Code Division Multiple Access: A Tutorial

Wireless Communications
Fall, 1999
Rowan University

November 24, 1999

By Amol Shah

Table of contents:

Introduction

One of the most important concepts to any cellular telephone system is that of "multiple access", meaning that multiple, simultaneous users can be supported. In other words, a large number of users share a common pool of radio channels and any user can gain access to any channel. Different types of cellular systems employ various schemes to achieve this multiple access. The traditional analog cellular systems, such as those based on the Advanced Mobile Phone Service (AMPS) and Total Access Communications System (TACS) standards, use Frequency Division Multiple Access (FDMA). Another common multiple access method employed in new digital cellular systems is the Time Division Multiple Access (TDMA). TDMA digital standards include North American Digital Cellular (IS-54), Global System for Mobile Communications (GSM) and Personal Digital Cellular (PDC). TDMA systems commonly start with a slice of spectrum referred to as one "carrier." Each carrier is then divided into time slots. Only one subscriber at a time is assigned to each time slot, or channel. With CDMA, unique digital codes, rather than separate RF frequencies or channels, are used to differentiate subscribers. The codes are shared by both the mobile station and the base station, and are called "pseudo-Random Code Sequences." All users share the same range of radio spectrum.

CDMA is a form of *spread-spectrum*, a family of digital communications techniques that have been used in military applications for years. Originally there were two motivations for using CDMA: either to resist enemy efforts to jam the communications, or to hide the fact that communication was even taking place. The use of CDMA for civilian mobile radio applications was proposed 40 years ago, but did not take place till recently. In March 1992, the Telecommunications Industry Association (TIA) established the TR-45.5 subcommittee with the charter of developing a spread-spectrum digital cellular standard. In the July of 1993, the TIA gave its approval to the CDMA IS-95 standard [1].

In recent times, CDMA has gained widespread international acceptance by cellular radio system operators as an upgrade that will increase both their system capacity and the service quality. The rest of this tutorial aims at giving the reader a basic understanding of the concepts behind CDMA, and various issues that need to be considered in the design of a CDMA system.

Spread Spectrum

Understanding the concept of CDMA begins with the basic concept of spectrum and the process of spectrum spreading. The term *spectrum* refers to the power spectrum associated with a baseband signal and the term *spread spectrum* refers to the spreading of the power spectrum of the baseband signal over a given bandwidth [2].

First consider a simple square pulse that has the following boundary conditions:

V(t) = V -T/2 < t < T/2
V(t) = 0 elsewhere

To determine the frequency and power spectrum of this signal, we apply the Fourier Transform:

$$S(\omega) = \int_{-T/2}^{T/2} V \cdot e^{-j\omega t}\, dt$$
$$= \left(2\frac{V}{\omega}\right)\sin\left(\frac{\omega T}{2}\right)$$
$$= VT\left[\frac{\sin\left(\frac{\omega T}{2}\right)}{\frac{\omega T}{2}}\right]$$

which reveals that a square pulse signal is composed of an infinite number of harmonically related sinusoidal waves having different amplitudes.

Figure 1 shows the power spectrum of this signal along with the harmonic components. We see that most of the power is in the main lobe whose bandwidth is given by 1/T, where T is the bit duration.

Figure 1 – Square pulse and power spectral density before and after spreading [2]

Spectrum spreading can be accomplished by increasing the frequency of the discrete time signal. Thus we consider a waveform with amplitude V and frequency f, and then increase the frequency of the same waveform by a factor of n, i.e. T is now reduced by n. The conditions describing this waveform are as follows:

V(t) = V -T/(2n) < t < T/(2n)
= 0 elsewhere

Applying the Fourier transform, we get the following spectral components:

$$S(\omega) = \int_{-T/2}^{T/2} V \cdot e^{-j\omega t} dt$$
$$= VT \left[\frac{\sin(\frac{\omega T}{2n})}{\frac{\omega T}{2n}} \right]$$

Figure 1 shows the power spectrum for n = 1 and n = 2. The energy of the signal does not change as a result of spreading. But since the energy of the signal is effectively the area under the curve for the power spectral density, having a larger bandwidth means that the amplitude of the curve must go down. This process is known as spread spectrum, resulting in a process gain that is defined as:

$$G_s = 10 \log(\frac{BW}{R_b})$$

where G_s is the process gain, BW is the transmission bandwidth, and R_b is the bit rate. For example, if BW=30 kHz, R_b=10 kHz, then G_s = 10 log(30/10) = 4.77 dB. Now if we increase the bandwidth to 1.25 MHz, the process gain would be G_s = 10 log (1250000/10) = 20.97 dB. The significance of the process gain is that it effectively gives us a "noise margin." It is as if the SNR has been enhanced due to spreading, making the signal more robust to interference from other users. This is important in CDMA because all other users in the cell are essentially interference for a given user.

