Advanced Techniques for Modulation, Error Correction, Channel Equalization, and Diversity

Nambi Seshadri
Carl-Erik W. Sundberg
Vijitha Weerackody

Signal processing techniques play an important role in creating and maintaining reliable wireless digital communications. Many of these, such as low-bit-rate speech coding, convolutional channel coding with Viterbi decoding, and equalization, are already part of the first generation of digital cellular systems. Other techniques that can enhance the voice quality and system capacity include transmit and receive antenna diversity techniques, coded modulation, interference randomization and reduction, and advanced channel-decoding techniques.

Introduction

First-generation digital cellular radio systems are being standardized in various parts of the world.[1] These schemes increase the system capacity for speech transmission by about a factor of three compared with existing analog systems, and promise to become even more efficient in the future. Standards have already been or are being set for data services as well. In this paper, we review the technologies that have made efficient digital transmission possible in a wireless cellular environment, and describe several future technologies that may improve the quality and capacity of digital wireless services.

Efficient transmission of speech is made possible primarily by low-bit-rate speech coding techniques. It is now possible to achieve almost toll-quality speech at rates of about 6 to 12 kilobits per second (kb/s), and high-quality coders that operate at about 3 to 6 kb/s will soon be available. This type of speech coder can double the capacity of a digital cellular system. Here, we concentrate on communications schemes that transport digitized speech bits reliably over a cellular radio channel. Such a digital communications scheme used with a speech coder consists of a channel encoder/decoder to detect and correct errors, an interleaver/de-interleaver to ensure that errors occurring in bursts are spread so they can be corrected efficiently, a modulator to convert bits to a waveform of specified bandwidth for transmission over the air, and a demodulator to convert the waveform back to bits. Figure 1 shows this type of digital communications scheme and a frozen moment, or *snapshot*, of the fading channel signal amplitude, described in more detail later in this paper. Imperative to this communications process is a good understanding of the channel and impairments that affect the performance of digital transmission schemes.

In this paper, we describe the transmission medium of wireless communications and consider several digital modulation schemes. Two have been standardized, one for North America, by Interim Standard 54 (IS-54), and one for Europe, by the Group Speciale Mobile (GSM). We omit consideration of the Japanese digital cellular (JDC) standard, because of its similarity to IS-54. (See Panel 1 for definitions of abbreviations, acronyms, and terms used in this paper.)

The topics covered in this paper include:

- The channel model, and how it is affected by multipath fading and co-channel interference
- Digital modulation schemes
- Diversity principles
- Combining techniques
- Multiple-access techniques
- Duplexing techniques
- Channel equalization
- Channel coding
- Digital wireless system standards
- Future techniques, namely, minimum-mean-squared-error diversity combining, diversity using multiple transmit antennas, coded modulation, and advanced decoding techniques.

Channel Model

In wireless multiple-access radio systems, the two dominant sources of impairment are multipath fading and co-channel interference. ***Multipath fading***, induced by multiple scatterers, is a result of the transmitted signal being received along different paths, each with a random gain and phase.[2] (See Panel 2 for a mathematical explanation of multipath fading.)

A ***deep fade*** occurs when these paths arrive at the receiver out of phase. These fades are separated by about a half-wavelength, which, at 900 MHz, is about 0.165 meter. A vehicle moving at 80 kilometers per hour can expect a fade every 7.5 milliseconds (ms). This type of fading is ***time-selective***, since the fade varies with time. Figure 1 shows a snapshot of a time-selective multipath fade for the channel parameters given above. The nominal expected signal level is 0 decibels (dB). For a significant fraction of the time, the signal level is below 0 dB; during deep fades, it may fall below –20 dB. Fading can also be ***frequency-selective***, in which a portion of the spectrum occupied by the signal is subject to a deep fade.

Limited radio spectrum is an inherent limitation in wireless transmission. Because radio spectrum must be shared and spatially reused by many subscribers, it is inevitable that two subscribers will interfere with each other by accessing the same frequency band for information transmission, causing ***co-channel interference***.

Digital Modulation Schemes

The digital modulation method chosen for the U.S. digital cellular system (IS-54) is a modified version of a differential quadrature four-phase shift keying (DQPSK) scheme with differentially coherent detection, known as $\pi/4$ shifted DQPSK or $\pi/4$-DQPSK ($\pi/4$-4DPSK).[3] The modification consists of rotating every second symbol by $\pi/4$ radians. During odd time intervals, the phase values $0, \pi/2, \pi$, and $3\pi/2$ are used to convey information; during even time intervals, the phase values $\pi/4, 3\pi/4, 5\pi/4$, and $7\pi/4$ are used. This modification reduces the amplitude, or ***envelope***, variations of the modulated signal, which increases the efficiency of the power amplifier. The modulation scheme, shown in Table I, uses the Gray-coded phase constellation, whose information is differentially encoded. The symbols are transmitted as changes in phase rather than as absolute phases. The binary data stream entering the modulator is converted into two parallel binary streams, (X_k) and (Y_k), where X_k and Y_k are the even and odd numbered bits,

Panel 1. Abbreviations, Acronyms, and Terms

AMPS — analog mobile phone system
CDMA — code-division multiple access
CRC — cyclic redundancy check
CT2 — cordless telephone 2
DCA — dynamic channel allocation
DECT — digital European cordless telephone
DPCM — differential pulse code modulation
DQPSK — differential four-phase shift keying
DS — direct sequence
FDD — frequency-division duplexing
FDMA — frequency-division multiple access
FH — frequency hopping
GMSK — Gaussian minimum-shift-keying
GSM — Group Speciale Mobile
GVA — generalization of the Viterbi algorithm
IS-54 — Interim Standard 54
JDC — Japanese digital cellular standard
LEO — low earth orbit
LVA — list Viterbi algorithm
MMSE — minimum-mean-squared error
PBX — private branch exchange
PCN — personal communications network
PCS — personal communications system or services
RMS — root mean square
SAD — speech activity detection
SOVA — soft output Viterbi algorithm
TDD — time-division duplexing
TDMA — time-division multiple access
UEP — unequal error protection

respectively. The digital data sequences (X_k) and (Y_k) are encoded onto two signals according to

$$I_k = I_{k-1}\cos[\Delta\Phi_k(X_k,Y_k)] - Q_{k-1}\sin[\Delta\Phi_k(X_k,Y_k)]$$
$$Q_k = I_{k-1}\sin[\Delta\Phi_k(X_k,Y_k)] + Q_{k-1}\cos[\Delta\Phi_k(X_k,Y_k)],$$

where I_{k-1}, Q_{k-1} are the amplitudes at the previous pulse time. The phase change $\Delta\Phi_k$ is determined according to Table I. The signals I_k, Q_k at the output of the differential phase-encoding block can take the values $0, \pm1, \pm1/\sqrt{2}$, which produce the 8-point constellation

