Spreading Codes for Direct Sequence CDMA and Wideband CDMA Cellular Networks

Esmael H. Dinan and Bijan Jabbari

George Mason University

ABSTRACT This article presents an overview of the spreading techniques for use in direct sequence CDMA cellular networks. We briefly review the theoretical background for sequences used in CDMA and wideband CDMA, and discuss the main characteristics of the Maximal length, Gold, and Kasami sequences, as well as variable- and fixed-length orthogonal codes. We also describe different methods of multiple spreading for channelization and scrambling in CDMA and W-CDMA realizations.

The continuous growth in traffic volume and emergence of new services have begun to change the structure of wireless networks. Future mobile communications systems will be characterized by high throughput, integration of services, and flexibility. The high capacity required to support these characteristics can be obtained by using the spectrum as efficiently as possible and by flexibility in radio resource management. Spread spectrum code-division multiple access (CDMA) approaches have been proposed for a variety of digital cellular mobile and wireless personal communications systems. Cellular CDMA systems offer the potential of high spectrum efficiency. This capacity advantage, together with other features such as soft capacity (or graceful degradation), multipath resistance, inherent frequency diversity and interference rejection, and the potential use of advanced antenna and receiver structures, have contributed to growing interest in this technology for proposed second- and especially third-generation cellular mobile systems.

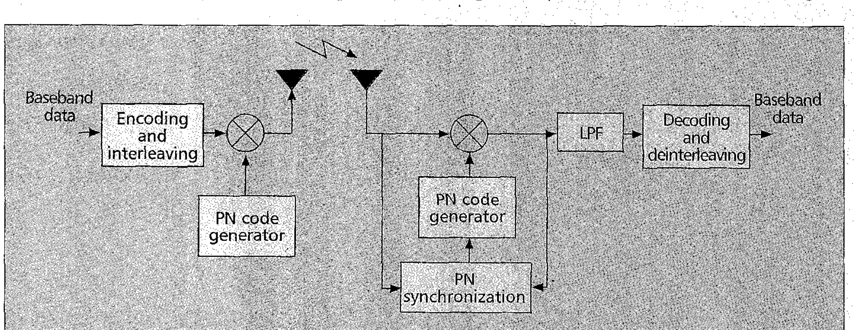
A spread spectrum CDMA scheme is one in which the transmitted signal is spread over a wide frequency band, much wider than the minimum bandwidth required to transmit the information being sent [1, 2]. It employs a waveform that for all purposes appears random to anyone but the intended receiver of the transmitter waveform. Actually, for ease of both generation and synchronization by the receiver, the waveform is pseudorandom, meaning that it can be generated by mathematically precise rules, but statistically it nearly satisfies the reuirements of a truly random sequence. In spread spectrum CDMA all users use the same bandwidth, but each transmitter is assigned a distinct code.

A block diagram of the baseband model of a direct sequence (DS) CDMA modulator and demodulator is shown in Fig. 1. Spreading consists of multiplying the input data by a pseudonoise (PN) sequence, the bit rate of which is much higher than the data bit rate. This increases the data rate while adding redundancy to the system. The ratio of PN sequence bit rate to data bit rate is called the *spreading factor* (SF). The resulting waveform is wideband, noiselike, and balanced in phase, and has a flexible timing structure. In case there are two different I and Q branches, each channel can be spread separately. In order to minimize the overall envelope variations and achieve high amplifier efficiency, a complex spreading technique can also be used, as shown in Fig. 2. When the signal is received, the spreading is removed from the desired signal by multiplying with the same PN code that is exactly synchronized to the received PN. When despreading is applied to the interference generated by other users' signals, there is no despreading. That is, each spread spectrum signal should behave as if it were uncorrelated with every other spread signal using the same band. Therefore, CDMA codes are designed to have very low cross-correlation. Orthogonal non-PN spreading codes with zero cross-correlation were employed to improve the bandwidth efficiency of cellular CDMA systems. The importance of the code sequence to spread spectrum communications is difficult to overemphasize, for the type of code used, its length, and its chip rate set bounds on the capability of the system that can be changed only by changing the code.

In this article, an analytical approach is taken to illustrate the main characteristics such as code length, auto-correlation, and cross-correlation of some important PN and orthogonal sequences. The Maximal-length, Gold, and Kasami sequences are described in the second, third, and fourth sections, respectively. Fixed- and variable-length orthogonal codes are discussed in the fifth and sixth sections. The seventh section outlines the importance of multiple spreading in cellular CDMA networks and presents different techniques used to spread the information signal, and finally the last section concludes the article.

MAXIMAL-LENGTH SEQUENCES

Figure 3 shows the structure of the well-known maximal-length linear feedback shift register sequences. Maximal-length sequences (m-sequences) are, by definition, the largest



■ **Figure 1.** *The baseband model of a DS-CDMA transceiver.*

codes that can be generated by a given shift register or a delay element of a given length [3]. Each clock time the register shifts all contents to the right. The sequence a_i is generated according to the recursive formula

$$a_i = c_1 a_{i-1} + c_2 a_{i-2} + \cdots + c_n a_{i-n} = \sum_{k=1}^{n} c_k a_{i-k} \quad (1)$$

Here, all terms are binary (0 or 1), and addition and multiplication are mudulo-2. We define the generating function of the sequence as

$$G(D) \overset{\Delta}{=} a_0 + a_1 D + a_2 D^2 + \cdots = \sum_{i=0}^{\infty} a_i D^i \quad (2)$$

where D is the delay operator.