Types of spread spectrum

There are two common spread spectrum techniques used to transmit signals. They are direct sequence (DS) and frequency hopping (FH). In frequency hopping, the data signal is transmitted as a narrow band signal with a bandwidth only wide enough to carry the required data rate. At specific intervals, this narrow band signal is moved, or hopped, to a different frequency within the allowed band. The sequence of frequencies follow a pseudo random sequence known to both the transmitter and receiver. A method known as *fast hopping* can be used where the system makes many hops for each bit of data that is transmitted. As a result, each bit is redundantly transmitted on several different frequencies. This allows interference to exist in the band which would under normal circumstances, block one or more narrowband channels. The downside of this approach is that the implementation of a fast hopper is complex and expensive [3].

As a result, another method known as direct sequence spreading grew out of the need for a cheaper system. In this method, the signal is multiplied by a pseudo random code sequence having a much faster bit rate. As a result, the bandwidth of the data signal gets spread. On the receiver side, this signal can then be multiplied with the same pseudo

noise signal and the original data signal is recovered. We take a detailed look at pseudo noise sequences and direct spread spectrum in the sections that follow.

PN Sequences

What are PN sequences?

A Pseudo-random Noise (PN) sequence is a sequence of binary numbers, e.g. ±1, which appears to be random; but is in fact perfectly deterministic. The sequence appears to be random in the sense that the binary values and groups or runs of the same binary value occur in the sequence in the same proportion they would if the sequence were being generated based on a fair "coin tossing" experiment. In the experiment, each head could result in one binary value and a tail the other value. The PN sequence appears to have been generated from such an experiment. A software or hardware device designed to produce a PN sequence is called a PN generator [4].

Pseudo-random noise sequences or PN sequences are known sequences that exhibit the properties or characteristics of random sequences. They can be used to logically isolate users on the same frequency channel. They can also be used to perform scrambling as well as spreading and despreading functions.

The reason we need to use PN sequences is that if the code sequences were deterministic, then everybody could access the channel. If the code sequences were truly random on the other hand, then nobody, including the intended receiver, would be able to access the channel. Thus, using a pseudo-random sequence makes the signal look like random noise to everybody except to the transmitter and the intended receiver [4].

PN Sequence Generation

A PN generator is typically made of N cascaded flip-flop circuits and a specially selected feedback arrangement as shown in figure 2. The flip-flop circuits when used in this way are called a shift register since each clock pulse applied to the flip-flops causes the contents of each flip-flop to be shifted to the right. The feedback connections provide the input to the left-most flip-flop. With N binary stages, the largest number of different patterns the shift register can have is 2^N. The all-binary-zero state, however, is not allowed because it would cause all remaining states of the shift register and its outputs to be binary zero. The all-binary-ones state does not cause a similar problem of repeated binary ones provided the number of flip-flops input to the module 2 adder is even. The period of the PN sequence is therefore 2^N-1.

For example, starting with the register in state 001 as shown, the next 7 states are 100, 010, 101, 110, 111, 011, and then 001 again and the states continue to repeat. The output

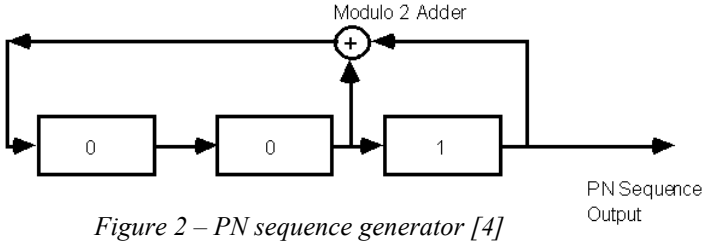

Figure 2 – PN sequence generator [4]

taken from the right-most flip-flop is 1001011 and then repeats. With the three-stage shift register shown, the period is 2^3-1 or 7.

The flip-flops that should be tapped-off and fed into the module 2 adder are determined by an advanced algebra that has identified certain binary polynomials called primitive irreducible or unfactorable polynomials. Such polynomials are used to specify the feedback taps. For example, IS-95 specifies the in-phase PN generator shall be built based on the characteristic polynomial [4]:

$$P(x) = x^{15} + x^{13} + x^9 + x^8 + x^7 + x^5 + 1$$

Picture a 15-stage shift register with the right-most stage numbered zero and the successive stages to the left numbered 1, 2, 3, ..., 14. Then the exponents less than 15 in the equation above tell us that stages 0, 5, 7, 8, 9, and 13 should be tapped and summed in a module 2 adder. The output of the adder is then input to the left-most stage. The shift register PN sequence generator is shown below.