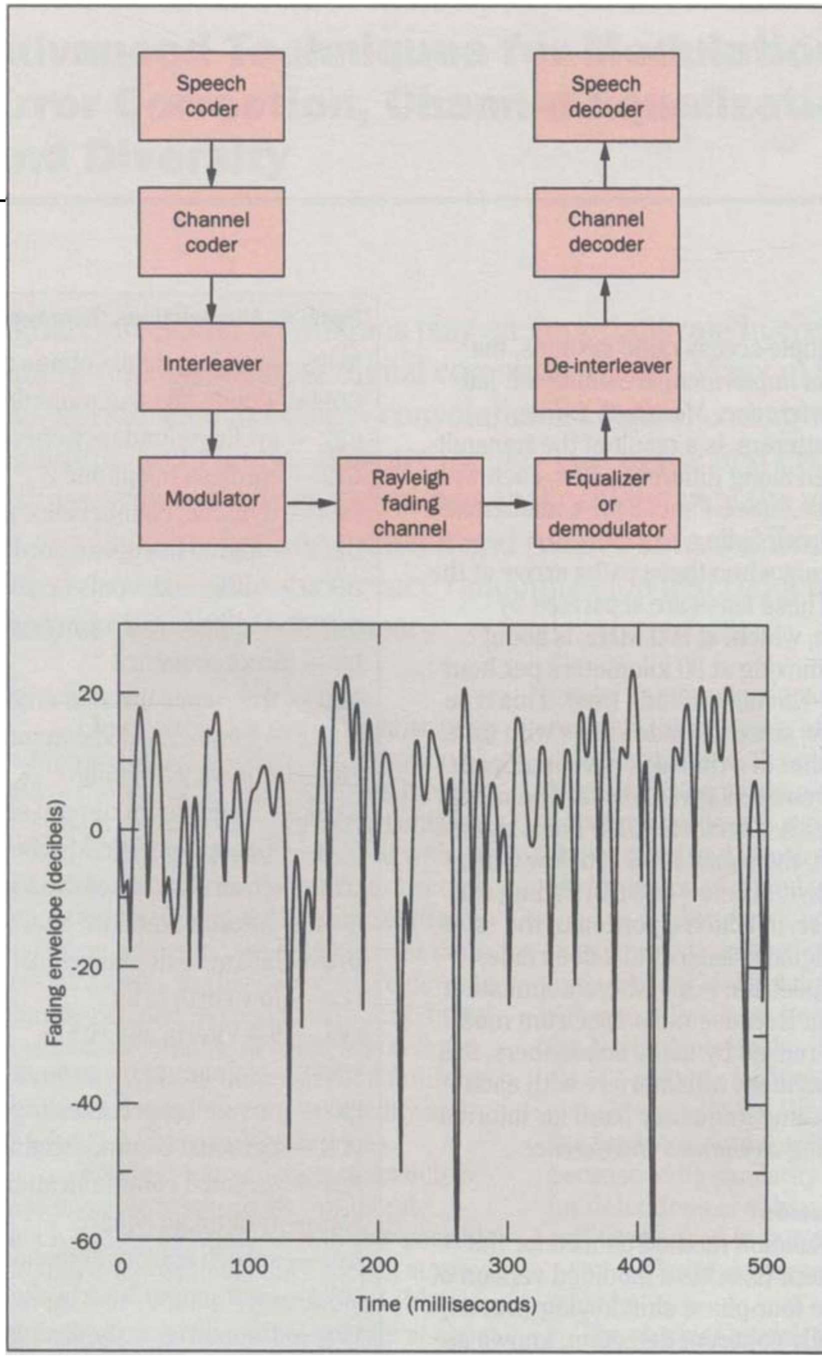

Figure 1. Block diagram of a digital speech communications system. A snapshot of the Rayleigh[2] fading amplitude with a carrier frequency of 900 MHz and a vehicle speed of 50 mph.

referred to earlier. Impulses I_k, Q_k are applied to the inputs of the I and Q baseband filters. The baseband filter, $g(t)$, has a linear phase and square-root-raised cosine-frequency response with an excess bandwidth of 0.35. The resultant transmitted signal $s(t)$ is given by

$$s(t) = \sum_k g(t-kT) I_k \cos\omega_c t - \sum_k g(t-kT) Q_k \sin\omega_c t \qquad (1)$$

where $g(t)$ is the pulse-shaping function, ω_c is the radian carrier frequency, and T is the symbol time.

The modulation scheme used in GSM is a Gaussian minimum-shift-keyed (GMSK) continuous-phase modulation scheme, which can be described by

$$s(t,\mathbf{a}) = A\cos(2\pi f_c t + \phi(t,\mathbf{a})) \ , \qquad (2)$$

where f_c is the carrier frequency, $\mathbf{a}$ is a binary data sequence that takes on a value of ± 1 at each instant, and $\phi(t,\mathbf{a})$ is the information-carrying phase given by

$$\phi(t,\mathbf{a}) = \pi \sum_{i=-\infty}^{+\infty} a_i q(t-iT) \ .$$

The phase response, $q(t)$, is obtained by integrating a frequency pulse $g(t)$

Panel 2. Multipath Fading

If the transmitted signal is an unmodulated carrier with frequency f_c, then the received signal is

$$r(t) = \mathrm{Re}\left\{ \sum_{k=1}^{K} \left[\beta_k(t) e^{j\theta_k(t)} \right] e^{j2\pi f_c t} \right\},$$

where $\beta_n(t)$ and $\theta_n(t)$ are the amplitude and phase of the n^{th} scatterer. The phase $\theta_k(t)$ is given by

$$\theta_{(t)} = 2\pi \frac{v}{\lambda} t + \phi_n(t)$$

where v is the velocity of the vehicle, λ is the wavelength, and $\phi_n(t)$ is a time-varying random phase. Vehicle movement causes time-varying fading whose bandwidth is determined by the maximum Doppler frequency v/λ.

The magnitude of $r(t)$ is Rayleigh distributed[2] with

$$p_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}, \quad r \geq 0$$

where σ^2 is the variance. Because the envelope is Rayleigh distributed, fading of this kind is called Rayleigh fading.

$$q(t) = \int_{-\infty}^{t} g(\tau)\, d\tau .$$

Here $g(t)$ is chosen to be a Gaussian pulse whose normalized bandwidth equals 0.30.

The major difference between the two schemes is that the IS-54 system, which uses $\pi/4$ shifted DQPSK, has an envelope that is constant only at the sampling instants, while the GMSK scheme has a constant envelope. The former scheme requires linear power amplifiers. Although the latter scheme can use more power-efficient, nonlinear, class C power amplifiers, these reduce the bandwidth efficiency by about 20 percent.

Diversity Principles

The effect of fading is normally mitigated by using diversity techniques, in which the information-bearing signal is received, ideally, along independently fading channels.

Table I. $\pi/4$-shifted DQPSK phase difference representation

X_k	Y_k	$(\Delta\Phi_k)$
1	1	$-3\pi/4$
0	1	$3\pi/4$
0	0	$\pi/4$
1	0	$-\pi/4$

In space diversity systems, the receiver uses L different antennas spaced to create independent fading channels, where L is the diversity order. An antenna separation of about a quarter of a wavelength is enough to cause almost independent fades at the receiving antennas. Depending on the antenna height, however, the base station needs a separation of 10 to 30 wavelengths.[2]

In frequency diversity systems, the same information is transmitted in L frequency bands, and in time diversity systems, the information is transmitted in L time slots, both spaced far enough apart in frequency and time, respectively, to cause independent fading. To produce diversity, the frequency and time diversity systems have thus used repetitive coding, a rudimentary form of channel coding. Each bit of the code is then transmitted either in a different frequency or time slot. Later in this paper, we discuss how to obtain diversity by using more sophisticated channel codes.

The last two techniques, frequency diversity and time diversity, may not be always feasible to implement. For example, frequency bands that are far enough apart to guarantee independent fading may not exist in an environment with limited radio spectrum. Similarly, because of delay restrictions, it may not be possible to transmit bits that are separated enough in time.

Combining Techniques

Independently Rayleigh[2] fading signals are combined to recover the information they contain, commonly using selection diversity, equal gain combining, or maximal ratio combining.

In selection diversity, the channel with the highest signal strength recovers the transmitted information. Selection diversity with two antennas, compared to a receiver with one antenna, offers about an 8- to 10-dB improvement in carrier-to-interference ratio at a bit-error rate of 10^{-3}.