If the sequence is periodic with period N (i.e., $a_i = a_{i+N}$), then $G(D)$ can be written as

$$\begin{aligned} G(D) &= \sum_{i=0}^{\infty} D^{iN}\left(a_0 + a_1 D + a_2 D^2 + \cdots + a_{N-1} D^{N-1}\right) \\ &= \frac{a_0 + a_1 D + a_2 D^2 + \cdots + a_{N-1} D^{N-1}}{1 + D^N} \end{aligned} \quad (3)$$

Combining Eqs. (1) and (2), we may reduce the latter to the finite recurrence relation

$$\begin{aligned} C(D) &\overset{\Delta}{=} \sum_{i=0}^{\infty} a_i D^i = \sum_{i=0}^{\infty}\sum_{k=1}^{n} c_k a_{i-k} D^i = \sum_{k=1}^{n} c_k D^k \left[\sum_{i=0}^{\infty} a_{i-k} D^{i-k}\right] \\ &= \sum_{k=1}^{n} c_k D^k \left[a_{-k} D^k + \cdots + a_{-1} D^{-1} + G(D)\right] \end{aligned}$$

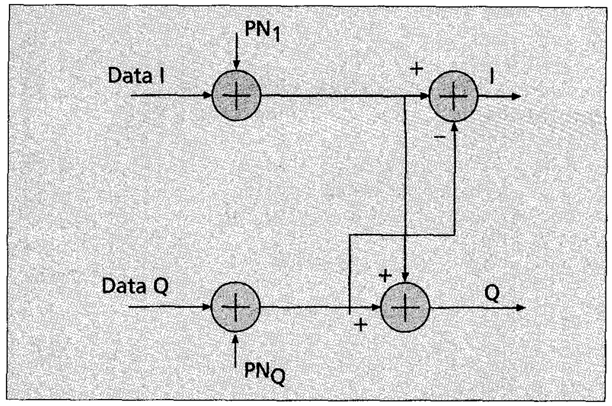
From this, $G(D)$ can be expressed as a ratio of finite polynomials

$$G(D) = \frac{\sum_{k=1}^{n} c_k D^k \left(a_{-k} D^{-k} + \cdots + a_{-1} D^{-1}\right)}{1 + \sum_{k=1}^{n} c_k D^k} \overset{\Delta}{=} \frac{g_0(D)}{f(D)} \quad (4)$$

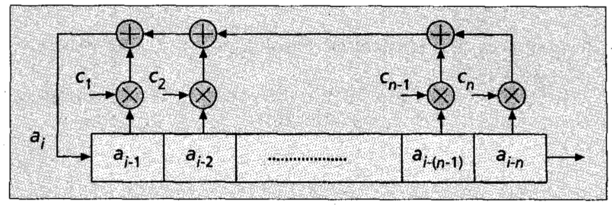
$f(D)$ is called the characteristic polynomial of the linear feedback shift register (LFSR) sequence generator, depends solely on the connection vector $c_1, \ldots, c_n$, and determines the main characteristics of the generated sequence. The polynomial $g_0(D)$ depends as well on the initial condition vector $a_{-n}, a_{-(n-1)}, \ldots, a_{-1}$ and determines the phase shift of the sequence. If the sequence generated by $G(D)$ has period N, it can be shown from Eqs. (3) and (4) that $f(D)$ must divide $1 + D^N$. It has been proven that every LFSR sequence is periodic with period $N \le 2^n - 1$. This leads us to define the m-sequence as an LFSR sequence whose period $N = 2^n - 1$ for all nonzero initial vectors. A necessary condition for $G(D)$ to generate an m-sequence is that $f(D)$ be irreducible (nonfactorable), that is, $f(D) \ne f_1(D)f_2(D)$. Suppose that $f(D) = f_1(D)f_2(D)$ with $f_1(D)$ of degree n_1 and $f_2(D)$ of degree n_2 (it is clear that $n_1 + n_2 = n$). Then we can write, by partial fractions,

$$\frac{1}{f(D)} = \frac{\alpha(D)}{f_1(D)} + \frac{\beta(D)}{f_2(D)}; \qquad \deg\alpha(D) < n_1, \qquad \deg\beta(D) < n_2$$

The maximum period of the sequence generated by the first term is $2^{n1} - 1$ and that of the second term is $2^{n2} - 1$. Hence, the period of the sequence generated by $1/f(D)$ is less than or equal to the least common multiple of $(2^{n1} - 1, 2^{n2} - 1) < 2^n - 3$. This is a contradiction, since if $f(D)$ were maximal length, the period of the sequence would be $2^n - 1$. Unfortunately this condition is not sufficient. The irreducible polynomials that generate an m-sequence are called *primitive*. A primitive polynomial of degree n is simply one for which the period of the coefficients of $1/f(D)$ is $2^n - 1$. These polynomials exist for all degrees $n > 1$. Golomb [3] showed that the number of primitive polynomials of degree n is equal to[1]



■ **Figure 2.** *The complex spreading technique in W-CDMA.*



■ **Figure 3.** *M-sequence generator structure.*

$$N_p(n) = \frac{2^n - 1}{n} \prod_{i=1}^{k} \frac{P_i - 1}{P_i} \quad (5)$$

where $\{P_i, i = 1, 2, \ldots, k\}$ is the prime decomposition of $2^n - 1$, that is,

$$2^n - 1 = \prod_{i=1}^{k} P_i^{n_i}, \text{ where } n_i \text{ is an integer.}$$

Table 1 presents the code length and the number of m-sequences for some values of n. The number of codes increases very fast when n increases. References [4, 5] give tables of primitive polynomials for $m \le 40$, sufficient to generate sequences of period up to $2^{40} - 1 \cong 10^{12}$, enough for most purposes.