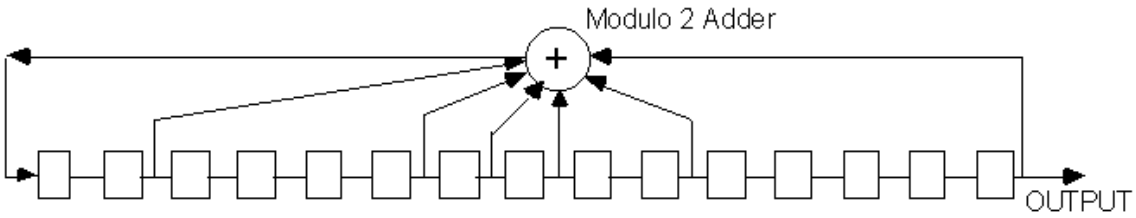

Figure 3 – 15 stage shift register to generate PN sequences [4]

Properties of PN sequences

PN signals are deterministic signals. They can be shown to have the following three properties:

a) *Balance property*: The numbers of zeros and ones of a PN code are different only by one.
b) *Run Property*: The number of times the ones and zeros repeat in groups or runs appear in the same proportion they would if the sequence were actually generated by a coin tossing experiment. For an N bit code, there would be N zero (or one) runs. 1/2 of these runs would be of length one, 1/4 of length 2, 1/8 of length 3 and so on.
c) *Correlation Property*: The correlation value of two N-bit sequences can be obtained by counting the number of similar (N_s) and dissimilar (N_d) bits and inserting them into the following equation:

$$P = (1/N) * (N_s - N_d)$$

Types of PN sequences in CDMA

IS-95 uses two PN generators to spread the signal power uniformly over the physical bandwidth of about 1.25 MHz. The PN spreading on the reverse link also provides near-orthogonality of and hence, minimal interference between signals from each mobile. This allows reuse of the band of frequencies available, which is a major advantage of CDMA. In IS-95, two types of maximum-length PN Sequences or PN codes are used: the short PN code and the long PN code.

Short Code: The short PN code is generated by a 15-stage linear shift register. Therefore, the maximum length of the Short PN Code is

$L = 2^N-1 = 2^{15}-1 = 32{,}768-1$

By implementation, an extra chip is inserted at the end of the sequence, yielding a sequence of length L=32,768 chips. The short PN code runs at a speed of 1,228,800 chips per second. This yields a repetition cycle of 32,768/1,228,800=26.67 ms

Long Code: The PN chips from the long code are used to provide several randomizing functions in the IS-95 system. These include providing chips for message-scrambling on the forward and reverse links, for identifying individual mobiles and access channels on the reverse links by using unique offsets for each entity and for randomizing the location of the power control bits on the forward traffic channels.

The long PN code is generated by a 42-stage linear shift register. Therefore, the maximum length of the long PN code is $L = 2^N-1 = 2^{42}-1 = 4.4 \times 10^{12} = 4.4$ trillion chips. The Long PN Code also runs at a speed of 1,228,800 chips per second. This yields a repetition cycle of $4.4 \times 10^{12}/1{,}228{,}800 = 41$-42 days.

The long PN code is generated in a 42 stage linear shift register generator with the output of the 42nd stage input into the first stage and modulo-2 added with the outputs of stages 1, 2, 3, 5, 6, 7, 10, 16, 17, 18, 19, 21, 22, 25, 26, 27, 31, 33, and 35. The output of the long code generator is taken after the output of each flip-flop in the generator has been added with a corresponding bit in a 42-bit mask which is unique to each user, access, and paging channel [4].

Correlation between PN sequences

The correlation of two random variables x(t) and y(t), is a *time-shift* comparison which expresses the degree of similarity or the degree of likeness between the two variables.

The Auto-Correlation function R, provides the degree of similarity between a random variable x(t) and a time-shifted version of x(t). Likewise, the cross-correlation function provides the degree of similarity, or the degree

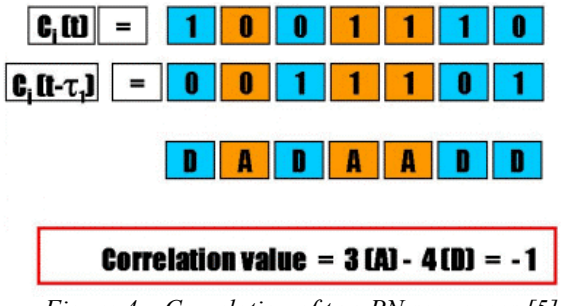

Figure 4 – Correlation of two PN sequences [5]

of likeness between a random variable x(t) and time-shifted version of another random variable y(t). To get the average value of the auto-correlation or cross-correlation, a normalization by the sequence length L is required.