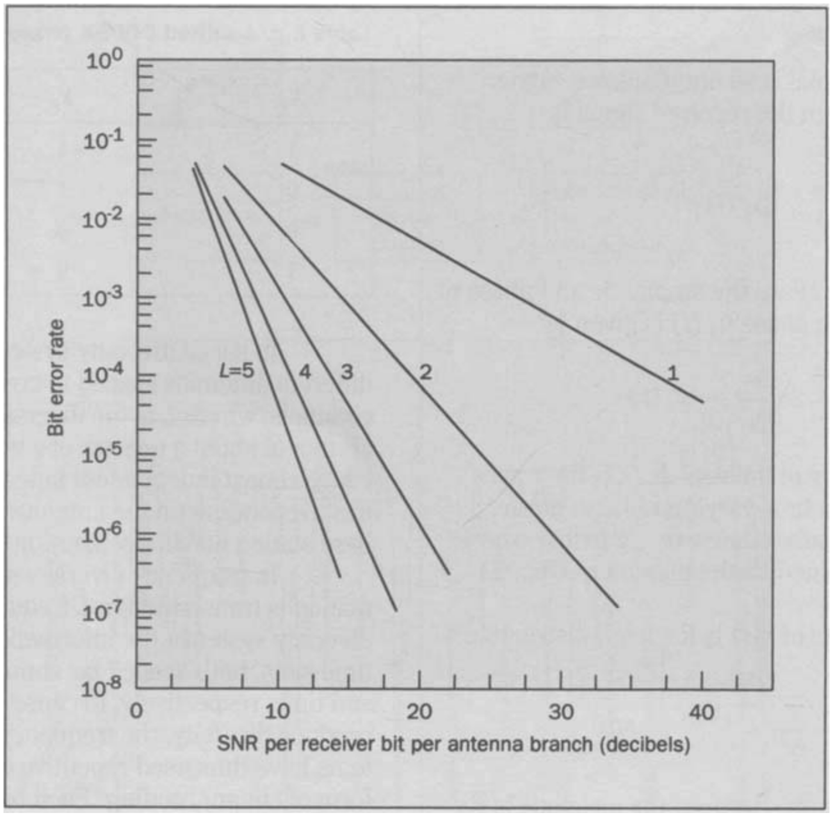

Figure 2. Bit error probability vs. SNR for various orders of (ideal) diversity with independent Rayleigh fading for binary PSK signaling.

In equal gain combining, the L received signals are co-phased and combined to retrieve information. In maximal ratio combining, the L received signals are co-phased and amplitude-weighted by the respective channel amplitudes to yield the information-bearing signal. When the interference is normally modeled as additive noise, to yield a signal-to-noise ratio of SNR, the error probability at large SNRs for a digital communications system using maximal ratio combining is given by

$$P_e \simeq c\left(\frac{1}{\mathrm{SNR}}\right)^L$$

where c is a proportionality constant. Figure 2 shows the bit-error probability behavior for different orders of diversity (L) for binary differential phase-shift keying (DPSK) using maximal ratio combining.

Multiple-Access Techniques

First-generation digital techniques in North America, Japan, and Europe use narrowband time-division multiple-access (TDMA), in which each subscriber transmits signals in a designated time slot in a time frame, where a frame consists of many slots. In Japan and North America, the available spectrum is divided into narrow frequency slots of 25 and 30 kHz bandwidth, respectively. Each frequency slot is used by three subscribers on one carrier frequency, using time-division multiplexing. In GSM, the spectrum is divided into 200-kHz-wide bands, with eight subscribers sharing one carrier frequency.

In frequency-division multiple-access (FDMA), for example, two or three subscribers divide each 30-kHz band into two or three narrower bands, using either digital or analog transmission. However, band splitting incurs significant base-station costs. Motorola has proposed a 10-kHz narrowband analog mobile phone system (AMPS) that can triple the capacity of the existing North American cellular systems.

Recently, Qualcomm, Inc.,[1] which provides communications systems and services, proposed a direct-sequence (DS) code-division multiple-access (CDMA)

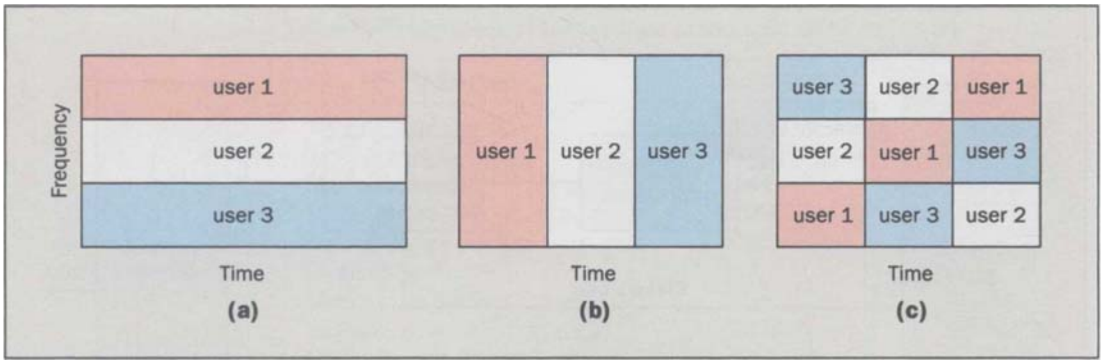

Figure 3. Time-frequency plane tiling for various access techniques. (a) FDMA, (b) TDMA, and (c) FH-CDMA. For DS-CDMA, each subscriber uses the entire frequency band at all times.

scheme that uses a 1.25-MHz bandwidth that is simultaneously shared by many subscribers. More details about this spread-spectrum system appear elsewhere in this issue.[4]

Other forms of access include frequency-hopping (FH) multiple access, where each subscriber is assigned a set of frequencies and dwells in a frequency for a certain period of time. This form of access reduces interference by assigning well-designed hopping sequences that minimize the probability of a subscriber experiencing collisions in too many frequencies. Frequency-selective fading can be avoided by limiting the time a subscriber spends in one frequency, and error-correction coding is used to maintain reliable communications. Figure 3 shows the access techniques described.

Duplexing Techniques

The commonly used duplexing schemes are frequency-division duplexing (FDD) and time-division duplexing (TDD). In FDD, the base-to-mobile (***downlink***) and mobile-to-base (***uplink***) transmit simultaneously on different frequency bands. TDD systems use the same frequency band in both directions, which requires the downlink and uplink transmissions to occur in different time slots.

Interference is managed better in FDD-based cellular systems, since an uplink transmission can only interfere with another uplink transmission operating in the same frequency. In TDD systems, a high-power base-to-mobile transmission can interfere with a low-power mobile-to-base transmission if the entire cellular system is not ***time synchronous***. Because the mobile receiver is simple, TDD systems have been proposed for cordless private branch exchange (PBX) and personal communications network (PCN) type standards.

Channel Equalization

Intersymbol interference, the result of frequency-selective fading, causes significant distortion in digital wireless communications. This type of fading occurs when the difference in delay between the various reflections arriving at the receiver is a significant fraction of a data symbol interval. Typical values of delay expressed are 1 to 4 microseconds (μs) for a macrocellular urban propagation environment,[5] and within a few hundred nanoseconds (ns) in microcells and inside large buildings.[6]

A typical model of a frequency-selective fading channel contains two reflections of the transmitted signal. Such a model produces fading over a part of the frequency band occupied by the transmitted signal.

To recover the data signal, equalization techniques are needed. The simplest of equalization techniques amplifies the attenuated part, and attenuates the amplified part, of the spectrum, thereby equalizing the channel distortion. This restores the output signal after equalization to its original spectral shape. When additional noise (interference) is present at the input to the equalizer, the equalizer must minimize the channel distortion without adversely enhancing noise. In these cases, more sophisticated equalization techniques are needed.[7,8]

Equalization methods have been applied before to equalization of very slowly varying channels. The task in mobile radio is more difficult, because channel variations are rapid. First, these problems have been alleviated by using a short training sequence—14-30 symbols every 150 to 200 data symbols—to compute the equalizer tap coefficients. Second, by using novel joint data and channel estimation algorithms, data can be reliably recovered even after the channel undergoes fading.[9]