Figure 4 illustrates a three-stage LFSR sequence generator. Considering the feedback connections, the sequence a_i is generated according to the formula

$$a_i = a_{i-2} + a_{i-3}$$

Let us consider the initial condition of the registers to be $a_{-3} = 1, a_{-2} = a_{-1} = 0$; then the sequence a_i is equal to ($i \ge 0$)

$$a = \{1011100,1011100,101\ldots\}$$

This sequence is periodic with period 7.

On the other hand, we have $f(D) = 1 + D^2 + D^3$, $g(D) = 1$, and $G(D) = 1/(1 + D^2 + D^3)$; this division yields to

$$\begin{aligned} G(D) &= 1 + D^2 + D^3 + D^4 + D^7 + D^9 + D^{10} + D^{11} + D^{14} + \cdots \\ &= \frac{1 + D^2 + D^3 + D^4}{1 + D^7} \end{aligned}$$

[1] *This is also given by [2]*

$$\frac{\Phi_p(N)}{n},$$

where Φ_p(m) is the Euler totient function, that is, the number of integers less than m *which are relatively prime to* m.

The coefficients of $G(D)$ present exactly the same sequence obtained by the recursive equation.

It can be demonstrated that m-sequences satisfy the following three randomness properties in every period of length $N = 2^n - 1$:

- The number of ones and zeros is nearly equal.
- Half the runs of ones and zeros have length 1, 1/4 have length 2, 1/8 length 3, and $1/2^k$ length k ($k < n$).
- Sequence autocorrelation

$$C(k) = \sum_{n=1}^{N} a'_n a'_{n+k} = N\delta(k)$$

where $\delta(k)$ is the Kronecker delta and $a'_n = 1 - 2a_n$ is the $\pm$ 1 sequence.

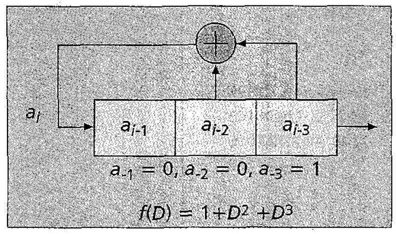


Figure 4. *An example of a maximal LFSR sequence generator.*

If the code waveform $p(t)$ is the squarewave equivalent of the sequence a'_n with pulse duration T_c, and if we define

$$q(\tau) \stackrel{\Delta}{=} \begin{cases} 1 - \dfrac{|\tau|}{T_c}, & |\tau| \le T_c \\ 0, & \text{otherwise} \end{cases}$$

then if $N >> 1$,

$$C(\tau) \cong N \sum_i q(\tau - iNT_c),$$

this function is shown in Fig. 5.

Autocorrelation refers to the degree of correspondence between a sequence and a phase-shifted replica of itself. This characteristic autocorrelation is used to great advantage in communications, and it is of most interest in choosing code sequences that give the least probability of false synchronization. With the above characteristics an m-sequence is indistinguishable from a pure random code when N is large. Note, however, that when correlation is over a much shorter span than N, $C(k)$ is a random variable with a distribution similar to the sum of binary i.i.d variables just as for random sequences.

Cross-correlation of two codes is of similar importance. Cross-correlation is the measure of agreement between two different codes and is given by

$$R_c(k) = \sum_{n=1}^{N} a'_n b'_{n+k} \tag{6}$$

where a_n and b_n are the elements of the two sequences with period N. The m-sequences are not immune to cross-correlation problems, and they may have large cross-correlation values. Welch obtained the following lower bound on the cross-correlation between any pair of binary sequences of period N in a set of M sequences [6]:

$$R_c(k) \ge N\sqrt{\frac{M-1}{MN-1}} \cong \sqrt{N} \tag{7}$$

m-sequences are of great interest in synchronous cellular spread spectrum networks. Since the correlation between different shifts of an m-sequence is almost zero, they can be used as different codes with an excellent correlation property. In synchronous networks all base stations and users use a common time reference using the Global Positioning System (GPS) and the pilot sent by the base stations, respectively. Then each base or mobile station is identified by a unique offset of its PN binary sequence in forward and reverse channels.

The composite codes presented in the next sections are of great utility when cross-correlation is a prime consideration (especially when the users are asynchronous to each other). It is also of some interest to note that even when the codes used exhibit excellent cross-correlation properties when averaged over their entire length, short-term cross-correlation, quite effective in disturbing communications, can occur. A number of papers in the literature deal with this problem, among them those by Lindholm [7], Wainberg [8], and Fredricsson [9].

Gold Sequences

A goal of spread spectrum system designers for a multiple access system is to find a set of spreading codes or waveforms such that as many users as possible can use a band of frequencies with as little mutual interference as possible. Gold sequences are useful because of the large number of codes they supply. They can be chosen so that over a set of codes available from a given generator, the cross-correlation between the codes is uniform and bounded [4, 10, 11]. In this section we describe how we can generate a set of Gold sequences.

Consider an m-sequence represented by a binary vector a of length N, and a second sequence a' obtained by sampling every qth symbol of a. The second sequence is said to be a *decimation* of the first, and the notation $a' = a[q]$ is used to indicate that a' is obtained by sampling every qth symbol of a. $a' = a[q]$ has period N if and only if $\gcd(N, q) = 1$, where "gcd" denotes the greatest common divisor. Any pair of m-sequences having the same period N can be related by $a' = a[q]$ for some q.