Consider $C_i(t)$ and the time-shifted version of itself, say $C_i(t-1)$

Ci(t) = 1 0 0 1 1 1 0
Ci(t-1) = 0 0 1 1 1 0 1

When corresponding bits from the two sequences have the same parity (or match each other), we call the match an agreement "A". Likewise, when corresponding bits from the two sequences do not have the same Parity (do not match each other), we call the mismatch a disagreement "D" (see figure 4). By counting all the agreements and all the disagreements over the full length L of the sequence, a measure of correlation can be estimated as [4]:

Correlation = Total number of "A" - Total number of "D"

Now, consider the reference PN code $C_i(t)$ and its time-shifted versions as shown below. Now let us compute the correlation of $C_i(t)$ and $C_i(t-\tau)$, for all suitable values of τ (here from 0 to 7).

Time Shifts	Shifted Sequence	Correlation
Reference	1 0 0 1 1 1 0	
τ_0	1 0 0 1 1 1 0	+7
τ_1	0 0 1 1 1 0 1	-1
τ_2	0 1 1 1 0 1 0	-1
τ_3	1 1 1 0 1 0 0	-1
τ_4	1 1 0 1 0 0 1	-1
τ_5	1 0 1 0 0 1 1	-1
τ_6	0 1 0 0 1 1 1	-1
τ_7	1 0 0 1 1 1 0	+7

In general, it can be shown that the full-length auto-correlation function (R) of PN codes or PN sequences is characterized by a large positive number equal to the length of the PN sequence ($R=2^n-1$) when time shift=0, and -1 for all time-shifts equal or greater than the duration of one chip. So when normalized by the length, the auto-correlation function is equal to 1 at time-shift zero and is very small (-1/L) for all values of time shifts equal or greater than one chip.

In summary, the auto-correlation function of PN codes is a two-value function. Its maximum value occurs when the time-shift parameter is zero. For all other values equal to or greater than one chip, the correlation function is -1.

Orthogonality of PN sequences

Consider the reference PN Code $C_j(t)$ and the time-shifted versions of another code $C_i(t)$ as shown below. Let us compute the cross-correlation of $C_j(t)$ and $C_i(t-\tau)$ for all suitable values of τ (0 to 7).

Time Shifts	Shifted Sequences	Correlation
Reference	1 0 0 1 0 1 1	
τ_0	1 0 0 1 1 1 0	+3
τ_1	0 0 1 1 1 0 1	-1
τ_2	0 1 1 1 0 1 0	-1
τ_3	1 1 1 0 1 0 0	-5
τ_4	1 1 0 1 0 0 1	+3
τ_5	1 0 1 0 0 1 1	+3
τ_6	0 1 0 0 1 1 1	-1
τ_7	1 0 0 1 1 1 0	+3

Two PN sequences $C_i(t)$ and $C_j(t)$ are said to be orthogonal if and only if their respective normalized correlation function is equal to 1 at a time-shift of zero and their cross-correlation function is equal to zero for all time-shift values. As shown above, averaged over the code length, the cross-correlation function of PN sequences is not zero. As a result, PN sequences are not perfectly orthogonal.

Code Usage

Short PN Codes

Shifted versions of the same PN sequence exhibit very small correlation properties, (-1/L). So, we can use the various offsets of the same sequence as different codes. In principle, all 32,768 offsets of the Short PN Code can be respectively treated as different codes since they are mutually isolated by a cross-correlation factor of -1/L, while maintaining an auto-correlation factor of 1. However, if two adjacent offsets are used, a multipath of the leading sequence (delayed by exactly one chip) would look identical to the lagging sequence. As a result, in IS-95, a 64-chip separation is recommended between short PN offsets. A 64-chip offset separation on the Short PN Code yields 512 offsets that will be used to isolate cells and sectors.

Let us assign the short PN offsets $PN(t- t_A)$ and $PN(t- t_B)$ to cell A and cell B respectively. The signal from cell A to mobile A is of the form $di(t)*PN(t- t_A)$. Likewise the signal from cell B to mobile B is of the form $dj(t)*PN(t- t_B)$.