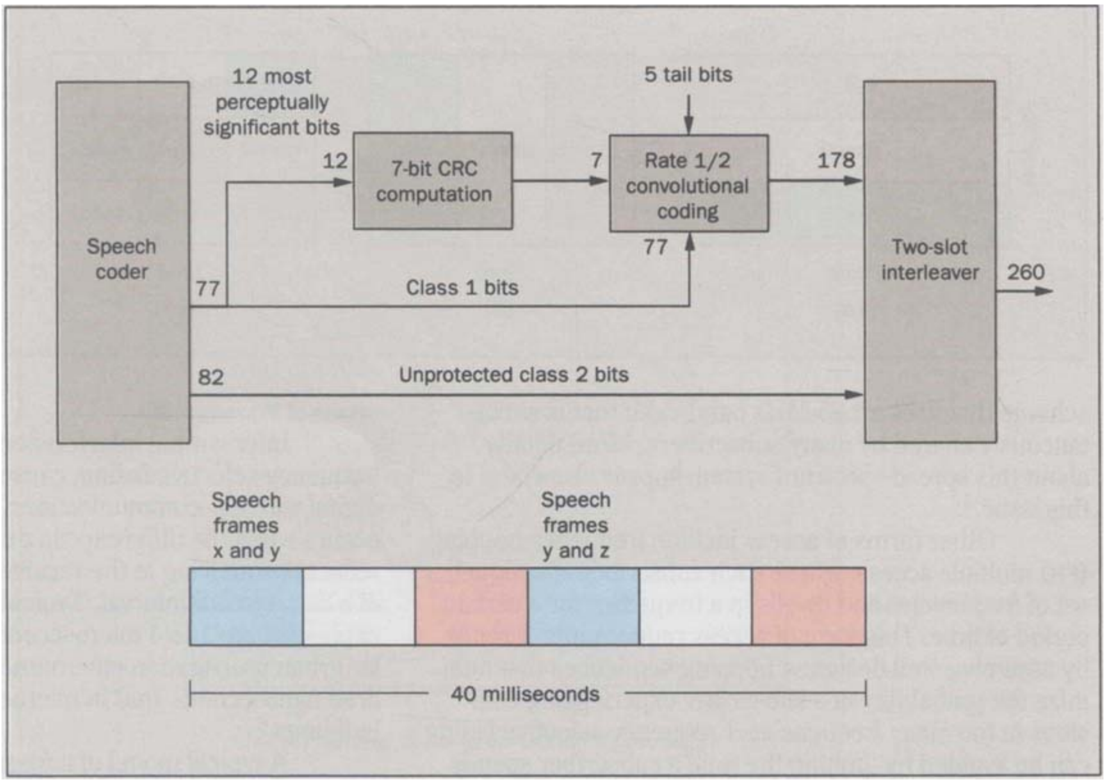

Figure 4. Channel coding for IS-54 digital speech transmission.

Both GSM and IS-54 have standardized the training sequence, but have not specified the method of equalization. The first-generation digital system in Japan (JDC) uses either two-fold space diversity or equalization.[1]

Channel Coding

All digital cellular systems use powerful channel coding to ensure reliable transmission. Conventionally, channel coding adds redundancy to information bits. These redundant bits are used at the receiver to reliably decode the information, even if the channel produces errors in part of the transmitted information.

Convolutional codes are an efficient class of channel codes into which redundancy is introduced by convolving digital information with one or more known patterns of bits. These codes can be efficiently decoded by the Viterbi algorithm,[10] using soft-decision (or analog) information. *Soft-decision information* is a measure of the reliability of data at the input to the channel decoder. This reliability information can substantially enhance the error-correction capability of the code.

Choosing a channel-coding technique for transmitting digitized speech over fading radio channels requires consideration of several system properties. Speech bits differ in their sensitivity to transmission errors, implying ***unequal error protection***.[11,12] Fading causes errors to occur in bursts. Since convolutional channel codes are efficient for correcting random errors, interleaving (to separate contiguous bits from the channel encoder output) is required (see Figure 1) at the expense of delay.

In general, the goal of interleaving is to achieve independent fading values at the input to the decoder by using a matrix interleaver. The number of rows in the matrix is determined by the expected maximum fade duration, and the number of columns is equal to the decoding depth of the code. The encoded data is written row by row and transmitted column by column. Accommodating a wide range of fade durations requires a large interleaver, which can cause substantial delay. In the IS-54 system, the interleaving delay for speech transmission is restricted to 40 ms, which corresponds to a total of two encoded speech frames of 520 bits. Reference 3 contains more information about the IS-54 interleaver.

Table II. Parameters of two cellular standards: GSM and IS-54

Parameter	GSM	IS-54
Forward band (MHz)	935-960	869-894
Reverse band (MHz)	890-915	824-849
Multiple access	TDMA	TDMA
Duplex type	FDD	FDD
Carrier spacing (kHz)	200	30
Channels per carrier	8	3
Channel rate (kb/s)	271	48.6
Modulation	GMSK	$\pi/4$-DQPSK
Modulation efficiency (b/s/Hz)	1.35	1.62
Voice rate (kb/s) (with channel coding)	22.8	13
Channel code type	Convolutional UEP	Convolutional UEP
Equalizer	Yes	Yes

When the channel decoder makes errors on more error-sensitive speech, these errors must be detected and concealed using the speech-based redundancy that exists from one speech frame to the next. This redundancy stems from a limited frame duration of 20 ms. An error-detection block code makes it easier to detect errors on these bits. The channel-coding scheme for the North American digital cellular system (IS-54) uses this type of error-detection scheme.

The IS-54 channel error-control scheme[3] for the speech coder data shown in Figure 4 uses a combination of all the error-control techniques mentioned earlier to mitigate channel errors. First, 7 cyclic redundancy check (CRC) bits are added to 12 of the most error-sensitive digitized speech bits to detect errors. Next, an error-correcting convolutional code is used to protect the 77 most error-sensitive bits of the speech coder data stream and the 7 CRC bits. The remaining 82 digitized speech bits are left unprotected. In addition, to prevent inter-frame memory and reduce decoding delay, the contents of every frame of the code's memory are wiped clean by encoding 5 known bits (*tail bits*) within the frame. In Figure 4, the 77 protected speech bits are called class 1 bits, and the 82 uncoded bits are class 2 bits. The third technique interleaves the transmitted data for each speech coder frame over two time slots (40 ms) to randomize the effects of Rayleigh fading.

The order of time diversity achieved in the presence of an ideal interleaver is 8. This number is equal to the minimum number of positions (in bits) between any two distinct coded sequences. It is also called the *free distance*, or minimum Hamming distance, of the code. At the receiver, the output of the equalizer is first de-interleaved to restore the bits to the original order. Any analog information used by the channel decoder should also be de-interleaved. References 12 and 13 describe methods of generating and decoding analog information.

The error-detecting (19,12) CRC block code checks for errors on the 12 most error-sensitive bits. If an error is detected, a "bad frame masking" algorithm is initiated. This repeats, or repeats and attenuates, the previous speech frame. Reference 3 describes this process in detail.

Note that the decoder for speech transmission operates on terminated blocks with the overhead of tail bits. For data transmission in the so-called slow associated control channel, a rate $R = 1/2$ convolutional code, with continuous Viterbi decoding, is used. For the so-called fast associated control channel, a rate $R = 1/4$ convolutional code occasionally injects data blocks into the speech channel. Here, the starting and ending states are the same, but the decoder estimates that state as a part of decoding. No overhead is required for transmitting known tail bits. Such "tail-biting" complicates the Viterbi decoding process. Reference 14 presents and evaluates a novel technique, using a "circular Viterbi algorithm" concept, to efficiently decode convolutionally encoded data blocks without a known tail, i.e., so-called tail-biting convolutional codes.

Channel coding, as described in this section, reduces the bandwidth efficiency below 2 bits/symbol, which is the efficiency of the 4-DPSK modulation scheme. Bandwidth becomes more efficient and transmission more reliable by combining coding with modulation schemes such as 8-DPSK. We describe this in "Examples of Future Techniques," later in this paper.

Channel coding techniques similar to the one described earlier for the IS-54 system are also used for the pan-European system. In general, robust speech coding strategies for cellular radio are computationally intensive and use low-bit-rate speech coding, channel coding, and equalization techniques. A low-cost, high-quality approach to wireless speech communications is proposed in Reference 15. Such a technique might be suitable for PCN applications. The method uses low-delay differential pulse code modulation (DPCM) and matched modulation with built-in unequal error-protection (UEP) properties for low-delay PCN applications. This technique, called pseudo-analog speech transmission, quantizes, protects against errors, and transmits spectral information using a digital modulation scheme. The residual signal is transmitted using either multilevel matched digital modulation schemes or analog modulation schemes.

Summary of Digital Wireless Systems Standards

The first digital wireless systems, now being introduced commercially in Europe, North America, and Japan, are designed according to several incompatible standards.[1,16,17] Digital cellular systems will increase capacity in North America and Japan without adding new spectrum allocations. These digital cellular systems also provide the first pan-European cellular system.