Two m-sequences a and a' are called the *preferred pair* [4] if:

- $n \neq 0 \pmod 4$; that is, n odd or $n = 2 \pmod 4$
- $a' = a[q]$, where q is odd and either $q = 2^k + 1$ or $q = 2^{2k} - 2^k + 1$
- $\gcd(n,k) = \begin{cases} 1 & \text{for } n \text{ odd} \\ 2 & \text{for } n = 2(\text{mod}4) \end{cases}$

The cross-correlation spectrum between a preferred pair is three-valued, where those three values are $-t(n)$, -1, $t(n) - 2$ where

$$t(n) = \begin{cases} 1 + 2^{\frac{n+1}{2}} & \text{for } n \text{ odd} \\ 1 + 2^{\frac{n+2}{2}} & \text{for } n \text{ even} \end{cases}$$

Finding preferred pairs of m-sequences is necessary in defining sets of Gold codes. Let a and a' represent a preferred pair of m-sequences having period $N = 2^n - 1$. The family of codes defined by $\{a, a', a + a', a + D\,a', a + D^2a', \ldots, a + D^{N-1}\,a'\}$ where D is the delay element is called the set of Gold codes for this preferred pair of m-sequences. It can be proved that the $N + 1$ elements of a Gold codes set have the property that any pair of codes in the set has the above three-valued cross-correlation spectrum. With the exception of sequences a and a', the set of Gold sequences are not maxi-

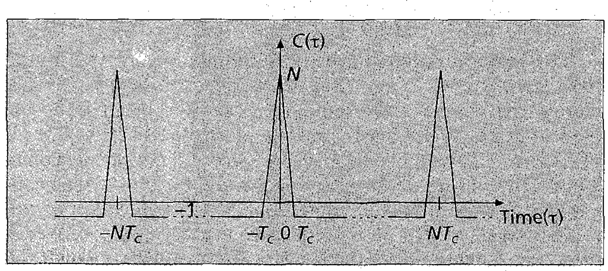
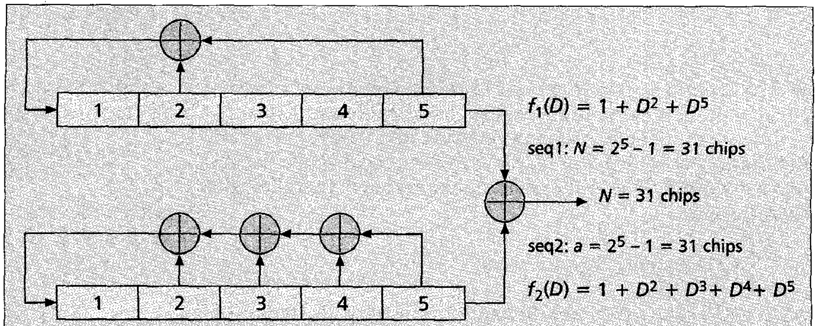


Figure 5. *Autocorrelation function of* p(t).

mal sequences. Hence, their autocorrelation functions are not two-valued, and it takes the same three values as cross-correlation. It is interesting to compare the peak cross-correlation value of Gold sequences with the known lower bound developed by Welch. For large N, the Welch bound is lower by $\sqrt{2}$ for n odd and by 2 for n even.

Regarding the characteristic polynomial of the preferred pairs, let α be any primitive element of $GF(2^n)$. Let $f_1(D)$ be an irreducible polynomial of degree n with the root of α, $f_1(D)$ is termed the minimal polynomial of α; and let $f_t(D)$ be the minimal polynomial of $\alpha^{t(n)}$, where both $f_1(D)$ and $f_t(D)$ are of degree n. Then the sequences generated by $f_1(D)$ and $f_t(D)$ are two preferred m-sequences. It can be shown that the characteristic polynomial of the sum of these two sequences is $f_1(D)f_t(D)$. The significance of this theorem is that it tells how to select shift register connections, which will generate maximal linear sequences with a known bound on the cross-correlation function. A list of characteristic polynomials of the preferred pairs can be found in [12]. Figure 6 illustrates the generation of 33 Gold codes of length 31 by adding the output of two LFSR sequence generators. The complete family of Gold codes for this generator is obtained using different initial loads of the shift register. The period of any code in the family is N, which is the same as the period of m-sequences.

For example [10], for $n = 13$ ($N = 2^{13} - 1 = 8191$) there are 630 maximal sequences and there exist pairs of these sequences whose correlation values are at $R_c = 703$, while Gold sequences guarantee the selection of pairs of sequences such that $R_c \leq 129$. We note further that for purely random sequences with this length we would expect the cross-correlation function to exceed $2\sigma = 2 \times 8192^{1/2} = 180$ for 5 percent of the correlation values; hence, linear sequences chosen in accordance with this technique perform better with respect to their cross-correlation properties than purely random sequences.



The 25 + 1 = 33 codes generated by this structure:

Sequence 1: 1111100011011101010000100101100
Sequence 2: 1111100100110000101101010001110
0 shift combination: 0000000111101101111101110100010
1 shift* combination: 0000101010111100001010000110001
.....
30 shift combination: 1000010001000101000110001101011

*Shift of sequence 2 to the left

■ **Figure 6.** *An illustration of generating a Gold code set.*

KASAMI SEQUENCES

Kasami sequence sets are one of the important types of binary sequence sets because of their very low cross-correlation [4, 13]. There are two different sets of Kasami sequences. A procedure similar to that used for generating Gold sequences will generate the *small set* of Kasami sequences with $M = 2^{n/2}$ binary sequences of period $N = 2^n - 1$, where n is even. In this procedure we begin with an m-sequence a and we form the sequence a' by decimating a by $2^{n/2} + 1$. It can be verified that the resulting a' is an m-sequence with period $2^{n/2} - 1$. For example, if $n = 10$, the period of a is $N = 1023$ and the period of a' is 31. Hence, if we observe 1023 bits of sequence a', we will see 33 repetitions of the 31-bit sequence. Now, by taking $N = 2^n - 1$ bits of sequences a and a' we form a new set of sequences by adding, modulo-2, the bits from a and the bits from a' and all $2^{n/2} - 2$ cyclic shifts of the bits from a'. By including a in the set, we obtain a set of $2^{n/2}$ binary sequences of length $N = 2^n - 1$. If m-sequence a is generated by $f(D)$ and a' with $f'(D)$, all the elements of the small set of Kasami sequences can be generated by the generating function $f(D)f'(D)$. The autocorrelation and cross-correlation functions of these sequences take on the values from the set $\{-1, -(2^{n/2} + 1), 2^{n/2} - 1\}$. Hence, the maximum cross-correlation value satisfies the Welch lower bound for a set of $2^{n/2}$ sequences of length $N = 2^n - 1$, and the small set of Kasami sequences is optimal.