The received signal at mobile A is the composite signal from cell A and cell B. Likewise, the received signal at mobile B is the same composite signal from cell A and cell B. The composite signal received by each Mobile can be represented as:

$R_x = di(t)*PN(t- t_A) + dj(t)*PN(t- t_B)$

Here, we have ignored the propagation delay that can be resolved by the synchronization process at the mobiles. Mobile A is assumed to be synchronized with cell A. So, it is tuned to cell A with its local short PN code set to $PN(t - t_A)$. Mobile A will then "multiply" its local code $PN(t - t_A)$, with the received or incoming input signal. The resulting output signal is:

$Output = di(t)*PN(t - t_A)*PN(t - t_A) + dj(t)*PN(t - t_B)*PN(t - t_A)$

As noted earlier, the normalized auto-correlation $PN(t - t_A)*PN(t - t_A)$ is 1, and the normalized cross-correlation $PN(t - t_B)*PN(t - t_A)$ is -1/L. For large values of L therefore, the second term can be ignored. This simply reduces the output to di(t). In this way, the signal from cell B is ignored.

On the forward link, short PN offsets are assigned to cells and sectors (one PN offset per omnicell or per sector) on a dedicated basis. Used as such, they provide isolation between cells or sectors due to their relatively small cross-correlation factor.

As these codes exhibit randomness properties, short PN codes are also used to provide quadrature spreading on the Forward Link. So, two Short PN Sequences are actually used - an in-phase sequence $PN_i(t - t_A)$ and a quadrature sequence $PN_q(t - t_A)$. These two sequences have the same length, same speed of 1.2288 million chips per second, the same phase (t_A) for a given cell or sector, but they have different generating-polynomials (different Shift Register Taps).

In the Reverse Link, the zero-offset of the Short PN Code pair $PN_i(t - 0)$ and $PN_q(t - 0)$ is used to provide quadrature PN spreading at the mobile.

Long PN Codes

As with the short PN codes, we can use the various offsets of the same long PN sequence as different codes. In principle, all 4.4 trillion offsets of the Long PN Code can be treated as distinct codes since they are mutually isolated by a cross-correlation factor of -1/L, while maintaining a low correlation factor. However, if two adjacent offsets are used, a multipath of the leading sequence (delayed by exactly one chip) would look identical to the lagging sequence. As a result, in IS-95, a multiple of a 64 chip separation (typically 256 chips) is recommended between adjacent long PN offsets. A 64-chip separation between adjacent offsets of the long PN code would yield approximately 69 billion offsets. A 256-chip separation would yield about $2^{42}/2^8 = 17$ billion offsets.

A subset of these codes is used as access channels. They allow mobiles to register, originate a call, perform authentication and so on. These channels are accessed on a contention basis. Another subset is used to isolate mobiles in the reverse link. These offsets are assigned to mobiles on a dedicated basis (like short PN offsets are assigned to the cells and sectors on a dedicated basis.)

Direct Sequence Spread Spectrum

Direct sequence is a spread spectrum technique in which the bandwidth of a signal is increased by artificially increasing the bit data rate. This is done by breaking each bit into a number of sub-bits called "chips." For example, if this number is 10, each bit in the original signal would be divided into 10 separate bits, or chips. This results in an increase in the data rate by 10. By increasing the data rate by 10, we also increase the bandwidth by 10.

The signal is divided into smaller bits by means of a PN sequence. This can be accomplished by using a two-input exclusive OR gate, where one input is the low speed data and the other input is a high speed PN-sequence. As shown in figure 5, the data signal has a narrow power spectrum. The high-speed PN code, in comparison, has a wider power spectrum. The result is that the composite signal has the same transition rate as the PN sequence, because of its wideband power spectrum, but lower amplitude because the total energy is constant [2].

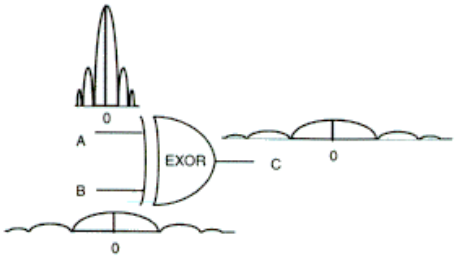

Figure 5 – Spreading a narrow band signal with a wideband PN sequence [2]

To now recover data from the composite spread-spectrum signal, another exclusive OR gate is used where the composite data C is applied to one input and an identical PN sequence is applied to the second input. The output Y is the original signal. It is important to note that the recovery process only works when the PN sequence is identical for both spreading and despreading; otherwise the desired signal will never be recovered.