GSM, the selected European standard (see Table II), is a medium-rate TDMA system with constant-amplitude Gaussian minimum-shift-keying (GMSK) modulation and convolutional channel coding with some UEP.[16]

IS-54, the digital cellular standard for North America (see Table II), is frequency compatible with the first-generation analog cellular system. Its low-rate TDMA system has linear modulation DQPSK and convolutional channel coding, with some UEP. In Japan, the first JDC is similar to IS-54, with a channel carrier spacing of 25 kHz, instead of 30 kHz, leading to proportionally lower bit rates. The mobile receiver uses either equalization or two-fold space diversity. Work is currently under way to establish standards for the so-called half-rate GSM, half-rate IS-54, and half-rate JDC. These systems will increase the capacity by establishing two digital voice channels where one currently exists.

Table III. Parameters of two PCN/cordless standards: CT2 and DECT

Parameter	CT2	DECT
Forward band (MHz)	864-868	1880-1900
Reverse band (MHz)	864-868	1880-1900
Multiple access	FDMA	TDMA
Duplex	TDD	TDD
Carrier spacing (kHz)	100	1728
Channels per carrier	1	12
Channel rate (kb/s)	72	1152
Modulation	FSK	GMSK
Voice rate (kb/s)	32	32

In addition to the IS-54 standard, a direct-sequence spread-spectrum (DS-SS) CDMA technology has been introduced for cellular communications.[1] This digital standard uses 1.25-MHz bandwidth channels, advanced power control, speech activity detection (SAD), and channel coding to increase the system capacity to almost three times that of IS-54. In this issue, T. P. Bursh, Jr., et al.[4] describe the system design, system performance, and field test results of CDMA digital cellular technology.

Table III shows parameters for two European digital cordless telephony standards, cordless telephone 2 (CT2) and digital European cordless telephone (DECT). These are developed for wireless PBX, digital cordless telephone, and applications of the personal communications network (PCN) and personal communications system (PCS). Both CT2 and DECT use TDD. DS-CDMA systems are also being developed in the U.S. for PCN and PCS applications.[1,16]

Other digital wireless systems currently being developed are wireless ethernet and other data network systems.[18] Numerous systems have been proposed in which low-earth-orbit (LEO) satellites would establish wide-ranging (up to global) coverage for wireless voice and data services.[19]

Considerations are now under way for future "second-generation digital" wireless (third-generation wireless) system standards, which will enable people to communicate with each other anywhere in the world at any time with a small, lightweight, pocket telephone. These systems should be able to carry voice, video, image, and data.

Examples of Future Techniques

This section briefly describes some of the techniques that can improve wireless systems in the future. Other important techniques not described here are SAD[20-22] and dynamic channel allocation (DCA),[23] both of which increase system capacity.

Minimum-Mean-Squared-Error Diversity Combining. Co-channel interference, a significant source of impairment in cellular systems, is caused by other interferers transmitting simultaneously in the same frequency band as the desired signal.[24] Normally, this interference is treated as background noise, and channel coding protects the data. However, channel coding is not effective when this data-like interference becomes comparable in power to the desired signal. In existing cellular systems, this is handled by limiting the availability of a frequency band to geographically well separated areas (called *frequency reuse*). One way to overcome a significant amount of interference, and thus increase frequency reuse, is through minimum-mean-squared-error (MMSE) antenna diversity combining. Here, signals received by more than one antenna at the mobile unit are linearly combined. The parameters of the linear combiner are chosen to eliminate interference. (See Panel 3 for a mathematical derivation of MMSE combining.)

Panel 3. MMSE Antenna Diversity Combining

MMSE antenna diversity combining overcomes a significant portion of interference. To explain this, consider a receiver with two-branch antenna diversity. At the n^{th} sampling instant corresponding to the n^{th} transmitted data symbol, the output of the first antenna is

$$r_{n1} = a_1 d_n + b_1 I_n + W_{n1} ,$$

where d_n and I_n are the desired and interfering signals at time n. The quantities a_1 and b_1 are complex channel amplitudes, where a_1 is the channel amplitude corresponding to the channel from the desired transmitter to the receiver, and b_1 is the amplitude for the channel from the interfering transmitter to the receiver. Similarly, at antenna 2, we have

$$r_{n2} = a_2 d_n + b_2 I_n + W_{n2} .$$

Here W_{n1} and W_{n2} are additive background disturbances that model nondominant co-channel and adjacent channel interferers and other sources of noise.

During a training phase, the transmitter sends a known data sequence $(d_1, \ldots, d_J)$, and the receiver linearly combines the antenna outputs to minimize the squared error between the combiner output and the known data sequence. Mathematically, this is equivalent to finding complex combiner coefficients x,y such that

$$\sum_{i=1}^{J} |xr_{i1} + yr_{i2} - d_i|^2$$

is minimized. When training is completed, the linear combiner with coefficients x and y is used for data demodulation.

It has been shown that with L branches of diversity, up to M co-channel interferers can be suppressed $(M<L)$ and a diversity benefit of $L-M$ can be provided.[23]

Diversity Using Multiple Transmit Antennas. Cost considerations and limitations in technology make it difficult to support more than one antenna at the portable receiver. In ideal situations, multiple antennas at the base station could provide diversity to the receiver. In contrast to receiver-based diversity, this is called transmit diversity. With L receive antennas and M transmit antennas, and paths that fade independently, diversity of order ML can be obtained at the receiver.

Transmit diversity techniques can be classified into three broad categories. The first category uses feedback information from the receiver to configure, or *train*, the transmitter. (Training configures the system to obtain diversity at the receiver.) The second category uses feed-forward information from the transmitter to train the receiver. The third category uses neither feedback nor feed-forward information; instead, it uses multiple transmit antennas and channel coding to provide diversity. Diversity available in TDD systems[25] and switched diversity in FDD systems[2,26] are examples of the first category.

In the second category of transmit diversity techniques, the signal is prefiltered at the base station and transmitted to the portable receiver using multiple transmit antennas.[27,28] With the aid of the feed-forward information from the transmitter, the receiver post-filters the signal to obtain diversity. For example, pre-filtering at the base station transmits delayed versions of the same signal using multiple antennas. At the receiver, the

Table IV. Simple coded modulation example

Information bits	Code word: Code symbol 1 c_1	Code word: Code symbol 2 c_2
000	0	0
001	1	5
010	2	2
011	3	7
100	4	4
101	5	1
110	6	6
111	7	3

signal is distorted by intersymbol interference, making the channel resemble one subject to frequency-selective fading. In a frequency-selective fading environment, an equalizer (post-filter) at the receiver provides the benefits of diversity. The feed-forward information is used to obtain the post-filter at the receiver.

The third category of transmit diversity technique is useful only in very slowly fading channels. Here, neither the transmitter nor the receiver needs information about the channel. For example, the variations of a slowly fading channel can be represented by a 2-mph mobile radio channel and a carrier frequency of 900 MHz. If the transmitted data symbol rate is 8 ksymbols/s, then the number of successive data symbols whose signal strength is below the root-mean-square (RMS) level of the signal, on average, is approximately 2000. If data symbols whose signal strength is 3 dB below the RMS level are considered, then the number is reduced to 1100 symbols. Randomizing large numbers of weak-signal data symbols together with strong-signal data symbols requires a very large interleaver/de-interleaver pair. This, in turn, incurs an intolerable delay. Multiple transmit antennas can simulate a fast fading effect on the very slowly fading channel.[29,30] Because channel codes reduce the effects of fast fading, they can be used to provide diversity to the receiver.

Consider two antennas, T_1 and T_2, transmitting the same signal simultaneously in a slowly fading Rayleigh channel. Two types of deep fades occur with two transmit antennas. In the first type, the received signals transmitted from the two antennas are weak. In the second type, both signals are strong and almost equal, and their phases are opposite to each other. While not much can be done about the first type, we can alleviate the fade caused by the second type. Suppose very small, time-varying phase offsets $\theta_1(t)$ and $\theta_2(t)$ are introduced at transmit antennas T_1 and T_2, respectively. This slowly rotates the signals received from the two antennas. If $\theta_1(t)$ and $\theta_2(t)$ are independent, the signals will cancel each other only for a small fraction of time. If we use a channel code with a suitable interleaver/de-interleaver pair, this technique can eliminate the deep fades that fall into the second type. Increasing the number of transmit antennas can substantially reduce weakness in all received signals. Introducing time-varying phase offsets will reduce the average signal strength at locations where the received signals were of the same phase before this new technique was introduced. While this new scheme will increase the bit error rate at locations where the received signals were co-phased, it will also reduce the bit error rate at a deep fade by a substantial amount, as well as the error bursts typical in fading channels. As a result, the average bit error rate is significantly improved.