n	$N = 2^n - 1$	$N_p^{(n)}$
2	3	1
3	7	2
4	15	2
5	31	6
6	63	6
7	127	18
8	255	16
9	511	48
10	1023	60
11	2047	176
12	4095	144
13	8191	630
14	16,383	756
15	32,767	1800
16	65,535	2048
17	131,071	7710
18	262,143	8064
19	524,287	27,594
20	1,048,575	24,000

■ **Table 1.** *Code length and the number of maximal LFSR sequences.*

The *large set* of Kasami sequences again consists of sequences of period $2^n - 1$, for n even, and contains both the Gold sequences and the small set of Kasami sequences as subsets. Let m-sequences a' and a'' be formed by the decimation of a by $2^{n/2} + 1$ and $2^{(n+2)/2} + 1$, and take all sequences formed by adding a, a', and a'' with different shifts of a' and a''. The number of such sequences is $M = 2^{3n/2}$ if $n = 0$ (mod 4), and even larger, $M = 2^{3n/2} + 2^{n/2}$, if $n = 2$ (mod 4). All the values of auto-correlation and cross-correlation from members of this set are limited to the five values $\{-1, -1 \pm 2^{n/2}, -1 \pm 2^{n/2} + 1\}$. With this larger set of sequences, we have $|R_c(k)|_{\max} \leq 2^{(n+2)/2}$. The Welch bound is not asymptotically approached, but the packing of signal space is more efficient than for the Gold codes. Extending the generation function of the small set, let $f''(D)$ be the generating function of sequence a''; it can be deduced that all the sequences generated by $f(D)f'(D)f''(D)$ form the large set of Kasami sequences. These sequences are one of the candidates for the scrambling code in W-CDMA systems.

ORTHOGONAL CODES

Orthogonal functions are employed to improve the bandwidth efficiency of spread spectrum systems. Each mobile user uses one member of a set of orthogonal functions representing the set of symbols used for transmission. While there are many different sequences that can be used to generate an orthogonal set of functions, the Walsh and Hadamard sequences make useful sets for CDMA. The orthogonal functions have the following characteristic:

$$\sum_{k=0}^{M-1} \phi_i(k\tau)\phi_j(k\tau) = 0, \qquad i \neq j \tag{8}$$

where $\phi_i(k\tau)$ and $\phi_j(k\tau)$ are ith and jth orthogonal members of an orthogonal set respectively, M is the length of the set, and τ is the symbol duration.

Walsh functions are generated by mapping codeword rows of special square matrices called Hadamard matrices. These matrices contain one row of all zeros, and the remaining rows each have equal numbers of ones and zeros. Walsh functions can be constructed for block length $N = 2^n$. The Hadamard matrix of desired length can be generated by the following recursive procedure:

$$H_1 = [0],\ H_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},\ H_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \tag{9}$$

$$H_{2N} = \begin{bmatrix} H_N & H_N \\ H_N & \overline{H_N} \end{bmatrix}$$

where N is a power of 2 and the overscore denotes the binary complement of the bits in the matrix.

Hadamard matrices can also be obtained using the following formula:

$$K_{2^n} = \underbrace{K_2 \otimes K_2 \otimes \cdots \otimes K_2}_{n \text{ times}} = (K_2)^{\otimes n} \tag{10}$$

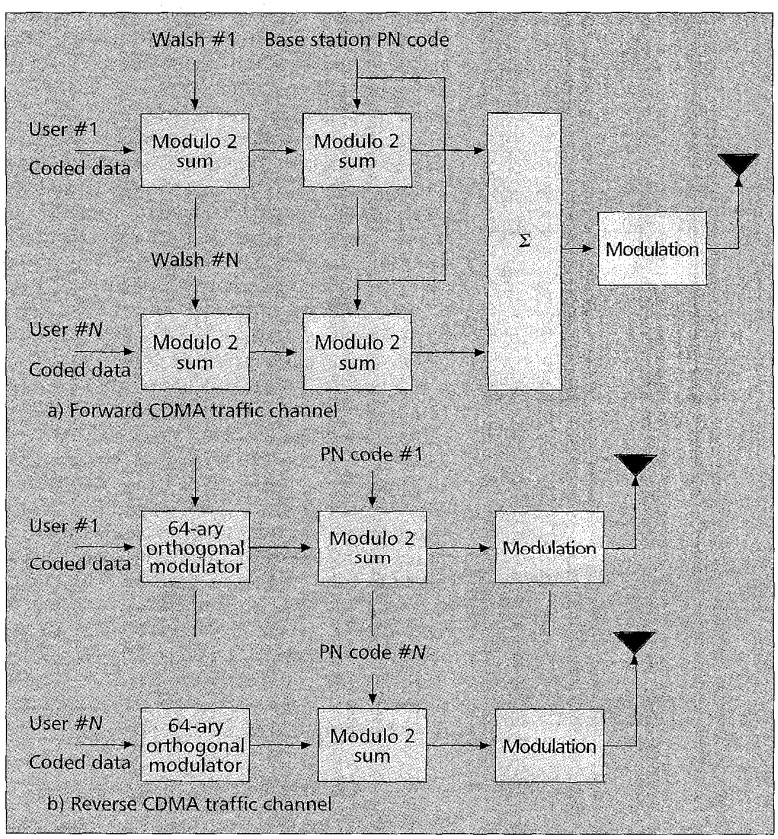
where the elements of K are the unique mapping of the elements of H, $\{0, 1\}$ onto the $\{1, -1\}$, respectively. Note that $\otimes$ is the Kronecker product and is defined as follows:

$$A_{n \times m} \otimes B_{k \times l} = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & A_{1m}B \\ \cdots & & \cdots & \cdots \\ & & & \\ a_{n1}B & a_{n2}B & \cdots & a_{nm}B \end{bmatrix}_{(n \times k) \times (m \times l)} \tag{11}$$