One of the main advantages of the spread-spectrum signal is its tolerance to interference. As shown in figure 6, consider a narrowband signal A and a PN sequence B that are combined to result in a spread-spectrum signal C. Noise n is then added to this signal, and the resulting signal is then despread by XORing with the same PN sequence at the receiver end. As the following derivation shows, the received signal Y is the original signal A with very low amplitude noise spread across the entire bandwidth of the spread signal. As a result, signal A can be recovered by bandpass filtering this received signal.

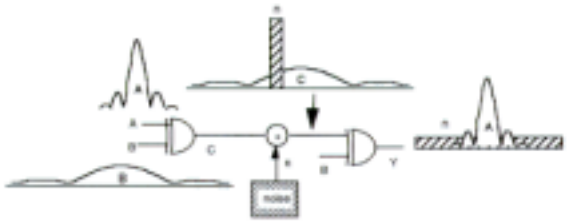

Figure 6 – Resistance of CDMA to noise/jamming [2]

$$C = A\overline{B} + \overline{A}B$$

$$Y = (C\overline{B} + \overline{C}B) + (n\overline{B} + \overline{n}B)$$

$$= (A\overline{B} + \overline{A}B)\overline{B} + \overline{(A\overline{B} + \overline{A}B)}B + (n\overline{B} + \overline{n}B)$$

$$= A + (n\overline{B} + \overline{n}B)$$

where A is the desired signal recovered due to despreading, and the second term is the interference component, spread over the entire band of the high speed PN code. Since the energy is conserved, the magnitude of the interference power is reduced proportionately, referred to as process loss.
Therefore, all users, other than the desired signal appear as noise. Figure 7 implies that there is a process loss for the interferer due to spreading and there is a process gain due to despreading of the desired signal. For *k* users this loss can be defined as

Process loss = 10 log(*k*)

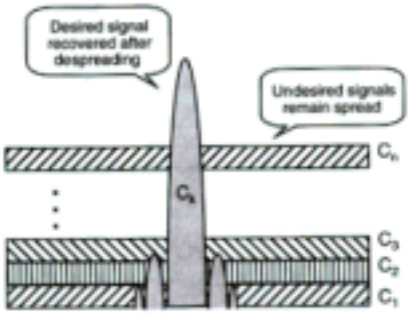

Figure 7 – Process gain for a CDMA system [2]

The overall system gain can then be defined as

CDMA gain = Process gain – Process loss due to k users
= 10 log (BW/R_b) – 10 log (k)
= 10 log (BW/kR_b)

Synchronization

The process of bringing two signals or the wideband components of two signals into time alignment is known as synchronization. Whenever a symbol stream is combined with a PN-chip sequence at the transmitter for scrambling, spreading, or addressing; the combining process must eventually be reversed or undone at the receiver to recover the symbol stream. In each case, the signal s(t) is in effect multiplied by a wideband signal PN(t) and the transmitted signal is of the form PN(t)*s(t). After propagation delay, the received signal is of the form PN(t-τ)*s(t-τ). At the receiver, a local replica of the PN sequence produces a signal PN(t-x), where x is an adjustable timing offset. The receiver forms the product PN(t-x)*PN(t-τ)*s(t-τ). The synchronization circuit adjusts x so that it is close enough to t that PN(t-x)*PN(t-τ) is approximately one. This recovers s(t-τ), the desired signal.

Several schemes have been developed to synchronize two digital streams. In one method used in Direct Sequence (DS) spread spectrum systems, initially the PN generator at the receiver, while the same generator as used at the transmitter, is clocked at a slightly different rate causing the chip rate of the receiver sequence to be slightly different that the chip rate of the signal received. Hence, the PN signals PN(t-x) and PN(t-τ) tend to slide by one another as time passes because their rates differ. A circuit continually

monitors the product PN(t-x)*PN(t-τ)*s(t-τ). The power in this signal will be spread over a wide bandwidth whenever x and τ differ by more than about half a chip time. When x and τ differ by less that half a chip, the power in PN(t-x)*PN(t-τ)*s(t-τ) lies mostly within the bandwidth of s(t). The synchronization circuit detects this condition and sets the chipping rate at the receiver equal to the chipping rate at the transmitter. Other tracking circuitry brings the two signals into closer alignment and keeps them aligned. When the received signal is very weak, the product PN(t-x)*PN(t-τ)*s(t-τ) may be processed an extended period of time [4].

Power Control

CDMA is a multiple-access system in which several users have access to the same frequency band. As a result, the received signal strength will be different for different mobiles, resulting in what is called *near-far interference*. "Near-far" refers to the ratio of signal strength from a near mobile to the signal strength from a mobile that is far away. This is critical for CDMA because many mobiles share the same frequency. Near-far interference degrades performance, reduces capacity and causes dropped calls.