Figure 5 illustrates a general scheme for inducing channel variations. In this figure, there are M transmit antennas. The signal at the i^{th} transmit antenna branch is prefiltered by $p_i(t)$, and then weighted by the amplitude $A_i(t)$, and phase $\theta_i(t)$.

The filter $p_i(t)$ can be chosen to obtain the second category of transmit diversity techniques, in which the signal is prefiltered at the base station and transmitted to a portable receiver by multiple transmit antennas. The channels can be varied by changing the phase or amplitude of these weights. Large frequency offsets to the signals at the transmit antennas will improve the randomization of fading at the input to the channel decoder, but will also increase the transmission bandwidth. This may cause difficulties in the demodulation process. Typically, frequency offsets in the range of 1 to 2 percent of the data rate are desirable for an indoor radio channel.

The amplitudes of the weights at the antennas can be varied to obtain diversity at the receiver. If two transmit antennas are used with a rate one-half repetition code, one antenna can transmit the first symbol in the block code, and the other antenna can transmit the next coded symbol. The resulting independent fades on coded symbols in the block code yield optimum performance of a block code in a Rayleigh fading channel. Reference 27 details such a scheme. Transmitting the signal from both antennas, with the phase of the weight at one transmit

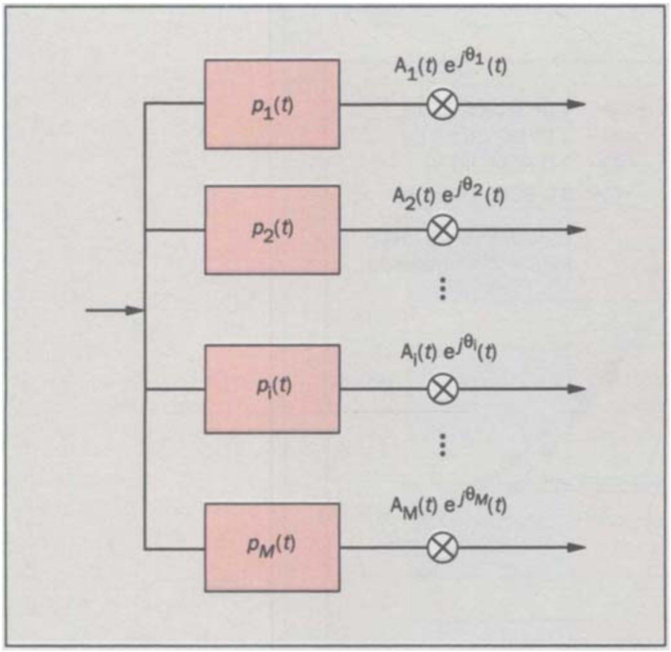

Figure 5. A general scheme for inducing channel variations. $A_i(t)$ and $\theta_i(t)$ are the amplitude and phase of the complex weight at the i^{th} transmit antenna, and $p_i(t)$ is the pre-filter in this antenna branch.

antenna being changed by 180 degrees at every symbol interval, can produce similar results.

This third category of transmit diversity techniques, useful only in very slowly fading channels where time diversity is not available, is simple to execute. It also improves the performance of the channel code by effectively randomizing the fade values at the input of the channel decoder. These techniques can be used with DS CDMA systems as well.[31]

Coded Modulation. In the section on channel coding, we saw how interleaved $R = 1/2$ channel coding, combined with $\pi/4$ shifted 4-DPSK (2 bits/symbol) modulation, can provide time diversity. The net bandwidth efficiency of this scheme is about 1 bit/symbol (bandwidth efficiency of uncoded modem × rate of the channel code) for the coded speech bits. The unprotected bits are transmitted at 2 bits/symbol without time diversity. To operate at 2 bits/symbol net efficiency and still provide time diversity, we must use more than four levels for modulation. For example, we can use eight levels of (3 bits/symbol) modulation combined with $R = 2/3$ channel coding. An example of eight-level modulation is 8-PSK with phase values $k\pi/4$, $k = 0, \ldots, 7$. In every symbol period, three bits are mapped to a symbol that takes values (k) from 0 through 7. The symbol k then selects one of eight phase values for modulating the carrier. The main code design criterion for fading channels is to maximize the minimum Hamming distance between any two code words, where each symbol of a code word is an 8-DPSK or (8-PSK) symbol. With adequate interleaving, we can show that the built-in number of time diversity branches is the minimum Hamming distance of the coded modulation system.

A channel code designed to maximize the minimum Hamming distance at a bit level cannot also be used directly to maximize the minimum Hamming distance after the channel code output has been mapped to a symbol. For example, in an 8-PSK constellation, symbols 0 and 7 are addressed by bits 000 and 111, which differ in three bit positions, but whose distance between symbols is only 1. The integration of channel coding with the signal constellation is called coded modulation.[32]

We offer a simple example of coded modulation design using the 8-PSK constellation, with a rate of only 1.5 bits/symbol. Here, three information bits address a code of length two 8-PSK symbols. This code, which contains eight code words, is shown in Table IV. Any two code words differ in exactly two symbol positions. The code is also designed to maximize the minimum Euclidean distance between any two different coded modulation sequences. This improves the error performance on a non-fading channel, and keeps SNRs on the fading channel low. Every three consecutive information bits select a code word, which, combined, are arranged in an $N\times 2$ interleaver, where N is the number of rows (interleaving depth), and the code word length equals the number of columns. The symbols are entered along rows and transmitted along columns. After they are deinterleaved, the two received symbols in each row (each symbol is complex and each component of the complex number is real) are correlated with each of the eight code words. The code word with the highest correlation determines the decoded information bits (maximum likelihood decoding). In practice, the symbols will be transmitted differentially. Further, a shifting of $\pi/8$ will be used from one symbol to the next to make the scheme a somewhat constant envelope ($\pi/8$-8DPSK) modulation.

Although the details are not presented, we can

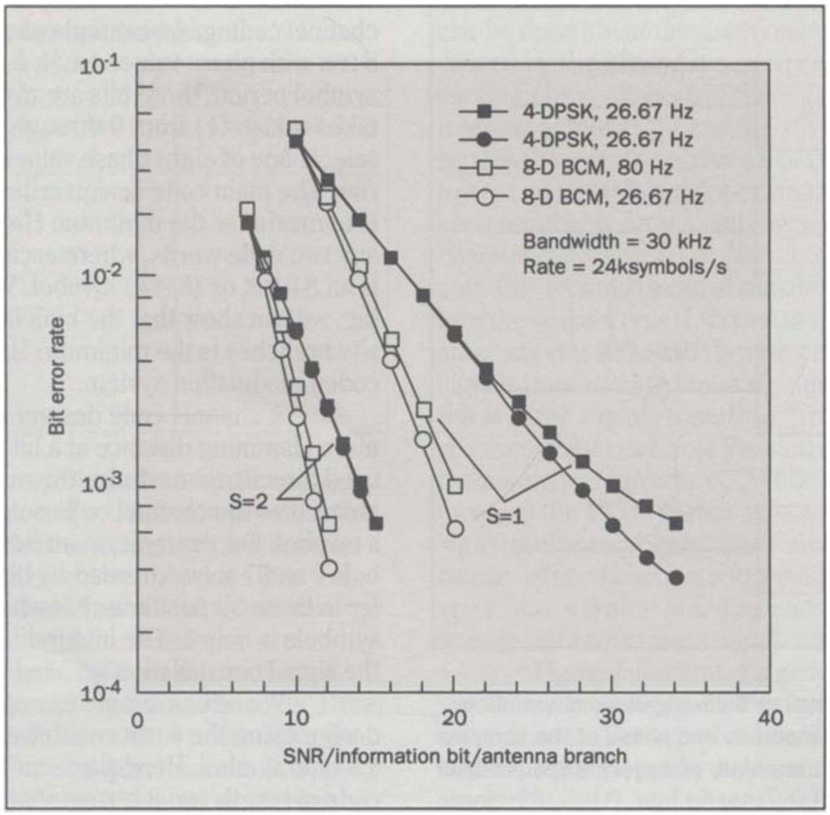

Figure 6. Simulated bit error rate for 2 bits/symbol coded modulation of length 4 symbols. The transmission rate is 24 ksymbols/s, and *S* is the number of branches of space diversity. The carrier frequency is 900 MHz, and the Doppler bandwidths are 26.67 and 80 Hz.

