16] ——, "The proof of a folk theorem on queueing delay with applications to routing in networks," *J. Assoc. Comput.*, vol. 30, pp. 834–851, 1983.

[17] C. S. Chen and W. S. Wong, "Bandwidth allocation optimization for 3G wireless multimedia systems," in *Proc. Int. Conf. Internet Computing*, vol. 1, Jun. 2001, pp. 75–81.

[18] T. Minn and K. Y. Siu, "Dynamic assignment of orthogonal variable-spreading-factor codes in W-CDMA," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 8, pp. 1429–1440, Aug. 2000.

[19] Y. Yang and T.-S. Yum, "Maximally flexible assignment of orthogonal variable spreading factor codes for multi-rate traffic," *IEEE. Trans. Wireless Commun.*, to be published.

[20] E. Arthur and J. S. Kaufman, "Sizing a message store subject to blocking criteria," in *Performance of Computer Systems*, M. Arato, A. Butrimenko, and E. Gelenbe, Eds. Amsterdam, The Netherlands: North Holland, 1979, pp. 547–564.

[21] J. S. Kaufman, "Blocking in a shared resource environment," *IEEE Trans. Commun.*, vol. COM–29, no. 10, pp. 1474–1481, Oct. 1981.
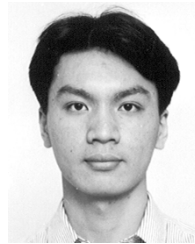
Chung Shue Chen (S'02) was born in China in 1977. He received the B.Eng. and M.Phil. degrees in information engineering in 1999 and 2001, respectively, from the Chinese University of Hong Kong, Shatin, Hong Kong, where he is currently working toward the Ph.D. degree.

His research interests include resource allocation, scheduling, QoS guarantee, and multiple access control in wireless communication systems.

Wing Shing Wong (M'81–SM'90–F'02) received the combined master and bachelor degree *(summa cum laude)*, from Yale University, New Haven, CT, in 1976, and the M.S. and Ph.D. degrees from Harvard University, Cambridge, MA, in 1978 and 1980, respectively.

He joined AT&T Bell Laboratories, Holmdel, NJ, in 1982. From 1987 to 1992, he managed a Technical Staff working on a number of research and development projects and consulting activities. He joined the Chinese University of Hong Kong in 1992 and is now a Professor of information engineering. He was the Chairman of the Department of Information Engineering, Chinese University of Hong Kong, from 1995 to 2003. Currently, he is serving as Science Advisor at the Innovation and Technology Commission of the HKSAR government. He is also a Visiting Professor of Southeast University, Nanjing, China. He has published over 100 refereed journal and conference papers. He is the Co-Editor-in-Chief of a new journal, *Communications in Information and Systems.* He is actively involved in a variety of research and development projects including topics such as mobile communication systems, search engine development, and information issues in estimation and control.

Dr. Wong was an Associate Editor of the IEEE TRANSACTIONS ON AUTOMATIC CONTROL for four years.