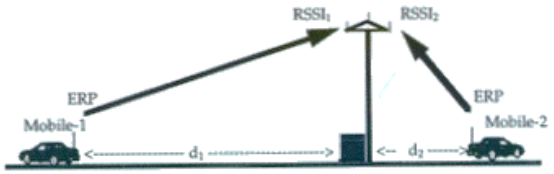

Figure 8 – Near far problem related to CDMA systems [2]

Refer to figure 8 where we have two mobiles, the first at a distance d_1 from the base, and the second at a distance d_2 from the base. According to [2], the ratio of the received signals at the base station will be

$$\frac{RSSI_1}{RSSI_2} = \left(\frac{d_2}{d_1}\right)^{\gamma}$$

where

$RSSI_1$ = received signal from mobile 1
$RSSI_2$ = received signal from mobile 2
d_1 = distance between mobile 1 and base station
d_2 = distance between mobile 2 and base station
γ = path-loss slope

This means that if d_1 and d_2 are not the same, the received signal will be different depending on the propagation environment and the respective distances. For example, if $d_2 = 4d_1$ and $\gamma = 4$ (typical dense urban environment), the received signal from mobile 1 will by 256 times (24 dB) stronger than the received signal from mobile 2, and the base station receiver will be unable to recover signal 2. Therefore, the transmitting power of each mobile has to be controlled so that its received power at the cell site is constant to a predetermined level, irrespective of distance. According to IS-95, CDMA power control is a three-step process [2]:

1. Reverse link open-loop power control (coarse)
2. Reverse link closed-loop power control (fine)
3. Forward link power control

Reverse Link Open-Loop Power Control: Reverse link (mobile to base) open-loop power control is accomplished by adjusting the mobile transmit power so that the received signal at the base station is constant irrespective of the mobile distance. The dynamic range of this power control is 85 dB.

Reverse Link Closed-Loop Power Control: Reverse link closed-loop power control is accomplished by means of a power-up or power-down command originating from the cell site. A single power control bit (1 for power down by 0.5 dB and 0 for power up by 0.5 dB) is inserted into the forward encoded data stream, every 1.25 ms. Upon receiving this command from the base station, the mobile responds by adjusting its power as instructed. The dynamic range of this power control is +/- 24 dB.

Forward Link Power Control: Forward link (base to mobile) power control is a one-step process. The base station controls its transmitting power so that a given mobile receives extra power to overcome fading, interference, BER, etc. In this mechanism, the cell site reduces its transmitting power while the mobile computes the frame error rate (FER). Once the mobile detects 1% FER, it sends a request to stop the power reduction. This adjustment process occurs once every 15 to 20 ms. The dynamic range of this type of power control is limited to only 6 dB in a 0.5 dB step because all mobiles are affected during this process.

Handoff

In AMPS, hand-off is a process of changing the carrier frequency. The primary task is to assign a new channel when a mobile moves into an adjacent cell or sector. For AMPS, the voice is muted for approximately 200 ms during this process.

In CDMA, handoffs are handled slightly differently. There are three possible types of handoffs in CDMA which are described as follows [2]:

CDMA to CDMA Soft Hand-Off: Soft hand-off is a process in which a mobile is directed to hand off to the same frequency, assigned to an adjacent cell or an adjacent sector without dropping the original RF link. The mobile keeps two RF links during the soft hand-off process. Once the new communication link is well established, the original link is dropped. This process is also known as “make before break,” which guarantees no loss of voice during hand-off.

CDMA to CDMA Hard Hand-Off: CDMA to CDMA hard hand-off is the process in which a mobile is directed to hand off to a different frequency assigned to an adjacent cell or a sector. The mobile drops the original link before establishing the new link. The voice is muted momentarily during this process.

CDMA to AMPS Hard Hand-Off: CDMA to AMPS hand-off is a process where a dual-mode mobile is directed to hand off to an AMPS channel. Voice is also muted momentarily during this process.

Capacity of a CDMA system

Single cell capacity for a CDMA system

While in FDMA and TDMA systems the capacity is bandwidth limited, the capacity of a CDMA system is interference limited, i.e. the performance for each user increases as the number of users decrease and vice versa. To calculate the capacity of the system, we first consider a single cell system. The energy-to-bit ratio is given as

$$\frac{E_b}{N_o} = \frac{\text{Energy per bit}}{\text{Power spectral density of noise + interference}}$$

Now, the energy per bit is related to signal power and data rate

$$E_b = \frac{P_s}{R}$$

and the noise is calculated as the interference from the other users divided by the bandwidth. Given this, the energy-to-bit ratio can be rewritten as [1]:

$$\frac{E_b}{N_o} = \frac{\frac{P_j}{R}}{W^{-1} \sum_i P_i}$$

where P_j is the received power from the intended user, P_i is the received power for the i^{th} user, W is the bandwidth of the signal and R is the bit rate.