use the principles illustrated to obtain a code of length 4 symbols, rate $R = 2$ bits/symbol with time diversity 2 using $\pi/8$-8DPSK modulation. Figure 6 shows the performance of this code. Gains of more than 10 dB are achieved over the uncoded $\pi/4$ shifted DPSK system for error rates below 10^{-3}. Figure 6 also shows how these schemes perform in the presence of receiver space diversity. The parameter S refers to the number of branches of space diversity. At larger error rates, these gains are smaller. These codes can be constructed[33-35] to provide unequal error protection and to double the capacity of cellular systems. Work is in progress to double the capacity of the North American cellular system (IS-54) by halving the total speech and channel coding rate from 13 kb/s to about 6 kb/s per subscriber.

If the modulation were fixed at 4 DPSK, it would be possible to use 4 kb/s for speech coding and 2 kb/s for channel coding. On the other hand, if the modulation were changed to 8-DPSK, then 6 kb/s could be used for speech coding, and an appropriate coded modulation, such as the one described earlier, could be used to provide reliable communications. The main benefit is higher speech quality, which is extremely important if wireless systems are to compete with wireline telephony. For moderate and high SNRs (above 15 dB, for example), these schemes improve the error rate significantly, without expanding bandwidth. To produce a comparable improvement at low SNRs, the bandwidth must be expanded. Here, conventional error control with coding and four-level modulation is adequate.

Advanced Decoding Techniques. Many wireless communications systems are, in a broad sense, concatenated communications systems (e.g., IS-54 and GSM). Figure 7 shows five such systems. A speech coder, followed by channel coding in example 2 of this figure, falls into this category. Normally, the inner decoder 1 decodes the symbols received from the channel. The decoded output symbols are then processed by the second decoder.

In example 2, the inner decoder is a Viterbi decoder, and the outer decoder is a speech decoder. System performance can be improved if the Viterbi decoder is modified to show whether its decoded output is

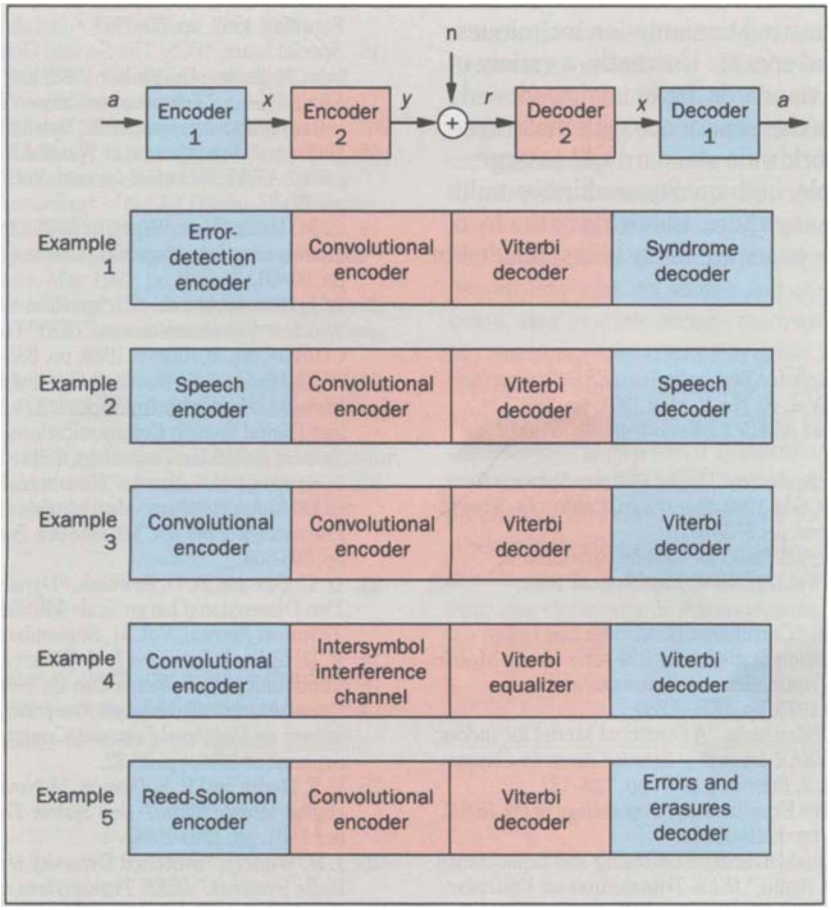

Figure 7. Examples of communications systems with concatenated decoders. Examples 1 and 2 represent LVA, and examples 3 through 5 represent SOVA.

reliable. If it is unreliable, the speech decoder can decide whether to accept or reject it by using interframe speech redundancy. If the decoded output is rejected, interframe redundancy can be used to conceal errors. In general, the soft output Viterbi algorithm (SOVA)[13] can test the reliability of the decoded output at the decoded bit or symbol level.

The list Viterbi algorithm (LVA)[36] is another generalization of the Viterbi algorithm (GVA). Here, the inner decoder produces a rank ordered list of the $\Lambda > 1$ best estimates of the ***sequence*** encoded by the inner coder. The outer decoder can then decide which estimate should be accepted as the correct candidate. To illustrate this concept, consider example 1 of Figure 7. Each of the Λ candidates, according to its rank, is checked for an error by the outer error-detection code. The first one with no errors detected is accepted as the decoded candidate. We can also use this algorithm when the outer decoder is a speech decoder. The interframe speech redundancy selects the best of the Λ estimates for speech decoding. Error-detection information from the Viterbi decoder can be supplied by comparing how close the metrics of the two best candidates are to a preset threshold.[36]

Combined or separately, LVA and SOVA can provide an extra 2 to 3 dB of interference tolerance, which increases the capacity or quality. In Figure 7, examples 1 and 2 demonstrate how LVA is useful, and examples 3 through 5 indicate where SOVA is useful. Examples 3 and 5 are classic, concatenated coding schemes; in example 4, the inner "coder" is an intersymbol interference channel, also present in IS-54 and GSM. For further details on SOVA and LVA schemes, see References 13, 36, 37, and 38.

Discussion and Conclusions

In this brief overview of current and future techniques for reliable digital transmission in wireless systems, we have emphasized transmission techniques for transporting digitized speech. Currently, a variety of digital wireless system standards are being introduced around the world. From the experience with these systems, we hope that a worldwide standard will emerge, one that will offer flexible, high-quality, multirate, multimedia service anytime, anywhere. Elements of the techniques described in this paper are likely to be included in such a system.