Each row of K presents a Walsh function. These functions have zero correlation between each other. In the spread spectrum transmitter, each bit is spread by a Walsh function; therefore, the spreading factor is equal to N. Orthogonal spreading codes can be used if all the users of the same channel are synchronized in time to the accuracy of a small fraction of one chip, because the cross-correlation between different shifts of Walsh functions is not zero (some functions are just the shifted version of others).

Another method can be used to modulate the orthogonal functions into the information stream of the CDMA signal. These functions can form orthogonal modulation symbols. With orthogonal symbol modulation, the information bitstream can be divided into blocks so that each block represents a nonbinary information symbol associated with a particular transmitted code sequence. If there are n bits per block, one of the set of $N = 2^n$ functions is transmitted in each symbol interval. The signal at the receiver is correlated with a set of N-matched filters, each matched to the code function of one symbol. The outputs from correlators are compared, and the symbol with the largest output is taken as the transmitted symbol. In this technique the signal spectrum is spread by the factor $2^n/n$, since all the users use the same orthogonal functions, the signal of one user cannot be distinguished from others. Therefore, in CDMA systems each user's signal is also spread by a distinct PN sequence after orthogonal modulation.

Figure 7 illustrates the application of Walsh functions and PN codes in IS-95-based cellular CDMA systems. The concept of multiple spreading will be expanded in this article. Walsh code spreading can be used if the receiver is synchronized with the transmitter. In the forward link, the base station can transmit a pilot signal to enable the receiver to recover synchronization. In IS-95 system, a pilot signal is not sent in the reverse link. Therefore, Walsh symbol modulation is used from the mobile station to the base station. In W-CDMA the pilot is sent in both directions. So the multiple spreading technique is also used in the reverse channel.

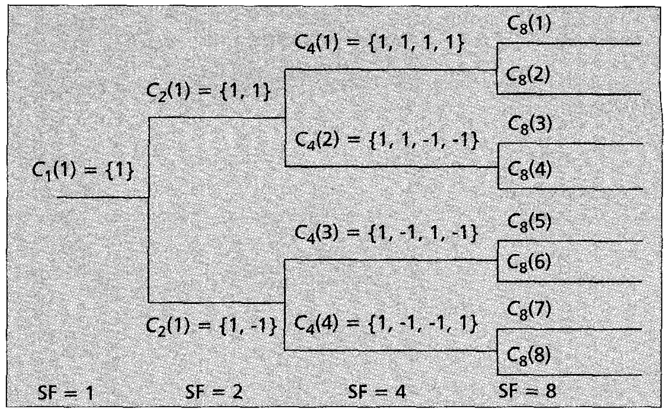


■ **Figure 7.** *Application of Walsh functions and PN codes in the forward and reverse links of cellular CDMA.*

Variable-Length Orthogonal Codes

Improving the capability of multimedia communications is one of the targets for W-CDMA mobile communications systems.

W-CDMA is designed to support a variety of data services from low to very high bit rates. Since the spreaded signal bandwidth is the same for all users, multiple-rate transmission needs multiple spreading factors (SF) in the physical channels. Consider that each bit of the lowest-bit-rate service (R_{min}) is spread by a code of length $N = 2^n$. Since the bit duration for rate $2R_{min}$ is half the duration of a bit in the minimum rate, we need a spreading code of length $N/2 = 2^{n-1}$ for spreading. Generally, a code length of 2^{n-k} is needed for bit rate $2^k R_{min}$. Depending on the maximum and minimum supported bit rates in the system and the spreading bandwidth, a range for the code length can be obtained. A method to obtain variable-length orthogonal codes that preserve orthogonality between different rates and SFs based on a modified Hadamard transformation is presented in [14].



■ **Figure 8.** *Code tree for generation of variable length orthogonal codes (SF: spreading factor).*

Let C_N be a matrix of size $N \times N$ and denote the set of N binary spreading codes of N-chip length, $\{C_N(n)\}_{n=1,\dots,N}$, where $C_N(n)$ is the row vector of N elements and $N = 2^n$; it is generated from $C_{N/2}$ as

$$C_N = \begin{bmatrix} C_N(1) \\ C_N(2) \\ C_N(3) \\ \vdots \\ C_N(N-1) \\ C_N(N) \end{bmatrix} = \begin{bmatrix} C_{N/2}(1)C_{N/2}(1) \\ C_{N/2}(1)\overline{C_{N/2}(1)} \\ C_{N/2}(2)C_{N/2}(2) \\ C_{N/2}(2)\overline{C_{N/2}(2)} \\ \vdots \\ C_{N/2}(N/2)C_{N/2}(N/2) \\ C_{N/2}(N/2)\overline{C_{N/2}(N/2)} \end{bmatrix} \quad (12)$$