The key to high capacity of commercial CDMA is realized if rather than using constant power, the transmitters can be controlled in such a way that rather than using constant power, the received powers from all users are roughly equal. The noise plus interference is now

$$N_o = (N-1)P_s$$

where N is the number of users and P_s is the power received from a user. Then the energy-to-bit ratio becomes

$$\frac{E_b}{N_o} = \frac{P_s / R}{(N-1)P_s / W} = \frac{W/R}{N-1}$$

where I is thermal interference. This equation ignores background noise, due to spurious interference as well as thermal noise. Including this term in the equation makes the energy-to-bit ratio

$$\frac{E_b}{N_o} = \frac{W/R}{(N-1)+(I/P_S)}$$

Maximum capacity is then achieved if the power control is adjusted so that the energy-to-bit ratio is exactly what is needed for an acceptable error rate. Solving this equation for the number of users N, we get

$$N = 1 + \frac{W/R}{E_b/N_o} - \frac{I}{P_S}$$

To calculate the pole capacity (maximum possible capacity) we assume that P_s goes to infinity. As a result, the capacity equation can be approximated as

$$N \approx \frac{W/R}{E_b/N_o}$$

Using the numbers for IS-95A CDMA with the 9600 kb/s transmission rate set and an energy-to-bit ratio of 6 dB (based on field tests), we find

$$N \approx \left(\frac{W}{R}\right)_{dB} - \left(\frac{E_b}{N_o}\right)_{target-dB} \approx 21.1 - 6\,\text{dB} = 15.1\,\text{dB}$$

or about N=32 [1]. This example shows that once power control is available, the system designer has the freedom to trade quality of service for capacity by adjusting the target energy-to-bit ratio.

Increasing CDMA capacity

As the interference is averaged in CDMA systems, anything that can be done to reduce the interference from other users helps increase the capacity. Human speech has been shown to have an activity factor of 35-40%. CDMA takes advantage of this and monitors the voice activity such that the transmission rate is lowered during periods of no voice activity. If we denote the voice activity factor by α, then the capacity equation changes to

$$N = \frac{1}{\alpha}\left[1 + \frac{W/R}{E_b/N_o} - \frac{I}{S}\right]$$

In general, α is taken to be around 2. This leads to a doubling of the capacity for the system [6].

Another method used to increase capacity is sectorization. The CDMA capacity is largely unaffected by sectorization, i.e. the capacity equation can be applied to each sector. As a result, with only small modifications because of the interference leakage between sectors, this increases the capacity of the system by a factor of 3. Thus, using sectorization and voice coding, we have been able to gain a six-fold increase in the capacity of the system.

Conclusion

CDMA has been gaining wide acceptance in the last few years. Many companies are now turning towards CDMA as a digital alternative to the existing analog systems. CDMA technology offers numerous benefits to the cellular operators and their subscribers. Some of the benefits include capacity increases of 5 to 6 times that of an AMPS analog system, improved call quality, simplified system planning through the use of the same frequency in every sector of every cell, enhanced privacy and increased talk time for portables through power control. Thus, compared to other systems, CDMA is an attractive alternative for cellular providers. It is important to remember however, that since the quality of a CDMA system is proportional to the number of users, careful planning of the system before implementation is essential.

References

[1] CDMA Development Group. (1999). CDMA Technology. <http://www.cdg.org/frame_tech.html>.

[2] Faruque, Saleh. (1996). Cellular Mobile Systems Engineering. Norwood, MA: Artech House.

[3] Metric Systems Corporation. (1999). Spread Spectrum. <http://www.metricsystems.com/technology.html>

[4] CDMA Online. (1999). CMDA Online. <http://www.cdmaonline.com>

[5] Lee, W. (1991, May). Overview of Cellular CDMA. IEEE Transactions on Vehicular Technology, 40, 291-302.

[6] Rappaport, Theodore. (1999). Wireless communications: Principles & Practise. Upper Saddle River, NJ: Prentice Hall.

[7] Hale, R., Shupita, K., Mahoney, R., Xia, M. (1999). CDMA: Code Division Multiple Access. <http://www.ee.mtu.edu/courses/ee465/groupe/>