References

1. Special Issue on Digital Cellular Technologies, *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, May 1991, pp. 289–374.
2. W. C. Jakes, Jr., *Microwave Mobile Communications*, Wiley, New York, 1974.
3. C-E. W. Sundberg and N. Seshadri, "Digital Cellular Systems for North America," GLOBECOM 1990, San Diego, California, November 1990, *Conference Record*, pp. 533–537.
4. T. P. Bursh, Jr., et al., "Digital Radio for Mobile Applications," *AT&T Technical Journal*, Vol. 72, No. 4, July/August 1993, pp. 19–26.
5. D. C. Cox and R. P. Leede, "Correlation Bandwidth and Delay Spread Multipath Propagation Statistics for 910 MHz Urban Mobile Radio Channels," *IEEE Transactions on Communications*, Vol. COM-23, November 1975, pp. 1271–1280.
6. A. A. M. Saleh and R. A. Valenzuela, "A Statistical Model for Indoor Multipath Propagation," *IEEE Journal of Selected Areas in Communications*, Vol. SAC-5, No. 2, February 1987, pp. 128–137.
7. S. U. H. Qureshi, "Adaptive Equalization," *Proceedings of the IEEE*, Vol. 53, September 1985, pp. 1349–1387.
8. P. Balaban and J. Salz, "Dual Diversity Combining and Equalization in Digital Cellular Mobile Radio," *IEEE Transactions on Vehicular Technology*, Vol. 40, No. 2, May 1991, pp. . 342-354.
9. N. Seshadri, "Joint Data and Channel Estimation Using Blind Trellis Search Techniques," to be published in the *IEEE Transactions on Communications*, 1994.
10. G. D. Forney, Jr., "The Viterbi Algorithm," *Proceedings of the IEEE*, Vol. 61, No. 3, March 1973, pp. 268–278.
11. R. V. Cox, J. Hagenauer, N. Seshadri, and C-E. W. Sundberg, "Subband Speech Coding and Matched Convolutional Channel Coding for Mobile Radio Channels," *IEEE Transactions of Signal Processing*, Vol. 39, No. 8, August 1991, pp. 1717–1731.
12. J. Hagenauer, N. Seshadri, and C-E. W. Sundberg, "The Performance of Rate-Compatible Punctured Convolutional Codes for Digital Mobile Radio," *IEEE Transactions on Communications*, Vol. 38, No. 7, July 1990, pp. 966–980.
13. J. Hagenauer and P. Hoeher, "A Viterbi Algorithm with Soft-Decision Outputs and Its Applications," GLOBECOM 1989, Dallas, Texas, November 1989, *Conference Record*, pp. 1680–1686.
14. R. V. Cox and C-E. W. Sundberg, "An Efficient Adaptive Circular Viterbi Algorithm for Decoding Generalized Tailbiting Convolutional Codes," *Proceeding of the IEEE Vehicular Technology Conference*, Secaucus, New Jersey, May 1993, pp. 104–107.
15. T. Miki, C-E. W. Sundberg, and N. Seshadri, "Pseudo-Analog Speech Transmission in Mobile Radio Communication Systems," *IEEE Transactions on Vehicular Technology*, Vol. 42, No. 1, February 1993, pp. 69–77.
16. Special issue, "PCS: The Second Generation," *IEEE Communications Magazine*, December 1992, Vol. 30, No. 12, pp. 32–136.
17. Special issue, "Telecommunications Regulation," *IEEE Communications Magazine*, June 1992, Vol. 30, No. 6.
18. B. Tuch, "Development of WaveLAN®, an ISM Band Wireless LAN," *AT&T Technical Journal*, Vol. 72, No. 4, July/August 1993, pp. 27–37.
19. C. M. Rush, "How WARC'92 Will Affect Mobile Services," *IEEE Communications Magazine*, October 1992, Vol. 30, No. 10, pp. 90–96.
20. D. J. Goodman et al., "Packet Reservation Multiple Access for Local Wireless Communications," *IEEE Transactions on Communication*, COM-37, No. 8, August 1989, pp. 885-890.
21. W. C. Wong, C-E. W. Sundberg, and N. Seshadri, "Shared Time-Division Duplexing: An Approach to Low-Delay, High-Quality Wireless Digital Speech Communications," to appear in *IEEE Transactions on Vehicular Technology*, 1994.
22. N. Amitay and S. Nanda, "Resource Auction Multiple Access (RAMA) for Statistical Multiplexing of Speech in Wireless PCS," *Proceedings of the ICC'93*, Geneva, Switzerland, May 1993, pp. 605-609.
23. D. C. Cox and D. O. Reudink, "Dynamic Channel Assignment in Two-Dimensional Large-Scale Mobile Radio Systems," *Bell System Technical Journal*, Vol. 51, September 1972, pp. 1611-1630.
24. R. D. Gitlin, J. Salz, and J. H. Winters, "The Capacity of Wireless Communication Systems Can Be Substantially Increased by the Use of Antenna Diversity," *Proceedings of the 1st International Conference on Universal Personal Communications*, Dallas, Texas, September 1992, pp. 28-32.
25. P. S. Henry and B. S. Glance, "A New Approach to High-Capacity Digital Mobile Radio," *Bell System Technical Journal*, Vol. 60, October 1981, pp. 1891-1904.
26. J. H. Winters, "Switched Diversity with Feedback for DPSK Mobile Radio Systems," *IEEE Transactions on Vehicular Technology*, Vol. VT-32, Feb. 1983, pp. 134–150.
27. N. Seshadri and J. H. Winters, "Two Signaling Schemes for Improving the Error Performance of Frequency-Division-Duplex (FDD) Transmission Systems Using Transmitter Antenna Diversity," *Proceedings of the IEEE Vehicular Technology Conference*, May 1993, Secaucus, New Jersey, pp. 508–511.
28. A. Wittneben, "Base Station Modulation Diversity for Digital SIMULCAST," *Proceedings of the IEEE Vehicular Technology Conference*, St. Louis, Missouri, May 1991, pp. 848–853.
29. A. Hiroike, F. Adachi, and N. Nakajima, "Combined Effects of Phase Sweeping Transmitter Diversity and Channel Coding," *IEEE Transactions on Vehicular Technology*, Vol. 41, May 1992, pp. 170–176.
30. V. Weerackody, "Characteristics of a Simulated Fast Fading Indoor Radio Channel," *Proceedings of the IEEE Vehicular Technology Conference*, May 1993, Secaucus, New Jersey, pp. 231–235.
31. V. Weerackody, "Diversity for the Direct-Sequence Spread Spectrum System Using Multiple Transmit Antennas," *ICC'93 Proceedings*, Geneva, Switzerland, May 1993, pp. 1775–1779.
32. E. Biglieri et al., *Introduction to Trellis-Coded Modulation with Applications*, MacMillan, New York, 1991.
33. C-E. W. Sundberg and N. Seshadri, "Coded Modulations for Fading Channels: An Overview." Invited Paper, *European Transactions on Telecommunications and Related Technologies*, special issue on

Applications of Coded Modulation Techniques, May 1993, pp. 309–324.
34. N. Seshadri and C-E. W. Sundberg, "Multi-Level Trellis Coded Modulations with Large Time Diversity for the Rayleigh Fading Channel," *IEEE Transactions on Communications*, September 1993.
35. N. Seshadri and C-E. W. Sundberg, "Coded Modulation with Time Diversity, Unequal Error Protection and Low Delay for the Rayleigh Fading Channel," *Proceedings of the 1st International Conference on Universal Personal Communications*, ICUPC'92, Dallas, Texas, September 1992, pp. 283–287. Also see *European Transactions on Telecommunications*, May 1993, pp. 325–334.
36. N. Seshadri and C-E. W. Sundberg, "List Viterbi Decoding Algorithms with Applications," *IEEE Transactions on Communications*, 1994.
37. P. Hoeher and N. Seshadri, "On Post-Decision Symbol Reliability Generation," *Proceedings of ICC'93*, May 1993, Geneva, Switzerland, pp. 741–745.
38. C. Nill and C-E. W. Sundberg, "List Output and Soft Symbol Output Viterbi Algorithms: Extensions and Connections," *Proceedings of the 1993 IEEE International Symposium on Information Theory*, San Antonio, Texas, January 1993, p. 22.

(Manuscript approved August 1993)

Nambi Seshadri *is a member of technical staff in the Signal Processing Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey, where he works on various aspects of signal processing and communications systems. Mr. Seshadri joined AT&T in 1986, after receiving a B.S. from the University of Madras, India, and an M.S. and Ph.D. from Rensselaer Polytechnic Institute, Troy, New York, all in electrical engineering.*

Carl-Erik W. Sundberg *is a distinguished member of technical staff in the Signal Processing Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey. His responsibilities include work on source and channel coding, digital modulation, and multiple-access techniques for wireless systems. Mr. Sundberg joined AT&T in 1984 after receiving an M.S. and Ph.D. from the University of Lund, Sweden.*

Vijitha Weerackody *is a member of technical staff in the Signal Processing Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey, where he works on signal processing and communications for wireless systems. Mr. Weerackody joined AT&T in 1990, after receiving a B.S. from the University of Moratuwa, Sri Lanka, and an M.S. and Ph.D. from the University of Pennsylvania, Philadelphia, all in electrical engineering.*