As a result, these variable length orthogonal codes can be generated recursively using a tree structure as shown in Fig. 8. Starting from $C_1(1) = 1$, a set of 2^k spreading codes with the length of 2^k chips are generated at the kth layer. It can be understood from Eq. 12 that generated codes of the same layer constitute a set of Walsh functions and they are orthogonal, although the rows of C_N are not in the same order of H_N. Fortunately, any two codes of different layers are also orthogonal except for the case that one of the two codes is a mother code of the other; for example, all of $C_{16}(2)$, $C_8(1)$, $C_4(1)$, and $C_2(1)$ are mother codes of $C_{32}(3)$, and so are not orthogonal against $C_{32}(3)$. In other words, a code can be used in a channel if and only if no other code on the path from the specific code to the root of the tree or the sub-tree produced by the specific code is used in the same channel. From this observation, we can easily find that if $C_8(1)$ is assigned to a user, all the codes $\{C_{16}(1), C_{16}(2), C_{32}(1), \dots, C_{32}(4), C_{64}(1), \dots, C_{64}(8), C_{128}(1), \dots\}$ generated from this code cannot be assigned to other users requesting lower rates; in addition, mother codes $\{C_4(1), C_2(1)\}$ of $C_8(1)$ cannot be assigned to users requesting higher rates. This means that the number of available codes is not fixed but depends on the rate and spreading factor of each physical channel. These restrictions are imposed in order to maintain orthogonality.

MULTIPLE SPREADING

Multiple spreading or two-layered spreading code allocation provides flexible system deployment and operation. It is possible to provide waveform orthogonality among all users of the same cell while maintaining mutual randomness only between users of different cells. This is rendered possible by the wide bandwidth of a spread spectrum DS-CDMA system that provides considerable waveform flexibility. Orthogonality can be achieved by first multiplying each user's binary input by a short spread sequence which is orthogonal to that of every other user of the same cell. As mentioned, one class of such binary orthogonal sequences is the variable-length Walsh orthogonal set. This spread signal is followed by multiplication of a long pseudorandom sequence, which is cell-specific but common to all users of that cell in the forward link and user-specific in the reverse link. The short orthogonal codes are called *channelization codes*, the long PN sequences *scrambling codes*. Hence, each transmission *channel code* is distinguished by the combination of a channelization code and a scrambling code.

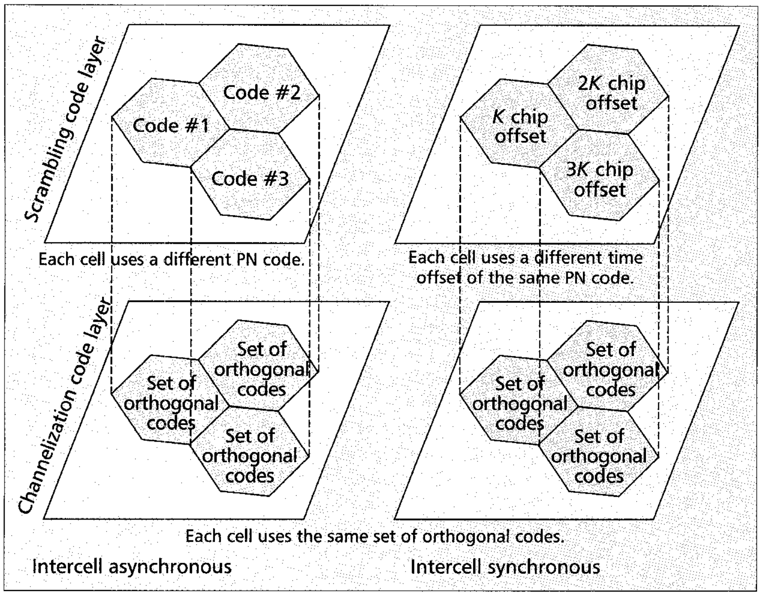
Now we present different spreading used in CDMA and W-CDMA reverse and forward links. Figure 9 shows two different channel code assignments used for forward links in CDMA and W-CDMA. Since the cells are assigned to different scrambling codes, each cell site can use short spreading codes independent from other cells. In *intercell asynchronous* operation each cell site is assigned to a distinct Gold sequence. As mentioned, a large number of long Gold spreading codes with a limited cross-correlation can be generated.

In *intercell synchronous* operation different cell base stations (different mobile users) use different time shifts of the same sequence in the forward (reverse) link. For example, in IS-95-based systems, the pilot signal is a pseudorandom binary sequence signal with a period of 32,768 chips. Since the chip rate is 1.2288 Mchip/s, the pilot pseudorandom binary sequence corresponds to a period of 26.66 ms. The pilot signals from all base stations use the same pseudorandom binary sequence, but each base station is identified by a unique time offset of its pseudorandom binary sequence. These offsets are in increments of 64 chips, providing 511 unique offsets relative to zero offset code. This large number of offsets ensures that unique base station identification can be obtained, even in dense microcellular environments. The reverse link uses the same 32,768-chip code used on the downlink. Each mobile station uses a different time shift on the PN code; therefore, the radio system can correctly decode the information from an individual mobile station.

Table 2 summarizes the spreading codes used in DS-CDMA and W-CDMA systems. The standardization of W-CDMA is underway, and different proposals are under investigation. The channelization code can be implemented using variable-length orthogonal sequences. As mentioned before, the scrambling codes can be based on the intercell synchronous or asynchronous techniques. The North American standard uses synchronous systems, while European and Japanese systems use the asynchronous technique where in the reverse link the mobile-station-unique scrambling code is allocated from the set of very large Kasami codes which are 256 chips long. Since this set of codes includes more than one million different codes, no extensive code planning is needed. Each cell is allocated a suitable subset of these

	CDMA (IS-95)		WCDMA	
	Forward link	Reverse link	Forward link	Reverse link
Channelization code	Walsh orthogonal sequences of length 64	—	Variable-length orthogonal sequences	Variable-length orthogonal sequences
Scrambling code	Different offsets of an m-sequence with a period of 32,767 (2^{15}-1) chips (a common PN for all users of a cell)	Different offset of an m-sequence with a period of 32,767 (2^{15}-1) chips (a distinct offset for each user)	a) 10 ms of a 2^{18}-1 chip Gold sequence. (a common PN for all users of a cell)	Very large set of Kasami sequences Optional: 10 ms of a 2^{41}-1 Chip Gold sequence

■ **Table 2.** *Spreading codes in CDMA and WCDMA cellular networks.*



■ **Figure 9.** *Intercell synchronous and asynchronous spreading code assignment.*

codes to be used by an active mobile station within the cell. Also, a second scrambling code using Gold codes of length 2^{41} may be used for the reverse link. They are truncated to form a cycle of 2^{16} bits (10 ms frame) and are selected based on computer simulation such that cross-correlation is minimum. Since the asynchronous technique does not need timing synchronization between base stations, it makes system deployment from outdoors to indoors very flexible; however, it makes cell search and code synchronization more complex.

CONCLUSIONS

This article reviews the important aspects of spreading codes used in the DS-CDMA technique. The spreading codes in CDMA are designed to have random behavior and very low cross-correlation. We discuss the main characteristics of m-sequences and their generation by means of the LFSR structure. A technique to select m-sequences with good cross-correlation properties (preferred pairs) is presented which leads to the generation of Gold codes. Kasami sequence sets are also important because of the large number of codes they supply and their low cross-correlation. Finally, we explain how the use of orthogonal codes and multiple spreading techniques provides flexible code allocation to the base station and mobile user.

REFERENCES

[1] A. J. Viterbi, *CDMA Principles of Spread Spectrum Communication*, Addison-Wesley, 1995.

[2] R. L. Pickholtz, D. L. Schilling, and L. B. Milstein, "Theory of Spread-Spectrum Communications — A Tutorial," *IEEE Trans. Commun.*, vol. COM-30, no. 5, May 1982, pp. 855–84.

[3] S. W. Golomb, *Shift Register Sequences*, Aegean Park Press, 1992.

[4] P. Fan and M. Darnell, *Sequence Design for Communication Applications*, U.K.:R esearch Studies Press, 1996.

[5] W. Stahnke, "Primitive Binary Polynomials," *Math. Comp.*, vol. 27, Oct. 1973, pp. 977–80.

[6] L. R. Welch, "Lower Bounds on the Maximum Cross-Correlation of Signals," *IEEE Trans. Info. Theory*, vol. IT-20, May 1974, pp. 397–99.

[7] J. H. Lindholm, "An Analysis of the Pseudo-Randomness Properties of Subsequences of Long m-Sequences," *IEEE Trans. Info. Theory*, vol. IT-14, July 1968, pp. 569–76.

[8] S. Wainberg and J.K. Wolf, "Subsequences of Pseudo-Random Sequences," *IEEE Trans. Commun. Tech.*, vol. COM-18, Oct. 1970, pp. 606–12.

[9] S. Fredricsson, "Pseudo-Randomness Properties of Binary Shift Register Sequences," *IEEE Trans. Info. Theory*, vol. IT-21, Jan. 1975, pp. 115–20.

[10] R. Gold, "Optimal Binary Sequences for Spread Spectrum Multiplexing," *IEEE Trans. Info. Theory*, vol. IT-B, Oct. 1967, pp. 619–21.

[11] R. Gold, "Maximal Recursive Sequences with 3-Valued Recursive Cross-Correlation Functions," *IEEE Trans. Info. Theory*, vol. IT-4, Jan. 1968, pp. 154–56.

[12] R. C. Dixon, *Spread Spectrum Systems with Commercial Applications*, Wiley, 1994.

[13] T. Kasami, "Weight Distribution Formula for Some Class of Cyclic Codes," *Coordinated Science Lab.*, Univ. IL, Urbana. Tech. Rep., R-285, Apr. 1966.

[14] F. Adachi, M. Sawahashi, and K. Okawa, "Tree-Structured Generation of Orthogonal Spreading Codes with Different Length for Forward Link of DS-CDMA Mobile Radio," *Elect. Lett.*, vol. 33, no. 1, Jan. 1997, pp. 27–28.

BIOGRAPHIES

ESMAEL H. DINAN received B.S. and M.S. degrees from Sharif University of Technology, Tehran, Iran, in 1993 and 1995 respectively, in electrical engineering. He also received the M.S. degree from Ecole Nationale Supérieure d'Informatique et de Mathématiques Appliquées de Grenoble (ENSIMAG) in France, in 1997 in computer science. He is currently pursuing his Ph.D. degree at the School of Information Technology and Engineering of George Mason University, Fairfax, Virginia. His area of active research is on access methods and performance modeling of wireless communication networks.

BIJAN JABBARI (bjabbari@gmu.edu) received a Ph.D. degree from Stanford University, Stanford, California, in 1981, in electrical engineering. He has held positions with Hewlett Packard, Southern Illinois University, Satellite Business Systems, and M/A-COM Telecommunications. In 1988, he joined George Mason University as an associate professor of electrical and computer engineering. He is also a professeur associé at Ecole Nationale Supérieure des Télécommunications, Paris, France. He is editor for Wireless Multiple Access for the *IEEE Transactions on Communications* and is on the editorial board of *Proceedings of the IEEE* and a few other journals. He has served as guest editor of the IEEE publications on wireless communication networks. He is on the board of several international telecommunications advisory committees and serves as a consultant to government and industry.